

**PROCEEDINGS OF**  
**The Second SIMS EUROSIM Conference**  
**on Simulation and Modelling**  
**SIMS EUROSIM 2024**

**The 65<sup>th</sup> SIMS Conference**  
**on Simulation and Modelling**  
**SIMS 2024**



**SIMS** Scandinavian  
Simulation  
Society



Editors: Esko Juuso, Jari Ruuska, Gaurav Mirlekar, and Lars Eriksson

Organized by University of Oulu  
Finnish Simulation Forum (FinSim)  
and  
Finnish Society of Automation

Co-sponsored by



**Proceedings of  
The Second SIMS EUROSIM Conference  
on Simulation and Modelling  
SIMS EUROSIM 2024**

**The 65th SIMS Conference on Simulation and Modelling  
SIMS 2024**

**Oulu, Finland, 11-12 September 2024**

**Editors:**

Esko Juuso, Jari Ruuska, Gaurav Mirlekar, and Lars Eriksson

**Published by:**

Scandinavian Simulation Society and Linköping University Electronic Press

ISBN: 978-91-8075-984-7

Series: Linköping Electronic Conference Proceedings, No. 212

ISSN: 1650-3686

eISSN: 1650-3740

DOI: <https://doi.org/10.3384/ecp212>

The article DOI:s are mounted as 10.3384/ecp212.XXX where XXX is the article id:s 001, 002 etc.

**Organized by:**

University of Oulu

Finnish Simulation Forum (FinSim)

and

Finnish Society of Automation

**in cooperation with:**

Scandinavian Simulation Society (SIMS)

**co-sponsored by:**

International Federation of Automatic Control (IFAC)

© 2024 The author(s) All articles are published under CC BY 4.0 license

<https://creativecommons.org/licenses/by/4.0/>



# Preface

The Second SIMS EUROSIM conference on Simulation and Modelling (SIMS EUROSIM 2024) and 65th SIMS conference on Simulation and Modelling (SIMS 2024) were organized in Oulu, Finland. The background of this conference series is in the 65-years history of Scandinavian Simulation Society, SIMS. The first SIMS EUROSIM Conference was organized as a virtual conference in 2021. The SIMS conferences are annual and every third of them is a joint SIMS EUROSIM conference.

The Scandinavian Simulation Society consists of members from five Nordic countries: Denmark, Finland, Norway, Sweden, and Iceland. The goal of SIMS is to further develop the science and practice of modelling and simulation in all application areas and to be a forum for information interchange between professionals and non-professionals in the Nordic countries. SIMS is a member society of The Federation of European Simulation Societies (EUROSIM) was set up in 1989. The purpose of EUROSIM is to provide a European forum for regional and national simulation societies to promote the advancement of modelling and simulation in industry, research and development. EUROSIM consists of 17 European Simulation Societies. The Scandinavian Simulation Society (SIMS) had the Annual Meetings during the conference. International Federation of Automatic Control (IFAC) co-sponsored the conference via technical committees:

- TC 3.2. Computational Intelligence in Control
- TC 6.1. Chemical Process Control
- TC 6.2. Mining, Mineral and Metal Processing
- TC 6.3. Power and Energy Systems
- TC 6.4. Fault Detection, Supervision & Safety of Techn. Processes-SAFEPROCESS

The conference program consisted of keynote presentations, regular presentations and a panel discussion. The call for papers resulted in 98 submissions prepared by 337 authors from 25 countries. Submissions were reviewed by six chairs, 26 IPC members and 21 international reviewers. Full articles were selected on the grounds of academic merit and relevance to the conference theme. Each submission had 2-5 reviews and the acceptance rate was 67% for the full articles. The proceedings include 66 articles prepared by 200 authors from 15 countries. The keynotes are included as abstracts. The IFAC conference templates were used for the regular papers.

The conference covered broad aspects of simulation, modelling and optimization in engineering applications. The most active area, energy systems, includes many articles on renewable energy, energy storage and power production. High number of contributions are for the circular economy, where the focus is the CO<sub>2</sub> capture and use, in the steel industry, where articles are in thermodynamics and computational methods, and in the automation where advanced process control, parameter estimation, process intensification and optimization are used in various application areas, including also mobile robots, warehouse optimization and oil recovery. Other session topics include water treatment and nature-based solutions, biosystems, fluid flow and heat transfer, transportation, machine learning, and modelling in process analysis and optimization.

Panel discussion was organized on future challenges and possibilities for simulation. The discussion focused on four areas: artificial intelligence, importance of the expert or domain knowledge, requirements of data and use of written or spoken expert data. Discussions emphasized using different forms of expertise together with the data-driven approaches. The audience had time for two questions. The limitations on relying only on expert knowledge was the question. This risk can be avoided by following mixtures of ideas also out of box. This answered the question of why to participate to the next SIMS? SIMS continues to focus on different industrial and environmental applications, modelling and simulation tools and provides strong support for PhD students continue for stimulating process development for model-based automation.

We would like to express our sincere thanks to the keynote speakers, authors, session chairs, members of the program committee and additional reviewers who made this conference such an outstanding success. Finally, we hope that you will find the proceedings to be a valuable resource in your professional, research, and educational activities whether you are a student, academic researcher, or a practising professional.

Esko Juuso, Jari Ruuska, Gaurav Mirlekar, and Lars Eriksson

# Table of Contents

Preface	I
Program	III
International Program Committee	IV
National Program Committee	IV
International reviewers	V
Keynote abstracts	VI
Panel discussions	X
Author index	XII
List of Articles	XV

## Conferences location

The conference was organized at **Original Sokos Hotel Arina in Oulu, Finland.**

**Oulu City Reception** at Oulu City Hall, 10 September 2024

**Opening**, 11 September 2024

Opening of The Second SIMS EUROSIM Conference on Simulation and Modelling SIMS EUROSIM 2024 and The 65th SIMS conference on Simulation and Modelling (SIMS 2024):

- *Adj. prof. Jari Ruuska, Conference Chair*
- *Address from University of Oulu, prof. Antti Niemi, University of Oulu, Faculty of Technology*
- *Address from Scandinavian Simulation Society (SIMS), prof. Tiina Komulainen, SIMS President*
- *Adj. prof. Esko Juuso, IPC Chair, University of Oulu, Finland*

## Keynote presentations

The Role of Simulation Governance in the AI Era: Applications in Structural Engineering  
*Prof. Antti H. Niemi, University of Oulu, Faculty of Technology, Finland*

Role of Physics-Based Realistic Simulation Environments for Research and Education in Robotics and AI  
*Associate Senior Lecturer Sumeet Gajanan Satpute, Robotics and AI, Luleå University of Technology, Sweden*

## Conference topics

The Proceedings include 66 papers in ten topic areas:

<b>Topics</b>	<b>Pages</b>
Renewable energy, energy storage and power production	1 - 105
Circular Economy: CO <sub>2</sub> capture and use	106 - 176
Steel Industry: thermodynamics and computational methods	177 - 241
Water treatment and nature-based solutions	242 - 265
Bio systems: processes and environmental impacts	266 - 296
Fluid flow and heat transfer	297 - 323
Transportation: engines, vehicles, fuels and electrification	324 - 361
Automation	362 - 424
Machine learning	425 - 455
Modelling in process analysis and optimization	456 - 494

## Conference program

Both conference days started with a keynote and continued with three parallel sessions. The Annual SIMS meeting was held in the end of the first day. The second day ended with

### **Panel Discussion on Future Challenges and Possibilities for Simulation**

More information is available at SIMS website (<https://www.scansims.org/>).

## Conference General Chair

Adjuct prof. Jari Ruuska, University of Oulu, Finland

## International Program Committee

- |   |  |
|---|--|
| Adj. prof. Esko Juuso, University of Oulu, Finland,<br>Chair                    | Prof. Kauko Leiviskä, University of Oulu, Finland  |
| Adj. prof. Jari Ruuska, University of Oulu, Finland,<br>Co-Chair                | Adj. prof. Esa Muurinen, University of Oulu, Finland                                       |
| Dr. Gaurav Mirlekar, University of South-Eastern<br>Norway, Norway, Co-Chair    | Assoc. prof. Idelfonso Nogueira, Norwegian<br>University of Science and Technology, Norway |
| Prof. Lars Eriksson, Linköping University, Sweden,<br>Co-Chair                  | Dr. Markku Ohenoja, University of Oulu, Finland  |
| Prof. Konstantinos Kyprianidis, Malardalen<br>University, Sweden, Co-Chair      | Assoc. prof. Adrian Pop, Linköping University,<br>Sweden                                   |
| Dr. Avinash Renuke, Malardalen University,<br>Sweden, Co-Chair                  | Prof. Vicenç Puig, Universitat Politècnica de<br>Catalunya, Spain                          |
| Dr. Timo Ahola, Outokumpu Stainless, Finland                                    | Assoc. prof. Michela Robba, University of Genova,<br>Italy                                 |
| Prof. Erik Dahlquist, Malardalen University, Sweden                             | Dr. Jani Tomperi, University of Oulu, Finland  |
| Prof. Tormod Drengstig, University of Stavanger,<br>Norway                      | Prof. Tero Tynjälä, LUT University, Finland  |
| Prof. Yrjö Hiltunen, University of Eastern Finland,<br>Finland                  | Dr. Stavros Vouros, Malardalen University, Sweden  |
| Prof. Biao Huang, University of Alberta, Canada                                 | Prof. Archana Balkrishna Yadav, Sardar Vallabhbhai<br>Patel Institute of Technology, India |
| Assoc. prof. David Hästbacka, Tampere University,<br>Finland                    | Assoc. prof. Ru Yan, University of South-Eastern<br>Norway, Norway                         |
| Assoc. prof. Wolfgang Kemmetmüller, Vienna<br>University of Technology, Austria | Prof. Chunhua Yang, Central South University,<br>China                                     |
| Prof. Tiina Komulainen, Oslo Metropolitan<br>University, Norway                 | Dr. Peter Ylen, VTT, Finland   |
| Prof. Juan Ignacio Latorre-Biel, Public University of<br>Navarre, Spain         | Adj. prof. Kai Zenger, Aalto University, Finland   |
|   | Prof. Borut Zupančič, University of Ljubljana,<br>Slovenia                                 |
|   | Prof. Lars Erik Øi, University of South-Eastern<br>Norway, Norway                          |

## National Organizing Committee

- Adj. prof. Jari Ruuska, University of Oulu, Finland, Chair  
Adj. prof. Esko Juuso, University of Oulu, Finland, Co-Chair  
Ms. Anu Randén-Siippainen, Finnish Automation Society, Finland  
Mr. Marko Vuorio, Finnish Automation Society, Finland

## International Reviewers

<b>Title</b>	<b>Givenname</b>	<b>Surname</b>	<b>Affiliation</b>	<b>Country</b>
Dr.	Hao	Chen	Mälardalen University	Sweden
M.Sc.	Beibei	Dong	Mälardalen University	Sweden
M.Sc.	Alessandra	Ghilardi	University of Pisa	Italy
M.Sc.	Kasuni	Guruvita	Mälardalen University	Sweden
Dr,	Heidi	Ivan	Mälardalen University	Sweden
Assoc. prof.	Payman	Jalali	LUT University	Finland
M.Sc.	Simon	Karlsson	Mälardalen University	Sweden
Prof.	Emil	Kurvinen	University of Oulu	Finland
Dr.	Maysam	Majidi Nezhad	Mälardalen University	Sweden
M.Sc.	Nima	Monghasemi	Mälardalen University	Sweden
Dr.	Aarne	Pohjonen	University of Oulu	Finland
Prof.	Mika	Ruusunen	University of Oulu	Finland
Assoc. prof.	Eivind	Haus	University of Stavanger	Norway
M.Sc.	Ladan	Samael	University of South-Eastern Norway	Norway
Dr.	Valentin	Scheiff	Mälardalen University	Sweden
Dr.	Xiaodan	Shi	Mälardalen University	Sweden
Dr.	Geir	Skaugen	SINTEF	Norway
Dr.	Aki	Sorsa	University of Oulu	Finland
M.Sc.	Kristian	Tiiri	University of Oulu	Finland
Dr.	Ravi Nath	Tiwari	University of Genoa	Italy
Dr.	Ari	Vuokila	University of Oulu	Finland

# The Role of Simulation Governance in the AI Era: Applications in Structural Engineering

*Antti H. Niemi*

*Dean, Professor, Faculty of Technology  
University of Oulu, Finland*

## Abstract

In the era of artificial intelligence (AI), machine learning (ML) and advanced computational techniques have revolutionized the design, analysis, and assessment processes in structural engineering. These methods enable engineers to optimize structural designs, predict behaviour under complex loading conditions, and evaluate risks more efficiently. However, the increasing reliance on simulations to inform critical decisions highlights the growing need for robust simulation governance. This governance ensures that computational models are credible, transparent, and aligned with real-world performance expectations.

Simulation governance encompasses a range of activities aimed at ensuring the quality and reliability of computational models and results. Key aspects include verification and validation (V&V) of models, uncertainty quantification (UQ), and adherence to principles of data integrity. Verification ensures that mathematical models are accurately solved, while validation confirms that models faithfully represent physical phenomena. UQ addresses the inherent uncertainties in material properties, boundary conditions, and loading scenarios, offering insights into the reliability of predictions. Together, these elements form the foundation for trust in simulations in safety-critical applications such as structural engineering.

One of the most pressing challenges in applying simulation governance to civil engineering is the disparity between time scales in structural performance. While simulations often focus on short-term responses (e.g., dynamic behaviour under seismic loads), many civil structures are designed for decades or even centuries of service life. Accurate modelling of long-term phenomena, such as material creep, fatigue, and environmental degradation, requires integrating multi-scale and multi-physics approaches. These considerations add complexity to the modelling process and necessitate careful calibration and validation to ensure predictive accuracy over extended time horizons.

Finite element analysis (FEA), a cornerstone of computational structural engineering, presents additional challenges requiring robust governance frameworks. Issues such as stress recovery at interfaces, the treatment of stress concentrations around geometric discontinuities, and the resolution of singularities demand careful numerical treatment. Classical finite element methods often struggle with accuracy in these areas, particularly under complex loading or irregular geometries. The Discontinuous Petrov-Galerkin (DPG) methodology has emerged as a promising approach to address these challenges. By using tailored test spaces and robust stability properties, DPG methods improve the reliability of stress predictions, especially in regions of interest like sharp corners or material interfaces. These advancements align with the principles of simulation governance by providing rigorously validated computational tools.

Another critical dimension of simulation governance in the AI era involves integrating sustainability considerations into computational frameworks. Civil engineers increasingly face pressures to minimize environmental impacts, optimize resource use, and design for adaptability and resilience. Simulation tools informed by AI can evaluate the lifecycle performance of structures, assess carbon footprints, and

explore trade-offs between competing design objectives. However, these applications rely on high-quality input data, calibrated models, and rigorous V&V processes to ensure meaningful results.

In summary, simulation governance is indispensable for harnessing the full potential of AI-driven tools in structural engineering. By addressing challenges such as time-scale disparities, numerical accuracy in FEA, and the integration of sustainability considerations, robust governance frameworks ensure that computational insights translate into reliable, safe, and efficient designs. As structural engineering continues to evolve in the AI era, the development and enforcement of these frameworks will play a pivotal role in shaping the future of the field.

### **Biography**

Dr. Antti H. Niemi earned his M.Sc. in Engineering Physics (2004) and D.Sc. in Mathematics (2009) from the Helsinki University of Technology (now Aalto University). Since 2020, he has been a professor of structural design and engineering mechanics at the University of Oulu, where he also serves as Dean of the Faculty of Technology. In this role, he leads a diverse and dynamic academic community, fostering innovation and collaboration across disciplines.

Dr. Niemi's research focuses on developing advanced mathematical models and numerical simulation methods in structural mechanics, addressing topics such as finite element methods, stress analysis, stability, and multi-scale modelling. His work integrates fundamental mathematics with practical applications to improve the reliability, efficiency, and sustainability of structural designs.

He has held research positions at international institutions, including Aalto University (Finland), KAUST (Saudi Arabia), and the Oden Institute at the University of Texas at Austin (USA), which have enriched his global perspective and interdisciplinary expertise. In addition to his academic roles, he has worked in engineering consulting as a bridge and structural designer, combining hands-on experience with theoretical advancements.



## Role of Physics-Based Realistic Simulation Environments for Research and Education in Robotics and AI

*Sumeet Gajanan Satpute,  
Associate Senior Lecturer,  
Robotics and AI, Luleå University of  
Technology, Sweden*

*George Nikolakopoulos,  
Chair Professor of Robotics and AI,  
Luleå University of Technology, Sweden*

**Abstract:** Physics-based realistic simulation environments are vital for advancing research and education in AI, providing an accurate and controlled platform for testing algorithms and models. These environments simulate real-world physics, including dynamics, collisions, and sensor interactions, allowing AI systems to learn and adapt in complex, lifelike scenarios. In research, they enable experimentation with AI-driven robotics, autonomous systems, and reinforcement learning without the constraints of physical setups. For education, they offer hands-on experiences for students to explore AI concepts and algorithms in dynamic environments, bridging the gap between theoretical learning and practical application, fostering innovation and understanding. During this talk, we will present our efforts in the development of relevant simulation environments within the Robotics and AI group, at Lulea University of Technology, Sweden. These environments are used in the ongoing courses as well as the advancement of autonomy algorithms towards their field implementation.





## Biography

### Education

2012	Master of Technology in Electrical Engineering with specialization in Control Systems, Mumbai University, India Thesis: Control of noisy underactuated mechanical systems
2021	PhD in Onboard Space systems, Luleå University of Technology, Luleå, Sweden. Thesis: Guidance and Control of Multiple Spacecraft Formation

### Professional Experience

2023	Associate Senior Lecturer, Robotics and AI Group, Luleå University of Technology, Luleå, Sweden.
2021-2023	Postdoctoral Researcher, Robotics and AI Group, Luleå University of Technology, Luleå, Sweden.
2015-2021	PhD in collaboration with OHB systems, Sweden and Onboard Space Systems, Luleå University of Technology, Kiruna, Sweden.
2012-2015	Lecturer, Electronics Department, Veermata Jijabai Technological Institute (VJTI), Mumbai, India

# Panel Discussion on Future Challenges and Possibilities for Simulation

*Chair: Adj. prof. Jari Ruuska, University of Oulu, Finland*

*Panelists:*

- *Associate Senior Lecturer Sumeet Gajanan Satpute, Division: Signals and Systems, Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Sweden*
- *R&D Manager Severi Anttila, Outokumpu, Tornio, Finland*
- *Prof. Tiina Komulainen, SIMS President, Oslo Metropolitan University, Oslo, Norway*
- *Senior prof. Erik Dahlquist, Past SIMS President, School of Business Society and Engineering, Division of Automation in Energy and Environmental Engineering, Västerås, Sweden*
- *Adj. prof. Esko Juuso, Conference IPC chair, Past EUROSIM President, Control Engineering, Environmental and Chemical Engineering, Faculty of Technology, University of Oulu, Finland*

The panel discussion was the last part the conference. The panellists were the keynote presenter Sumeet Gajanan Satpute (SS), the current president of SIMS Tiina Komulainen (TK) and the IPC chair of the conference Esko Juuso (EJ), who is also a past president of both Eurosim and SIMS. Industry was represented by Severi Anttila (SA). Erik Dahlquist (ED) has a long experience in simulation, including over 20 years activityiesin SIMS. The chair of the national organizing committee Jari Ruuska (JR) was the chair of the panel. The discussion focused on four questions: artificial intelligence, importance of the expert knowledge, requirements of measurement data and use of written or spoken expert data. These questions were presented by the panel chair. The audience raised a question on limitations on relying only on expert knowledge. The audience also asked why to participate in the next SIMS?

The answers of the panellists were collected during the panel discussions by Dr. Markku Ohenoja and doctoral student Henri Välikangas.

## 1. Artificial intelligence

- ED: AI is as good as the data fed into it.
- EJ: AI does not answer to everything, but a tool among others.
- SS: Requirements for AI are not reasonable, although the performance of them is getting better all the time.
- SA: If it helps to lessen our workload, it will allow us to focus on other things.

- EJ: Machine learning and AI are additional tools, not direct solution in process industries.

## 2. Importance of expert or domain knowledge

- ED: Go first to the experts to start on a higher level. Measurement data not enough, knowledge needed. Examples from mining safety, paper mill troubleshooting.
- SS: Working in robotics is a multidisciplinary area, so getting input from many experts of different areas is important to reach optimal solutions. More complex systems, more multidisciplinary expertise needed.
- EJ: Data-based solutions without expertise fail to reach the same level which they usually achieve with expert knowledge.
- EJ & TK: It is important to understand the system state, and the fact that companies lose the most amount of profit in error situations, importance of that are hard to model and need to be handled with expert knowledge.
- SA: Data quality combined with operation point detection with experts.

## 3. Requirements of measurement data

- ED: It is important to focus effort on the selected data, as it will define most the performance of the AI solution.
- TK: Data quality combined with operation point detection with experts.

- EJ: More focus should be on the data and the uncertainties present in it, rather than imagining that more data is always better.
- ED: Expertise is needed to guide efforts of data-analysis to the right direction.

#### **4. How to facilitate use of written or spoken expert data in system development data**

- ED: Expert knowledge vs. expert guess: data quality also there, i.e. verification of data is important
- SA: Reliability of prediction/simulation as well! Difficult for operators to re-gain the trust for a monitoring/model once lost.
- EJ/ED/TK: Also operator decisions are subjective sometimes
- ED: Learning from other operators and mills, risk-taking to test and go outside comfort zone in training (for advances systems)
- SS: Experts might have 40 years of experience for good decision-making. AI also needs (some) time to learn.
- EJ: An active research area including new possibilities are available for the challenging problem.
- ED: AI-based advisory systems are needed to support operator to conduct best practices.
- TK: An example is spotting energy saving potential in WWTP.

#### **5. Audience: Can relying upon only expert knowledge can inhibit progress?**

- ED: Relying upon only expert knowledge can inhibit progress, thus a mixture of these should be utilized to achieve the best outcome. Different thinkers are preferred, because it will lead to innovation more than similar thinking people, because no one will want to do things differently or to think on what else their method could be utilized upon.
- SA: Such a risk exists. Relying on peers to innovate doesn't give enough varying information to create a lot of innovative ideas or solutions. It is like talking to a mirror, especially here in the Nordic countries, where there are only a handful of R&D personnel, and all of them come from a similar background.
- TK: It takes time and effort to standardize the culture and get everyone to use the new methods.

- EJ: It would be important for people who have innovative ideas to stand behind their ideas in the face of "old thinkers".
- SA: The field of development should be able to adapt to the new ideas, and not think that the methods haven't evolved with the time, even though they were not working solutions 40 years ago.
- SS: Braking boundaries is a difficult task.
- ED: Culture of work and development needs time to change and adapt to new ideas and new thinkers.

#### **6. Audience: Why to participate in the next SIMS?**

- ED: Mixture of ideas available.
- SA: Possibility for other area experts to hear out of the box.
- TK: Different application areas, different methods when listening others' work is inspiring / gives ideas to own research.

## Author index

Aarne Pohjonen	192, 198, 488	Esa Puukko	211
Abdullahi Adamu	370	Eshetu Janka	274
Aidin Heidari	217	Esin Iplik	223
Aki Lamponen	339	Esko K. Juuso	18, 481, X
Aki Sorsa	356	Fabian Mauss	162
Albert Likang Hu	257	Farjam Karim	441
Ali Moradi	392, 398	Fernando Russo Abegão	370
Antti H. Niemi	VI	Gamunu L Samarakoon	249, 266, 274
Ali Reza Pirouzfaz	362	Arachchige	
Alireza Kakoe	348	Gaurav Mirlekar	98, 362
Alisa Ala-Huikku	407	Geir Skaugen	71, 85
Amare Desalegn Fentaye	48, 433	George Nikolakopoulos	VIII
Amin Mahmoudzadeh	57, 348, 356,	Gudny Øyre Flatabø	249
Andwari	472	Hadi Amlashi	33
Amirhossein Ghazi	168	Hannu Hakalahti	407
Anas Fattouh	48	Heli Vääätäjä	378
Antonio Garcia Martinez	57	Henna Tiensuu	425
Antti Kajjalainen	185, 236	Henri Tervo	185, 236
Aparajita Tripathy	339	Henri Välikangas	456, X
Arnab Chaudhuri	257	Håvard Falch	85
Arthur Sousa de Sena	441	Ida K.B. Stenberg	274
Arvind Keprate	242	Isa Banagar	356
Aryan Bhusal	266	Ishan Rangajith Koralege	441
Assa Aravindh Sasikala Devi	198	Ishmael Nii Nyarko	289
Ayat Gharehghani	348	Solomon	
Britt M. E. Moldestad	138, 312, 318, 392, 398	Ismail Hossain Rafi	392, 398
Carlos Mico Reche	57	Jaakko Palosaari	333
Carlos Pfeiffer	98	Jaakko Suutala	425, 448
Charith Rajapaksha	398	Jamshid Moradi	472
Charlotte Stubenvoll	378	Jani Isokääntä	425
Chinthaka Attanayake	79	Jari Hyvönen	348
Christoffer Moen	130	Jari Juuti	333
Clemens Felsmann	464	Jari Larkiola	192
Damiano Rotondo	385	Jari Ruuska	356, 456, X
Dimuthu Lesthuruge	441	Joachim Rød Knarrum	289
Dong Trong Nguyen	25	Joanna Maria Badach	257
Eetu-Pekka Heikkinen	177	Johan Espelund	71
Eivind S. Haus	385	Johannes Sainio	236
Elisa Olson	211	Joram Wasserfall	464
Emil Kurvinen	356	Jordy Jorritsma	63
Eric Lendormy	348	Josefin Rojas Vasquez	48
Erik Dahlquist	48, X	Judit Kovács	236
		Juha Hirvonen	407
		Juho Alatalo	339

Juho Könnö	57, 356, 472	Mohsen Gholizadefalah	79
Jukka Kömi	204, 228, 236	Moksadur Rahman	1, 93
Jukka Säkkinen	339	Mostafa Abedini	289
Kamelia Boodhoo	370	Nabin Aryal	266, 281, 289
Katrine Marsteng Jansen	242	Neda Razi	122
Kian Golbaghi	348	Neville D'Souza	98
Klaus Hubacek	63	Nils Eldrup	106
Konstantinos Kyprianidis	10, 41, 63, 415, 433	Nirajan Raut	312, 318, 398
Koteswara Rao Putta	362	Niroj Koirala	312, 318
Kristian Sandström	41	Nora C.I. Furuvik	146
Kristian Thorsen	385	Nurul Huda Mahmood	441
Kristian Tiiro	114	Olli Väinölä	192
Kristofer Bölke	223	Omid Lotfizadeh	33
Ladan Samaei	79, 130	Oskar Lind Jonsson	324
Lars Eriksson	324	Oskari Seppälä	192
Lars Erik Øi	79, 106, 122	Otto Kankaanpää	223
Lars-André Tokheim	130, 168	Outi Ruusunen	114
Luis J. Yebra	18	Payman Jalali	297, 304
Maciej Mikulsky	348	Pekka Seppänen	339
Maher Azaza	48	Pekka Uusitalo	370
Mahmoud Elaraby	228	Pentti Karjalainen	204, 228
Majid Nejadseifi	297, 304	Per Morten Hansen	79
Malik Baqeri	242	Perttu Niskanen	356
Malin Harr Overland	385	Petri Hannula	378
Mamdouh Eissa	228	Prakash Bhattarai	138
Manuel González	456	Pramod Ghimire	289
Fernández		Qifeng Shu	177, 185
Marcell Gáspár	236	Rajan Jaiswal	312
Marion Powilleit	464	Rakhi	162
Markku Ohenoja	114, 370, 456, X	Reza Safavi Nick	211
Masoume Shabani	48	Riitta Keiski	114
Masoumeh Dehghanizadeh	106	Rita Kallio	177
Matti Aula	211	Robel S. Bekele	274
Maximilian Dylong	93	Roohallah Surki Aliabad	204
Meysam Majidi Nezhad	48	Sadegh Mehranfar	57
Miika Malin	448	Saeed Sadeghpour	204
Mika Ruusunen	114, 356, 370, 456	Saman Pershen	289
Mikko Leinonen	333	Sambeet Mishra	362
Mohammad Mahdi Salah	348	Samitha Gunarathne	441
Mohammad Rakibul	154, 289	Sarthak Acharya	339
Hasan Chowdhury		Severi Anttila	211, X
Mohammad Reza Babaei	433	Shervin Karimkashi	297, 304
Mohammed Ali	228	Simon Karlsson	41
Mohammed Taha	10	Simon Mählkvist	415
		Soheila Taghavi	138, 146, 154, 392, 398

Soudeh Shamsiri	122
Stavros Vouros	10, 41, 63
Stefan Kirschbaum	464
Stefan Thorburn	1
Stéphane Brochot	456
Subham Kandel	312, 318
Sumeet Gajanan Satpute	VIII, X
Sunil Prasad Lohani	312, 318
Svein Sævik	25
Tauno Tepsa	378
Tero Päivärinta	339
Tero Tynjälä	297, 304
Thea Indrebø	249
Thomas Helander	415
Tiina M. Komulainen	242, X
Timo Fabritius	217
Tomas Ekman	223
Tommi Kokko	378
Toni Luomanmäki	407
Tormod Drengstig	385
Touko Puro	198, 488
Trine Aas-Hansen	25
Trygve Kristiansen	25
Tuomas Alatarvas	177, 185
Uchechukwu Agha Isu	146
Vafa Ahmadi	266, 281
Vahid Farrokhi	79
Vahid Javaheri	204, 228, 236
Vegard Njøten Fagerbakke	25
Wenche Hennie Bergland	249
Virpi Leinonen	425
Zahir Barahmand	33
Zahra Sanidanesh	79

## List of Articles

### 1. Renewable energy, energy storages and power production (01 – 14)

Renewable Energy Resource Risk Quantification and Mitigation Assessment for Mining Micro-Grid <i>Moksadur Rahman and Stefan Thorburn</i>	1
Evaluation of environmental and economic impact of wind turbine blade manufacture at life-cycle level <i>Mohammed Taha, Stavros Vouros and Konstantinos Kyprianidis</i>	10
Dynamic simulation models in the planning of experiments for control development <i>Esko Juuso and Luis J. Yebra</i>	18
Experimental and Numerical Testing of a Multi-Modular Floating Structure with Varying Connection Stiffness <i>Trine Aas-Hansen, Vegard Njøten Fagerbakke, Trygve Kristiansen, Svein Sævik and Dong Trong Nguyen</i>	25
Life Cycle Assessment of Floating Offshore Wind Farms: The Case of Hywind Tampen in Norway <i>Omid Lotfizadeh, Zahir Barahmand and Hadi Amlashi</i>	33
Computationally Efficient Optimization of Long Term Energy Storage Using Machine Learning <i>Simon Karlsson, Stavros Vouros, Kristian Sandström and Konstantinos Kyprianidis</i>	41
A Battery Model for Transportation and Stationary Applications <i>Erik Dahlquist, Maher Azaza, Josefin Rojas Vasquez, Anas Fattouh, Meysam Majidiezhad, Masoume Shabani and Amare Desalegn Fentaye</i>	48
Numerical simulation of thermal runaway kinetic mechanisms and battery thermal model for safety assessment of different lithium-ion battery chemistries <i>Sadegh Mehranfar, Amin Mahmoudzadeh Andwari, Juho Könnö, Antonio Garcia Martinez and Carlos Mico Reche</i>	57
Assessing the impact of rising wind power with energy storage on grid resilience in Sweden to mitigate volatility and enhance grid flexibility <i>Jordy Jorritsma, Stavros Vouros, Konstantinos Kyprianidis and Klaus Hubacek</i>	63
CFD validation of optimized compact heat exchanger designs <i>Johan Espelund and Geir Skaugen</i>	71
Simulation of Ammonia Cracker Process with Aspen HYSYS <i>Per Morten Hansen, Chinthaka Attanayake, Vahid Farrokhi, Mohsen Gholizadefalah, Ladan Samaei, Zahra Sanidanesh and Lars Erik Øi</i>	79
Steady State and Transient Modelling of A Three-Core Once-Through Steam Generator <i>Håvard Falch and Geir Skaugen</i>	85
A Comparative Study of Conventional Lime Kilns and Plasma Calcination: Techno-Economic Assessment and Decarbonization Potential <i>Maximilian Dylong and Moksadur Rahman</i>	93
Assessment of Data-Driven Techniques for Flow Rate Predictions in Sub-sea Oil Production <i>Neville D'Souza, Carlos Pfeiffer and Gaurav Mirlekar</i>	98

### 2. Circular economy: CO<sub>2</sub> capture and use (15 – 23)

Simulation and Cost Estimation of CO <sub>2</sub> Capture with Alternatives for Doubled Capacity <i>Lars Erik Øi, Masoumeh Dehghanizadeh and Nils Eldrup</i>	106
---	-----

Simulation of Biogenic Carbon Capture and Utilization Process Chain	114
<i>Kristian Tiiro, Markku Ohenoja, Outi Ruusunen, Riitta Keiski and Mika Ruusunen</i>	
Simulation Model for an Amine-based CO <sub>2</sub> Capture Rig	122
<i>Soudeh Shamsiri, Neda Razi and Lars Erik Øi</i>	
Design of electrified fluidized bed calciner for direct capture of CO <sub>2</sub> from cement raw meal	130
<i>Ladan Samaei, Lars-Andre Tokheim and Christoffer Moen</i>	
Performance Analysis of Advanced Wells in Reservoirs Using CO <sub>2</sub> Enhanced Oil Recovery	138
<i>Prakash Bhattaraj, Soheila Taghavi and Britt M.E. Moldestad</i>	
CO <sub>2</sub> Enhanced Oil Recovery in Reservoirs with Advanced Wells: Simulations and Sensitivity Analysis	146
<i>Uchechukwu Agha Isu, Nora C.I. Furuvik and Soheila Taghavi</i>	
CO <sub>2</sub> Storage and Evaluation of Important Parameters Affecting the CO <sub>2</sub> Plume Distribution: Simulation and Sensitivity Analysis	154
<i>Mohammad Rakibul Hasan Chowdhury and Soheila Taghavi</i>	
Equilibrium analysis for methanation focusing on CO <sub>2</sub> derived substitute natural gas	162
<i>Rakhi and Fabian Mauss</i>	
Modelling and Simulation of CO <sub>2</sub> Capture through Aqueous Indirect Mineralization using CaO-containing By-products	168
<i>Amirhossein Ghazi and Lars-Andre Tokheim</i>	

### 3. Steel Industry: thermodynamics and computational methods (24 – 33)

Phase Transformations in Steelmaking Slags: A Thermodynamic Approach	177
<i>Tuomas Alatarvas, Rita Kallio, Eetu-Pekka Heikkinen and Qifeng Shu</i>	
Utilizing computational thermodynamics in characterization and classification of non-metallic inclusions in Ti-deoxidized steels	185
<i>Tuomas Alatarvas, Henri Tervo, Antti Kaijalainen and Qifeng Shu</i>	
Cellular automata model for austenite formation and grain growth during heating and holding above austenization temperature	192
<i>Aarne Pohjonen, Oskari Seppälä, Olli Väinölä and Jari Larkiola</i>	
Non-interacting lattice random walks for calculating diffusion controlled growth in solid state for dilute concentrations	198
<i>Aarne Pohjonen, Touko Puro and Assa Aravindh Sasikala Devi</i>	
On the Growth Kinetics of Lamellar and Blocky Austenite during Intercritical Annealing of Hot-Rolled Medium Manganese Steel: Thermodynamic and Diffusion-Controlled Transformation Simulations	204
<i>Roohallah Surki Aliabad, Saeed Sadeghpour, Pentti Karjalainen, Jukka Kömi and Vahid Javaheri</i>	
Effect of Slag Particle Diameter on the Re-melting of Ferrochrome Slag by means of Steelmaking Liquid Slag	211
<i>Reza Safavi Nick, Elisa Olson, Matti Aula, Severi Anttila and Esa Puukko</i>	
Optimizing Energy Consumption in Hydrogen Reduction of Iron Ore Pellet: Insights from HSC Chemistry Analysis	217
<i>Aidin Heidari and Timo Fabritius</i>	
Using an advanced simulation tool for successful conversion of reheating furnace to full oxyfuel operation	223
<i>Esin Iplik, Tomas Ekman, Kristofer Bölke and Otto Kankaanpää</i>	



Computational Designing Approach for Medium Manganese Steels with Potential Better Hydrogen Embrittlement Resistance	228
<i>Mahmoud Elaraby, Mohammed Ali, Mamdouh Eissa, Jukka Kömi, Pentti Karjalainen and Vahid Javaheri</i>	
Physical Simulation of Heat-Affected Zones in a Weld Metal Used with 500 MPa Offshore Steel	236
<i>Henri Tervo, Marcell Gáspár, Judit Kovács, Antti Kaijalainen, Vahid Javaheri, Johannes Sainio and Jukka Kömi</i>	

#### 4. Water treatment and nature-based solutions (34 – 36)

Comparison of ML and ASM models for effluent nutrient estimation in the Hias Process	242
<i>Tiina M. Komulainen, Malik Baqeri, Katrine Marsteng Jansen and Arvind Keprate</i>	
Anaerobic digestion of biosolid pyrolysis liquid and hydrolyzed sludge - simulation with extended ADM1 model	249
<i>Thea Indrebø, Gudny Øyre Flatabø, Wenche Hennie Bergland and Gamunu L. Samarakoon Arachchige</i>	
Green infrastructure for resilient urban design: the mapping and management of green roofs in Oslo	257
<i>Albert Likang Hu, Joanna Maria Badach and Arnab Chaudhuri</i>	

#### 5. Biosystems: processes and environmental impacts (37 – 40)

Process simulation for biogas upgrading and biomethane recovery using biofilm-based reactors	266
<i>Vafa Ahmadi, Aryan Bhusal, Gamunu L Samarakoon Arachchige and Nabin Aryal</i>	
Modelling and Simulation of Full-Scale Sequential Batch Reactor Biological Process Using GPS-X	274
<i>Ida K.B. Stenberg, Robel S. Bekele, Gamunu L.S. Arachchige and Eshetu Janka</i>	
Kinetic modelling and simulation of bioanode and biocathode in a bioelectrochemical cell for carbon dioxide reduction	281
<i>304Vafa Ahmadi and Nabin Aryal</i>	
Alternative fuels for the maritime industry and their impact on flue gas composition	289
<i>Mostafa Abedini, Mohammad Rakibul Hasan Chowdhury, Saman Pershen, Joachim Rød Knarrum, Ishmael Nii Nyarko Solomon, Pramod Ghimire and Nabin Aryal</i>	

#### 6. Fluid flow and heat transfer (41 – 44)

Performance of direct air capture process in honeycomb channel configuration: A CFD study	297
<i>Majid Nejadseifi, Shervin Karimkashi, Tero Tynjälä and Payman Jalali</i>	
Computational analysis of conjugate heat transfer in a 2D rectangular channel with mounted obstacles using lattice Boltzmann method	304
<i>Majid Nejadseifi, Shervin Karimkashi, Tero Tynjälä and Payman Jalali</i>	
Experimental and computational studies to investigate flow dynamics of Geldart A and Geldart B particles in a Circulating Fluidized Bed, CFB	312
<i>Subham Kandel, Niroj Koirala, Nirajan Raut, Rajan Jaiswal, Sunil Prasad Lohani and Britt Margrethe Emilie Moldestad</i>	

Impact of grid sensitivity and drag model along with the height of recirculating pipe on a cold flow circulating fluidized bed	318
<i>Subham Kandel, Nirajan Raut, Niroj Koirala, Sunil Prasad Lohani and Britt Margrethe Emilie Moldestad</i>	

## 7. Transportation: engines, vehicles, fuels and electrification (45 – 49)

The Application and Advantages of a Generic Component-Based SI/CI Engine Model with VVA Compatibility	324
<i>Oskar Lind Jonsson and Lars Eriksson</i>	
Modeling of a tire mounted energy harvester using an inertial and analytical tire deformation model	333
<i>Mikko Leinonen, Jaakko Palosaari and Jari Juuti</i>	
Interoperability Challenges and Opportunities in Vehicle-in-the-loop Testings: Insights from NUVE Lab's Hybrid Setup	339
<i>Sarthak Acharya, Aparajita Tripathy, Juho Alatalo, Pekka Seppänen, Aki Lamponen, Jukka Säkkinen and Tero Päivärinta</i>	
New Chemical Kinetics Mechanism for Simulation of Natural Gas/Hydrogen/Diesel Multi-fuel Combustion in Engines	348
<i>Mohammad Mahdi Salah, Amin Mahmoudzadeh Andwari, Alireza Kakoe, Kian Golbaghi, Jari Hyvönen, Ayat Gharehghani, Maciej Mikulsky and Eric Lendormy</i>	
Driving force model for a real-time control concept of a hybrid heavy duty vehicle	356
<i>Jari Ruuska, Aki Sorsa, Isa Banagar, Perttu Niskanen, Amin Mahmoudzadeh Andwari, Mika Ruusunen, Emil Kurvinen and Juho Könnö</i>	

## 8. Automation (50 – 57)

Model Predictive Control for Integrated Photovoltaic (PV) and Electrolysers System	362
<i>Ali Reza Pirouzfara, Sambeet Mishra, Gaurav Mirlekar and Koteswara Rao Putta</i>	
Dynamic Reactor Modelling and Operability Analysis of Xylose Dehydration to Furfural Using an Extractive-reaction Process in an Agitated Cell Reactor	370
<i>Markku Ohenoja, Pekka Uusitalo, Fernando Russo Abegão, Abdullahi Adamu, Kamelia Boodhoo and Mika Ruusunen</i>	
CARLA-based digital twin via ROS for hybrid mobile robot testing	378
<i>Charlotte Stubenvoll, Tauno Tepsa, Tommi Kokko, Petri Hannula and Heli Väättäjä</i>	
Identifiability and Kalman Filter Parameter Estimation Applied to Biomolecular Controller Motifs	385
<i>Eivind S. Haus, Malin Harr Overland, Kristian Thorsen, Damiano Rotondo and Tormod Drengstig</i>	
A Novel Approach to Simulating the Performance of Autonomous Inflow Control Devices	392
<i>Ismail Hossain Rafi, Ali Moradi, Soheila Taghavi and Britt M. E. Moldestad</i>	
Integration of Dynamic Multiphase Flow and Reservoir Models for Improved Oil Recovery Simulation	398
<i>Charith Rajapaksha, Ismail Hossain Rafi, Nirajan Raut, Ali Moradi, Soheila Taghavi and Britt M. E. Moldestad</i>	
Integration of Optimization Methods into Simulation Technology for Manufacturing via Warehouse Optimization	407
<i>Hannu Hakalahti, Alisa Ala-Huikku, Toni Luomanmäki and Juha Hirvonen</i>	

Evaluating Modelling Performance: Sensitivity Analysis of Data Volume in Industrial Batch Processes	415
<i>Simon Mählkvist, Thomas Helander and Konstantinos Kyprianidis</i>	

## 9. Machine learning (58 – 61)

Machine Learning -based Optimization of Biomass Drying Process: Application of Utilizing Data Center Excess Heat	425
<i>Henna Tiensuu, Virpi Leinonen, Jani Isokääntä and Jaakko Suutala</i>	
Enhanced Anomaly Detection in Aero-Engines using Convolutional Transformers	433
<i>Mohammad Reza Babaei, Amare Desalegn Fentaye and Konstantinos Kyprianidis</i>	
A Deep-Unfolding Approach to RIS Phase Shift Optimization Via Transformer-Based Channel Prediction	441
<i>Ishan Rangajith Koralege, Arthur Sousa de Sena, Nurul Huda Mahmood, Farjam Karim, Dimuthu Lesthuruge and Samitha Gunarathne</i>	
Data Center Resource Usage Forecasting with Convolutional Recurrent Neural Networks	448
<i>Miika Malin and Jaakko Suutala</i>	

## 10. Modelling in process analysis and optimization (62 – 66)

Evaluation of Model Uncertainty Propagation in Mineral Process Flowsheet Designs	456
<i>Henri Välikangas, Markku Ohenoja, Stéphane Brochot Brochot, Manuel González Fernández, Jari Ruuska and Mika Ruusunen</i>	
Optimizing Annual-Coupled Industrial Energy Systems with Sequential Time Dependencies in a Two-Stage Algorithm	464
<i>Marion Powilleit, Stefan Kirschbaum, Joram Wasserfall and Clemens Felsmann</i>	
Numerical Methods for the Flow Fields; A Comparative Review	472
<i>Jamshid Moradi, Amin Mahmoudzadeh Andwari and Juho Könnö</i>	
Nonlinearity Analysis of Variables for Modelling and Control	481
<i>Esko K. Juuso</i>	
GPU acceleration of average gradient method for solving partial differential equations	488
<i>Touko Puro and Aarne Pohjonen</i>	

**DOI: 10.3384/ecp212.0xx (where xx is your Article\_No)**

# Renewable Energy Resource Risk Quantification and Mitigation Assessment for Mining Micro-Grid

Moksadur Rahman Stefan Thorburn

*ABB Corporate Research Center, Forskargränd 7, 72226 Vasteras,  
Sweden*

---

**Abstract:** As one of the most energy-intensive industries, mining accounts for over one-third of industrial final energy consumption. With the growing mineral demand, combined with declining ore grades, it is expected that the energy demand in mining will only grow in the future, potentially increasing its already large greenhouse gas footprint. With rising energy costs, renewable energy presents a viable option not only to improve the environmental footprint but also to reduce overall costs with optimized operation of mines. While renewable energy generators i.e., solar photovoltaics and wind turbines offer numerous benefits like modularity, environmentally friendliness, and natural availability; the major drawbacks are their temporal intermittency and seasonal and long-term variability. Hence, these generators pose a resource risk that the actual quantity of wind and solar irradiation can be less than expected. The resource risk imposes uncertainty in short-, medium- and long-term energy generation and consumption. Hence such risk needs to be actively considered and mitigated during the evaluation and operational phase of renewable or hybrid energy system projects. This paper provides a comprehensive review of renewable resource risk quantification techniques. Subsequently, a list of renewable energy resource risk quantification methods is discussed i.e., renewable reliability (i.e., the percentage of demand met by renewables), energy deficit and energy oversupply index, probability of exceedance (PoE) for annual energy production (AEP), probability of generating at least  $k$  MW of renewable power, capacity factor. Finally, some selected matrices are used to assess the effect of different risk mitigation options, e. g. the optimal size of energy storage.

*Keywords:* Mining, renewable resource risk, resource reliability, sustainable energy.

---

## 1. INTRODUCTION

Mining is one of the most energy-intensive industries. It accounts globally for 11% of the total final energy consumption and 38% of industrial final energy consumption (McLellan et al., 2012). Also, being one of the largest expenses in mining, energy on average accounts for 15% to 40% of the total operational cost (Igogo et al., 2020). Having said that, the sector's final energy consumption is still heavily dependent on fossil fuels, with 62% of final energy consumption being made up of oil, gas, and coal directly, while 35% is made up of electricity from the grid that often includes fossil fuels (Maennling and Toledano, 2018). With the increase in mineral demand, combined with declining ore grades, it is expected that the energy demands in mining will only grow in the future, potentially increasing its already large greenhouse gas (GHG) footprint (Nasirov and Agostini, 2018). Under these circumstances, the mining industry has been under enormous pressure to reduce its environmental impacts. This is leading to an increasing interest in adopting renewable energy to power mining operations. With increasing energy costs, renewable energy like solar and wind present a viable option not only to improve the environmental

footprint but also to reduce overall costs with optimized operation of mines. While renewable energy generators i.e., solar photovoltaics (PV) and wind turbines have numerous benefits such as environmental friendliness, natural availability, and lower life-cycle cost; the major drawbacks are their temporal intermittency and seasonal and long-term variability. Therefore, renewable energy generators pose a resource reliability risk that can be manifested as a quantity risk—i.e., the risk that the quantity of wind and sunshine will be less than expected (Bolinger, 2017). The resource reliability risk imposes uncertainty in short-, medium- and long-term energy generation and consumption. Hence such risk needs to be actively considered and mitigated during the evaluation and development phase of renewable or hybrid energy system projects. Therefore, a methodology is required to quantify the energy supply risk in a renewable or hybrid energy generation system. Subsequently, such risk quantification method can be used to analyze the effect of different risk mitigation options, e. g. the optimal size of energy storage and/or backup/emergency energy generator or through grid or demand flexibility. In this paper, we have focused exclusively on battery energy storage as a risk mitigation option. Nevertheless, the methodology presented can be adapted to other mitigation strategies as well.

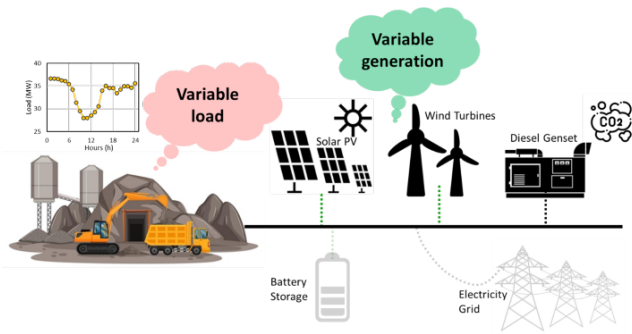


Fig. 1. Hybrid energy generation system for mines

## 2. MATERIALS AND METHODS

### 2.1 Hybrid Energy Generation System for Mines

A hybrid energy system combines multiple types of energy generators and/or backup energy sources like storage or grid in a complementary fashion to ensure dependable power supply at a competitive cost (Fathima and Palanisamy, 2015). One of the major benefits is that it can capitalize on existing grid infrastructure and add different components to help reduce costs, environmental impacts, and system disruptions. Hence, a hybrid energy system is a viable option that can help the mining industry to transition away from fossil fuel-based operations. Depending on the renewable resource availability and economic feasibility, a hybrid energy system for mines can consist of solar PV, wind turbine along with diesel generators (DG) and/or storage and grid as backup source, as shown in Fig. 1. Of course, the actual configuration will vary depending on site and mine-specific requirements. However, including solar and wind generators in mining energy generation systems comes with disadvantages like temporal intermittency, and seasonal and long-term variability. Traditional mining energy sources like diesel and grid can deliver energy whenever needed. Contrarily, solar and wind generators can only deliver energy when the sun is shining, and wind is blowing. This makes both the demand side and the generation side of the energy system variable. The challenge lies in the need to constantly balance energy demand with energy generation. Hence, backup sources like storage, diesel generators, or connection to regional electricity grids are essential for the security of supply. Due to the remote nature of mining sites, combining solar and wind energy with battery energy storage systems (BESS) is seen as the most viable option to initiate energy transition in the mining industry.

### 2.2 Wind Energy

The use of wind energy in electricity generation is widespread in today's world. Typically wind turbines, devices that convert the kinetic energy of wind into electrical energy, are used for this purpose. Wind energy can also be used to complement solar energy due to its availability during the night and on cloudy days.

#### Wind Resource Assessment

The economic value of wind energy generators depends on the availability of wind resources at the intended geographical location. Hence, the wind resource assessment is

a crucial part of the feasibility study. Even though the approaches for resource assessment typically vary depending on many factors like purpose, stage of development, and generator types under consideration, such calculations are often based on some on-site wind measurement, sometimes augmented by meteorological modeling, and likely to be combined with longer-term measurements from offsite (but ideally nearby) reference stations. Typically, Weibull distribution is used to represent the frequency of wind speeds at a specific location. The general form of the Weibull distribution for wind speed takes the following form as shown in equation (1) as presented in (Al Buhairi, 2006),

$$f_v(v) = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad (1)$$

#### Wind Turbine Modelling

A model of a wind turbine is typically represented by a power curve, which is a plot between power output and wind speeds at a particular hub height. In this work, a piecewise model of a power curve from Devrim and Eryilmaz, 2021 is used as described in equation (2),

$$P_i(v) = \begin{cases} 0 & \text{if } v < v_{ci} \text{ or } v \geq v_{co} \\ 1 & \text{if } v_{ci} \leq v < v_r \\ P_r & \text{if } v_r \leq v < v_{co} \end{cases} \quad (2)$$

where  $v_{ci}$ ,  $v_{co}$ , and  $v_r$  are cut-in, cut-off, and rated wind speeds respectively.  $P_r$  is the rated output power of the wind turbine.

Specification data for different wind turbines can be obtained freely from the wind turbine library maintained by Open Energy Platform.

### 2.3 Solar Energy

Solar energy is one of the fastest-growing renewable energy technologies available today. The most common options for utilizing solar energy are PV and solar thermal systems. In this paper, the focus will be on solar PV, which are electronic devices that convert sunlight directly into electricity.

#### Solar Resource Assessment

One of the major factors for the economic feasibility of solar PV systems is the availability of solar energy that can be utilized to produce electricity. Typically, solar irradiation, the amount of energy that reaches a unit area in a unit of time (expressed as  $Wh/m^2$ ), is used to quantify available solar energy. There are different methods available to estimate solar irradiation in a given location such as based on in situ data, derived from satellite data, or a combination of both. Typically, Beta distribution is used to represent the solar irradiation data. The general form of the Beta distribution is depicted in equation (3) as described in Liu et al. (2016),

$$f(r) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{r}{r_m}\right)^{\alpha-1} \left(1 - \frac{r}{A}\right)^{\beta-1} \quad (3)$$

where  $r$  and  $r_m$  are the actual solar intensity and the maximum intensity in a time period,  $\alpha$  and  $\beta$  are the shape parameters of Beta distribution,  $\Gamma$  is a function of Gamma.

#### Solar PV Modelling

The output power of solar PV is calculated from equation (4),

$$P_{PV} = \eta_{PV} A_{PV} PR_{PV} \times GHI \quad (4)$$

where,  $\eta_{PV}$ ,  $A_{PV}$ ,  $PR_{PV}$  and  $GHI$  denotes solar module yield, area, performance ratio (also known as a coefficient for losses that range between 0.9 and 0.5, the default value is 0.75), and global horizontal irradiance. It's worth noting that the above formula is an estimation, as the actual solar power generation depends on many factors such as temperature, shading, dust, and the age of the panel. Some of these factors can be covered by solar module yield.

#### 2.4 BESS

BESS plays an important role in renewable energy integration due to its ability to directly address intermittency issues that are inherent to renewable energies. Major benefits of BESS include assistance in peak shaving, load shifting, voltage and frequency regulation by adding virtual spinning reserve, etc. Typically, a BESS consists of battery cells connected in parallel and series configurations with inverters to facilitate charging and discharging.

##### BESS Modelling

A simplified battery model based on charge quantity and state of charge (SoC) calculation is used in this work. The charge quantity of battery storage at the time  $t$  is calculated by equation (5) according to Deshmukh and Deshmukh (2008),

$$E_B(t) = E_B(t-1)(1-\sigma) + \left( E_{GA}(t) - \frac{E_L(t)}{\eta_{inv}} \right) \times \eta_{bat} \quad (5)$$

where,  $E_B(t)$  and  $E_B(t-1)$  are the charge quantities of battery storage at the time  $t$  and  $(t-1)$ ,  $\sigma$  is the hourly self-discharge rate,  $E_{GA}(t)$  is the total energy generated by the energy source after energy loss in the controller,  $E_L(t)$  is load demand at the time  $t$ ,  $\eta_{inv}$  and  $\eta_{bat}$  are the efficiency of inverter and charge efficiency of battery storage. The charge quantity of battery storage is subject to the constraints represented by equation (6),

$$E_{Bmin} \leq E_B(t) \leq E_{Bmax} \quad (6)$$

where  $E_{Bmax}$  and  $E_{Bmin}$  are the maximum and minimum charge quantity of battery storage.

When referring to BESS, it is more common to use an empirical definition of *SoC*, as represented in equation (7),

$$SoC = \frac{E_B(t)}{E_{Bmax}} \quad (7)$$

#### 2.5 Renewable Resource Risk

Often the availability of renewable resources dictates the economic viability of renewable energy integration. Hence, a feasibility study for renewable energy projects must include resource assessment as a first step. Most often a “P50” estimate of wind speed or solar irradiance is used to calculate the annual energy production that forms the basis for economic calculation. This introduces two primary sources of potential error or bias: 1) the systematic bias from the resource measurement and/or modeling techniques used and 2) the random error related to the inherent short-, medium- and long-term variability of the resource over time. There is a third error of systemic type from energy converter models that are used to estimate the amount of energy generation. Another aspect that is often overlooked in such traditional approaches to the feasibility study is that during the operational phase, the energy demand must always be matched by the energy available instead of ensuring only an annual balance. This means energy must be balanced in short-terms like 15– minutes, hourly, etc., and medium-terms like daily, weekly, monthly, etc. to long-term like yearly and over the project lifetime. Thus, the traditional methods overlook the dynamic energy supply risk and are unable to analyze and provide risk mitigation options and their associated costs. In addition to this, for completeness, such a feasibility study should also consider options related to the other side of the energy balancing act i.e., the demand side flexibility options. Energy consumption peaks should be avoided to reduce the risk of emergency shutdowns and high peak price payments. At least, the decision makers need ways/tools to compare different risk mitigation alternatives related to both the supply- and demand-side that also include associated costs of corresponding options. For example, what are the overall costs and benefits of reducing peak energy demand at rear times with no or exceptionally low renewable generation against installing additional energy storage or emergency backup generators to cover rear peaks? There is a need for a well-defined method/tool to quantify, predict, and reduce the operational risks of energy supply and to evaluate means to reduce these risks.

There are multiple approaches used in the literature to quantify the renewable resource risk. The most common ways to quantify the renewable resource risk are:

- Probability of exceedance (PoE) for annual energy production (AEP).
- Renewable reliability (i.e., the percentage of demand met by renewables).
- Probability of generating at least  $k$  MW of renewable power.
- Energy deficit index and energy oversupply index
- Capacity factor

One of the most widely used matrices is the probability of exceedance for annual energy production (illustrated in Fig. 2(a)), which, with just a few key inputs, can be used to estimate the probability that, for example, the wind or solar generation at a given site will fall below a given level (Bolinger, 2017). This also allows comparison of the resource risk among multiple project sites in terms of probabilistic values. Probability of exceedance is often represented as “ $P$  –level” which ranges from  $P$  1 (i.e., there

is only a 1% chance that actual generation will exceed the  $P1$  estimate) to  $P99$  (i.e., there is a 99% chance that actual generation will exceed the  $P99$  estimate). In comparison to the central or median estimate that is equivalent to the “ $P50$ ” estimate, the probability of exceedance allows the project analyst to choose different “ $P - level$ ” for wind and/or solar generation. Another common way to quantify the renewable resource risk is by calculating the reliability (i.e., the percentage of demand met by renewables) (Tong et al., 2021). Devrim and Eryilmaz (2021) proposed calculating the probability of generating at least  $k$  kW of renewable power. Additionally, simple indicators like the energy deficit index, energy over-supply index, and capacity factor can be calculated to quantify energy supply risk.

According to Tong et al. (2021), the renewable energy system’s reliability index is the percentage of the total load that is met by renewables at a given time, as depicted in equation (8),

$$I_{reliability} = \frac{\text{Renewable generation at time } t}{\text{Load at time } t} \times 100\% \quad (8)$$

The energy deficit index is defined as the ratio between energy shortage and energy demand in a particular hour, as described in equation (9),

$$I_{deficit} = \frac{\text{Energy deficit at time } t}{\text{Load at time } t} \times 100\% \quad (9)$$

Similarly, the energy oversupply index is the ratio between energy oversupply and energy demand in a particular hour, as shown in equation (10),

$$I_{oversupply} = \frac{\text{Energy over supply at time } t}{\text{Load at time } t} \times 100\% \quad (10)$$

The capacity factor of a renewable energy system is a measure of how much electricity the system generates compared to its maximum potential output. It is calculated by dividing the actual energy generated by the system over a certain period of time by the maximum possible energy that could have been generated during that same period (Ahmad et al., 2018). The result is then expressed as a percentage. The formula for calculating the capacity factor can be expressed as below,

$$C_F = \frac{\text{Actual energy generated at time } t}{\text{Maximum possible energy generation}} \times 100\% \quad (11)$$

### 2.6 Renewable Resource Data

There are several ways to obtain solar and wind resource data for a specific location. The historical wind data can be obtained from various sources such as the National Renewable Energy Laboratory (NREL) or other national meteorological services. These datasets usually provide data in the form of wind speed and direction measurements at a specific location and time. Some of these datasets can be downloaded in bulk, while others require you to request access or download data on a case-by-case basis. Similarly, the historical solar radiation data can be

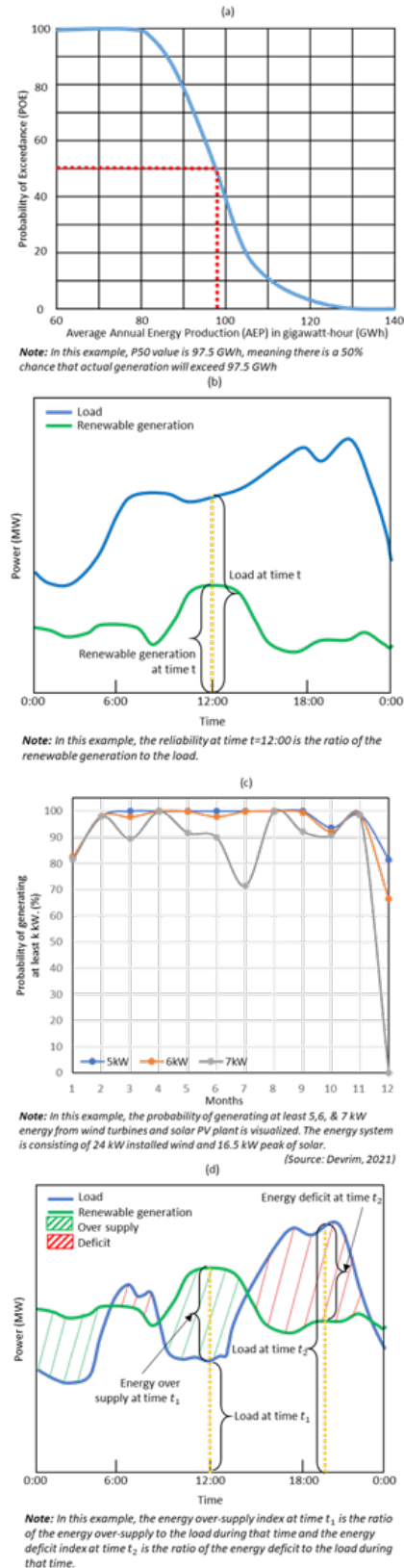


Fig. 2. Illustration of (a) probability of exceedance, (b) reliability (% of load met by renewables), (c) probability of generating at least k kW of renewable power, and (d) energy deficit and oversupply index, are visualized.



obtained from various sources such as NASA’s Surface Meteorology and Solar Energy (SSE) dataset, the NREL or other sources. These datasets usually provide data in the form of solar radiation measurements (usually in  $kWh/m^2$  or  $W/m^2$ ) at a specific location and time. Some other online databases and platforms provide solar and wind data, such as the European Renewable Energy Data Platform (EURODATA) and Renewable Resource Data Center (RReDC). It’s important to note that using historical data alone may not provide a complete picture of the renewable energy resources available in a specific location, and it’s recommended to combine with other sources of information, such as on-site measurements, local weather patterns, topography, and land use, etc. to get a more accurate assessment.

In this work, NASA’s Solar and meteorological resource data- “POWER data” are used for wind and solar resource assessment (NASA, 2024). This satellite and modeled-based database are accurate enough to provide reliable solar and meteorological resource data over regions where surface measurements are sparse or non-existent and offer two unique features – the data are global and contiguous in time (Pavlović et al., 2013). Microgrid design tools such as HOMER and RETScreen also use “Power data” as one of the data sources. Most importantly, the data from “POWER data” is available at multiple temporal levels: hourly, daily, and monthly.

For this work, the hourly data for wind speed, GHI, atmospheric temperature, and pressures are collected over 21 years from 2001 to 2021 for a location in Scotland (Latitude: 57.0161 and Longitude:  $-2.8719$ ; referred to as location-1). To get an overview of the data, wind speed, and solar irradiation are visualized in Figs. 3 and 4. The wind speed shows greater variability with a mean around  $7.4\text{ m/s}$ . Interestingly, for the selected location the wind speed is slightly higher in winter than summer. This is linked to the fact that the winter brings higher temperature gradients. On the other hand, as expected the solar irradiation peaks during summer and very low during winter.

### 3. RESULT AND DISCUSSION

#### 3.1 Prerequisites

To estimate power generation from available wind and solar resources, the wind turbine model described in Section 4.2 and the solar PV model described in Section 5.2 are used. The hourly electricity generation from wind and solar is calculated for the entire historical dataset of 21 years. The wind and solar park are sized such that it can on average meet 20% of the load assumed to be  $20\text{ MW}$ . In reality, the load will be variable but for the sake of simplicity, it is assumed to be constant here. Eventually, six different cases as presented in table 1 are formulated by considering different shares of solar and wind in the renewable energy share. The wind turbine and solar module specifications presented in table 2 and 3 are used for the calculation. Accordingly, the wind farm and solar park capacities are upscaled to fulfill the installed power needed for each use case.

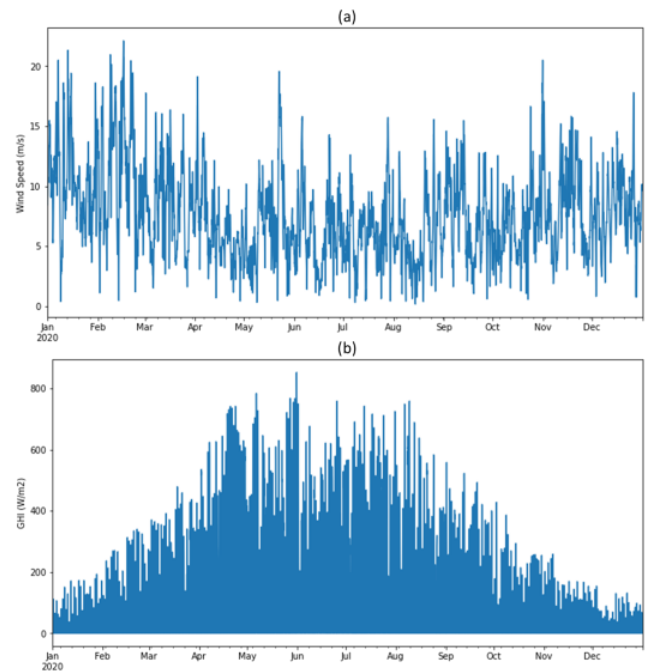


Fig. 3. Hourly (a) wind speed and (b) solar irradiation for the year 2020.

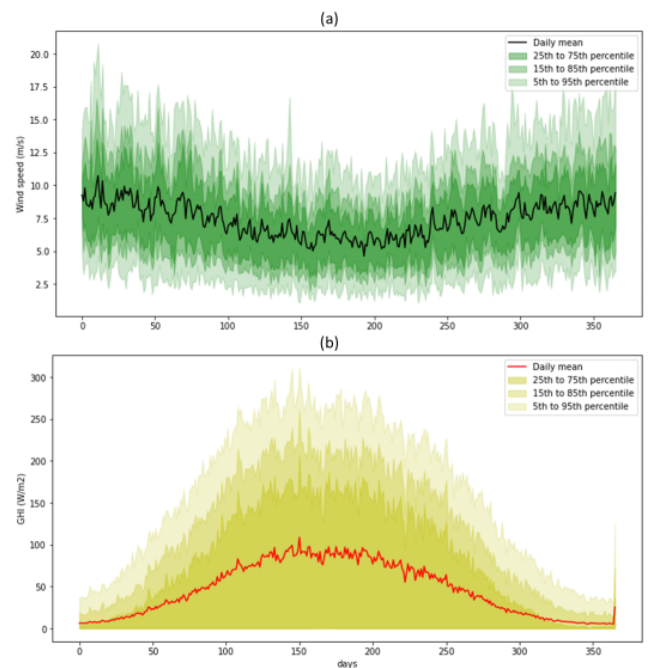


Fig. 4. Hourly (a) wind speed and (b) solar irradiation for the year 2011-2020.

#### 3.2 Renewable Resource Reliability

In Figure 11 the wind and solar energy generation corresponding to the historical dataset for a given location is visualized. To be able to include both wind and solar energy case-2 was selected. As expected, the energy generation from wind and solar follows the same trend as available wind and solar resources. However, the variability of available energy is something to note here. If we look at the hourly mean as well as percentile values, wind has



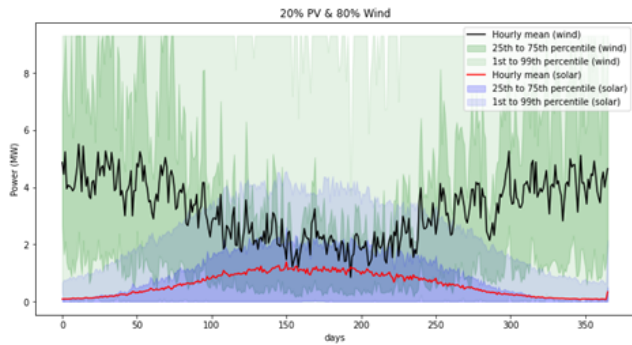


Fig. 5. Individual power generation from wind and solar plant for case 2.

much larger variability than solar. Of course, the wind speed variation is partly amplified due to the high share of wind in case-2. However, variability in wind energy generation comes from high wind speed variability. One interesting fact is that wind generation cannot be more than the total rated power of the wind park that is imposed by the cumulative power curves. Another observation is that over a year solar and wind can act as complementary energy sources for this location. By using both solar and wind energy together, it is possible to reduce the impact of the yearly variability of each source and have a more consistent supply of electricity.

Subsequently, the total energy produced combined by the wind and solar is calculated by adding individual generation. The cumulative power generation for each of the cases is chronologically visualized in Fig. 6. By looking at the average energy generated by each of the system configurations, it is obvious that case-2 and case-3 provide fairly stable average energy throughout the year. The results are summarized in table 4. For better visualization, the mean reliability for different cases is plotted in Fig. 7. The mean reliability decreases as the system configuration

Table 1. Cases with different share of solar and wind

Case no.	Description
Case-1:	0% PV & 100% Wind
Case-2:	20% PV & 80% Wind
Case-3:	40% PV & 60% Wind
Case-4:	60% PV & 40% Wind
Case-5:	80% PV & 20% Wind
Case-6:	100% PV & 0% Wind

Table 2. WT (Enercon e-53/800) specification

Parameter	Value	Unit
Rated power:	810.0	[kW]
Cut-in wind speed:	3.0	[m/s]
Rated wind speed:	12.0	[m/s]
Cut-out wind speed:	26.0	[m/s]
Rotor Diameter:	52.9	[m]
Hub height:	60/73	[m]
Swept area:	2198	[m <sup>2</sup> ]

Table 3. PV module specification

Parameter	Value	Unit
Module efficiency:	15	[%]
Performance ratio:	0.75	[-]
Life:	25	[years]

changes from “Case-1: 0% PV and 100% wind” to “Case-6: 100% PV 0% wind”. Meaning, for this specific location wind heavy systems offer higher mean reliability. On the other hand, the mean energy deficit and the oversupply index increase with solar-heavy systems. However, one must not get deceived by the facts or base their conclusion entirely by looking at the mean values only. The local variation must be considered as well. Mean value over such a long timescale often doesn’t tell the whole story.

Table 4. Renewable Reliability at location-1 for different cases

Case	Mean reliability	Mean P50 reliability	Mean energy deficit index	Mean oversupply index
1	19.8%	12.5%	12.5%	21.9%
2	19.2%	14.2%	10.8%	17.0%
3	18.5%	14.6%	10.8%	17.0%
4	17.9%	12.8%	11.3%	16.4%
5	17.3%	9.1%	11.7%	25.3%
6	16.6%	1.2%	14.5%	32.0%

Subsequently, the same calculation is performed for another location in central Australia (Latitude:  $-22.5909$  and Longitude:  $133.4432$ , referred to as location-2). As can be seen from table 5 and Fig. 8, the trends are reversed as this location has relatively higher solar irradiation and lower wind. This shows how renewable generation and their reliability can be very much location-dependent and thus the system configuration will vary based on renewable resource availability.

Table 5. Renewable Reliability at location-2 for different cases

Case	Mean reliability	Mean P50 reliability	Mean energy deficit index	Mean oversupply index
1	14.5%	12.0%	8.7%	11.5%
2	19.8%	17.9%	9.2%	11.6%
3	25.1%	16.6%	14.2%	24.4%
4	30.4%	12.7%	21.5%	37.6%
5	35.7%	12.7%	21.5%	50.6%
6	41.1%	2.2%	37.1%	63.8%

### 3.3 PoE for AEP

Once the preferred share of wind and solar for a specific location is known, the *AEP* of the system is calculated at different *PoE* levels. To do so, individual *AEP* with *PoE* for wind and solar PV is computed for location-1. For a fair comparison between wind and solar case-3 with 60% wind and 40% solar is selected for this analysis. For wind, the spread between *P99* and *P1* for *AEP* is around  $19.5 \text{ GWh}$  to  $34.5 \text{ GWh}$ . Subsequently, the spread between *P99* and *P1* for *AEP* is around  $10.6 \text{ GWh}$  to  $12.8 \text{ GWh}$  for solar park. For a combined system, the *AEP* values are just added together as presented in Fig. 9. Now, this graph can be used as the basis of financial calculation when a *P-value* is given. As mentioned earlier, typically a *P50 AEP* is used for such calculation. For a more conservative calculation, a higher *P-value* can be used.

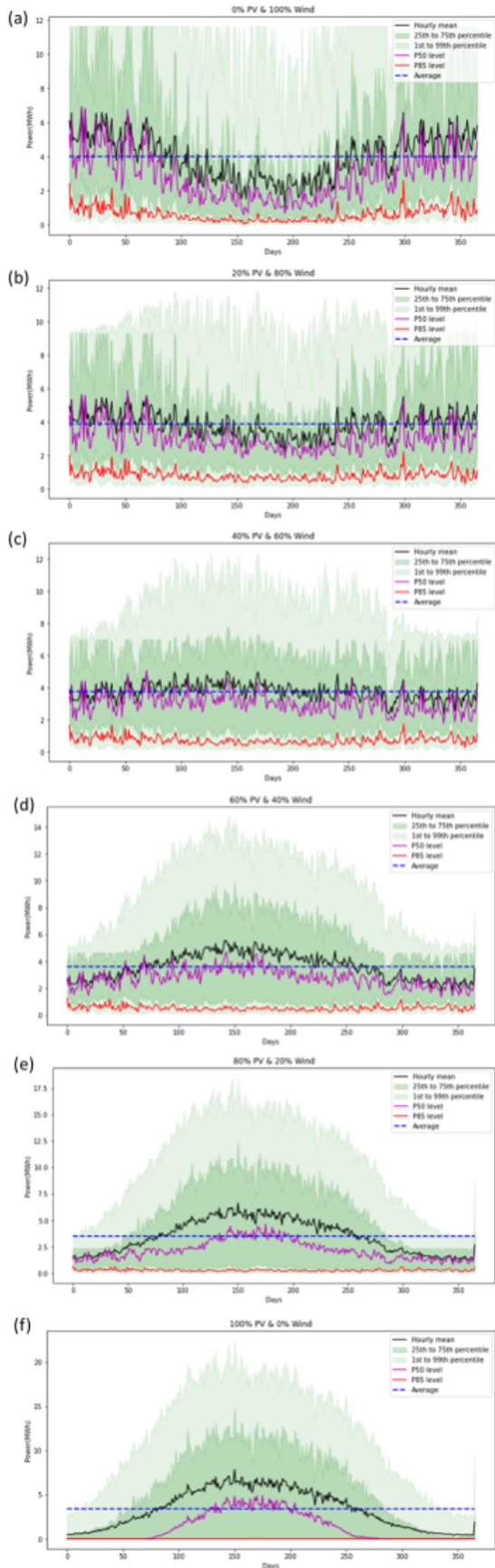


Fig. 6. Combined power generation from wind and solar plant for different cases.

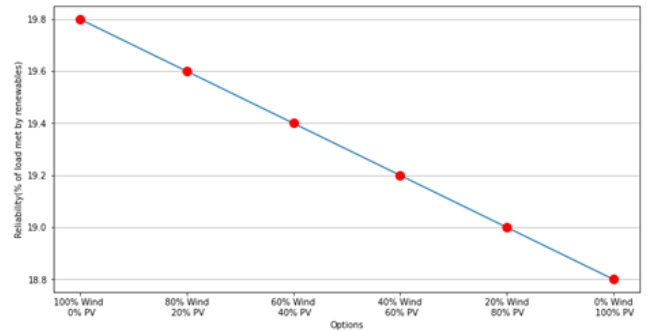


Fig. 7. Renewable reliability of different cases for a location-1.

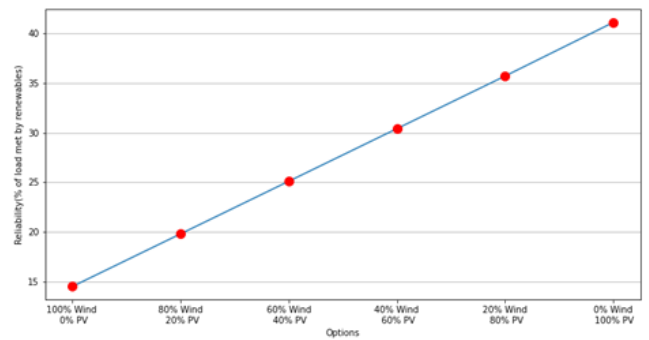


Fig. 8. Renewable reliability of different cases for a location-2.

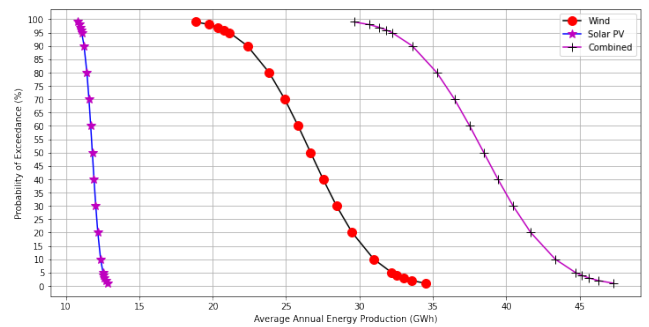


Fig. 9. AEP with PoE for case-3 at location-1.

### 3.4 Renewable Risk Mitigation with BESS

To get a further understanding of renewable energy variability, the energy deficit (or power shortage) and oversupply for case-3 over a year is visualized in Figs. 23 to 25. It is important to note here, a constant load is considered to calculate the energy deficit and the oversupply.

Fig. 10(a) shows, that the energy deficit and oversupply are spread out over the entire year except for some parts, which is preferable when considering a BESS. While analyzing the monthly trends, it was found that there are months where the energy deficit and oversupply are equally distributed (as in Fig. 10(b) and months where that is not the case (Fig. 10(c)). The argument here is that a BESS needs to be designed to cover a month where the renewable generation was particularly low. In that case, the BESS needs to be oversized and that is associated with high capital cost.

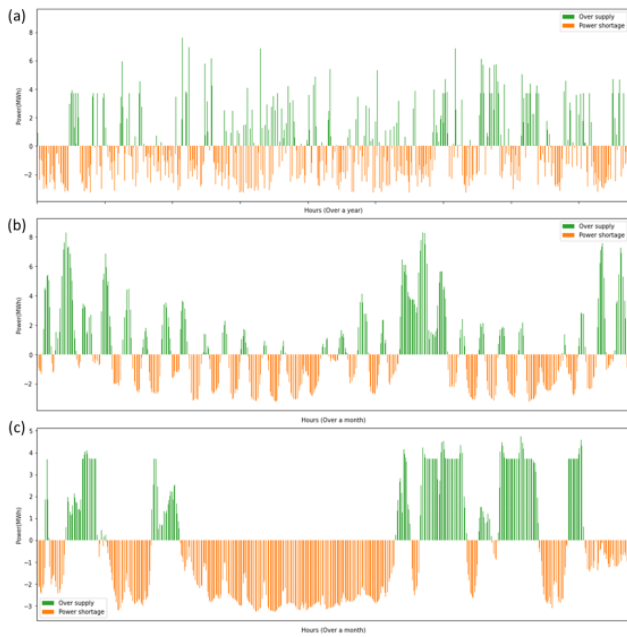


Fig. 10. Energy deficit and oversupply for (a) 2001, (b) June 2001, and (c) December 2001, (case-3, location-1)

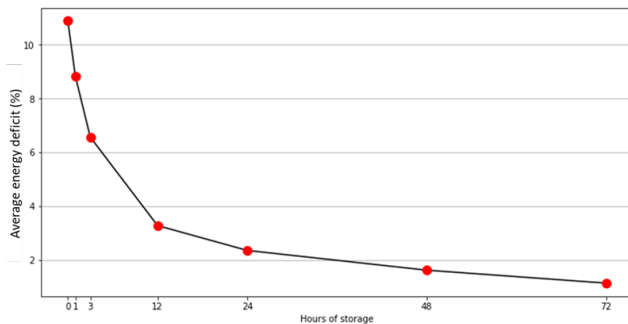


Fig. 11. Average energy deficit index for BESS with Solar and Wind system (case-3, location-1)

To analyze how the BESS can help mitigate some of the variability introduced by renewable energy, the BESS model described earlier is used. Subsequently, different battery size is used to calculate the corresponding average energy deficit index for systems with BESS, solar and wind (case-3). The results are visualized in Fig. 11 where the BESS capacities are represented as hours of storage. Here, “1 hour of storage” corresponds to a BESS size that can cover the entire load by an hour. As can be seen from Fig. 11, initially the average energy deficit index reduces sharply with increasing battery sizes. The slope of the curve diminishes as the BESS size increases.

To analyze how different shares of solar and wind change the energy deficit versus the BESS size graph, the calculations are repeated for different cases (case-1 to case-6). The result is summarized in Fig. 12. “Case-2: 20% PV and 80% wind” can have a lower energy deficit index than “case-1: 0% PV and 100% wind”. Interestingly, with further increase in PV share in the system results in a higher energy deficit index. Overall, the case-2 with BESS can provide the lowest energy deficit index.

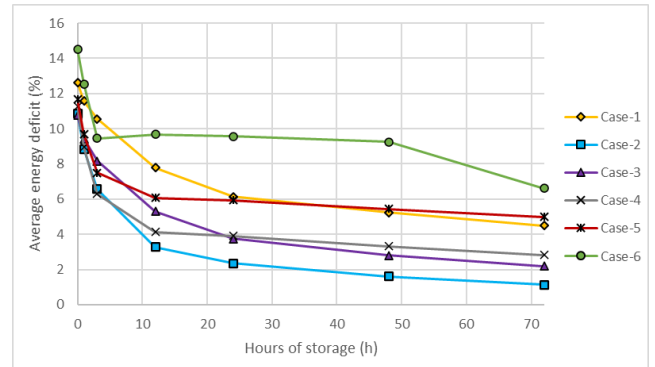


Fig. 12. Average energy deficit index for BESS with Solar and Wind system (case-1 to 6, location-1)

#### 4. CONCLUSIONS

The mining industry has huge potential for renewable energy to meet its energy needs while reducing environmental footprint and overall cost. Freely available solar and meteorological data sources provide a good starting point for the assessment of renewable energy potential, allowing for a fairly accurate and efficient evaluation of the feasibility of different renewable energy projects for mines. These datasets can provide information on factors such as solar radiation levels, wind speeds, and temperature, which are all important for determining the potential output of renewable energy systems. Indeed, the renewable energy generation potential of mines will vary depending on the onsite availability of renewable resources. Accordingly, the preferred share of different renewable sources, here solar and wind, in a mining energy grid will differ significantly at different sites. The reliability of renewable energy generation from the same solar-wind combination can be utterly different in different locations. Interestingly, the reliability trend can reverse for two different locations, meaning increasing the share of solar in a wind-solar mix can result in both decreasing or increasing reliability based on the location. Additionally, using both solar and wind energy together, it is possible to reduce the impact of yearly variability of each source and have a more consistent supply of energy. For financial calculation, annual energy production with the probability of exceedance can act as a better indicator. When it comes to the battery energy storage sizing, the benefit diminishes with increasing size. Meaning, the reduction in overall energy deficit from a solar-wind-battery system decreases exponentially with increasing battery energy storage size. Additionally, the lowest possible energy deficit is also heavily dependent on the share of solar and wind in the system.

#### REFERENCES

- Ahmad, J., Imran, M., Khalid, A., Iqbal, W., Ashraf, S.R., Adnan, M., Ali, S.F., and Khokhar, K.S. (2018). Techno economic analysis of a wind-photovoltaic-biomass hybrid renewable energy system for rural electrification: A case study of Kallar Kahar. *Energy*, 148, 208–234. doi:10.1016/j.energy.2018.01.133.
- Al Buhairi, M.H. (2006). A statistical analysis of wind speed data and an assessment of wind energy potential

- in Taiz-Yemen. *Ass. Univ. Bull. Environ. Res*, 9(2), 21–33.
- Bolinger, M. (2017). Using probability of exceedance to compare the resource risk of renewable and gas-fired generation. Technical report.
- Deshmukh, M.K. and Deshmukh, S.S. (2008). Modeling of hybrid renewable energy systems. *Renew. Sustain. Energy Rev.*, 12(1), 235–249. doi: 10.1016/j.rser.2006.07.011.
- Devrim, Y. and Eryilmaz, S. (2021). Reliability-based evaluation of hybrid wind-solar energy system. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.*, 235(1), 136–143.
- Fathima, A.H. and Palanisamy, K. (2015). Optimization in microgrids with hybrid energy systems – A review. *Renew. Sustain. Energy Rev.*, 45, 431–446. doi: 10.1016/j.rser.2015.01.059.
- Igogo, T., Lowder, T., Engel-Cox, J., Awuah-Offei, K., and Newman, A.M. (2020). Integrating Clean Energy in Mining Operations: Opportunities. *Challenges, Enabling Approaches (No. NREL/TP-6A50-76156)*, 43.
- Liu, Z., Liu, W.l., Su, G.c., Yang, H., and Hu, G. (2016). Wind-solar micro grid reliability evaluation based on sequential Monte Carlo. In *2016 Int. Conf. Probabilistic Methods Appl. to Power Syst.*, 1–6. doi: 10.1109/PMAPS.2016.7764073.
- Maennling, N. and Toledano, P. (2018). The renewable power of the mine. *Available SSRN 3661616*. doi: 10.2139/ssrn.3661616.
- McLellan, B.C., Corder, G.D., Giurco, D.P., and Ishihara, K.N. (2012). Renewable energy in the minerals industry: a review of global potential. *J. Clean. Prod.*, 32, 32–44.
- NASA (2024). <https://power.larc.nasa.gov/>.
- Nasirov, S. and Agostini, C.A. (2018). Mining experts’ perspectives on the determinants of solar technologies adoption in the Chilean mining industry. *Renew. Sustain. Energy Rev.*, 95, 194–202. doi: 10.1016/j.rser.2018.07.038.
- Open Energy Platform (2024). <https://openenergy-platform.org>.
- Pavlović, T.M., Milosavljević, D.D., and Pirsl, D.S. (2013). Simulation of photovoltaic systems electricity generation using homer software in specific locations in Serbia. *Therm. Sci.*, 17(2), 333–347.
- Tong, D., Farnham, D.J., Duan, L., Zhang, Q., Lewis, N.S., Caldeira, K., and Davis, S.J. (2021). Geophysical constraints on the reliability of solar and wind power worldwide. *Nat. Commun.*, 12(1), 1–12.

## Evaluation of environmental and economic impact of wind turbine blade manufacture at life-cycle level

Mohammed Taha. Stavros Vouros. Konstantinos Kyprianidis

Mälardalen University, Västerås Sweden (Tel: 021101300; e-mail:

[mohammed.taha@mdu.se](mailto:mohammed.taha@mdu.se)).

---

**Abstract:** Life cycle analysis is considered as a valuable decision-making tool to oversee the environmental impact of a product through its various stages. Starting from the raw material sourcing up to the end-of-life processes of the product. Life cycle costing is added to the life cycle analysis to augment the economic aspects. One of the main drawbacks of the life cycle analysis is the focus on single path for the life stages as it evaluates single option for each life stage and adds the impact to the following stages. this study presents a tool to evaluate the environmental and economic impact of different options in life cycle stages, determine the possible combination of different life cycle choices, and calculate the emissions, energy intensity and cost of each combination scenario. The study takes wind turbine blade as a case study, where glass fiber reinforced polymers and carbon fibers reinforced polymers are considered as a raw material alternatives with two supply options Europe or China markets, four manufacturing site options (onsite, Denmark, Germany, and China) and four end of life processing options (reuse, pyrolysis, landfill, and mechanical grinding). The results range the different combinations scenarios emissions in the range of (74 – 17) tons of CO<sub>2</sub> eq, the energy intensity between 261 GJ and 863 GJ, and the cost vary from 89000€ to 22,000€. This work presented a logical method for mapping, analyzing, and evaluating the environmental and economic sustainability of a wind turbine blade through different life cycle pathways.

**Keywords:** Life cycle analysis, life cycle cost, wind turbine, wind turbine blades, wind turbine blades end of life.

---

### 1. INTRODUCTION

Wind energy is considered one of the fastest growing renewable energies in Europe. Europe wind energy installed capacity increased from around 140 GW to 272 GW in 2023 (Costanzo et al., 2024). Sweden had even faster growth rate as it expanded from 5.1 GW to 16.3 GW in the same period (Swedish Wind Energy Association, 2024). Several studies proved high technical and economic potentials of wind energy in several areas of Sweden (Warners et al., 2023). In addition to the increased installed capacity, the wind turbines had also been scaled up to maximize wind energy exploitation. Between 2014 and 2023, Sweden's installed turbines count increased by less than half, while its installed capacity increased by more than three times (Swedish Wind Energy Association, 2024). This increase led to greater attention on wind energy sustainability with particular emphasis on the end-of-life treatment. Europe decommissioned 1.5 GW and repowered 736 MW of wind turbines in 2023 (Costanzo et al., 2024). Composite materials used to manufacture wind turbine blades and nacelles pose one of the main environmental challenges due to the difficulty of disposal and recyclability.

To better understand and manage sustainability, life cycle analysis and costing have been utilized as a valuable tool in this sector. Life cycle analysis (LCA) is a tool to evaluate the environmental impact of a product through its different life stages. The complete spectrum of stages starts from the initial raw material resources taken from the environment to the end-

of-life disposal of the product (Bjørn et al., 2017). A general framework has been adopted to perform the LCA, this framework standardized through ISO 14040. The main steps in the standardized framework are goal definition, scope definition, inventory analysis and impact assessment. This logical approach permits to identify parts of the life cycle to emphasis, such as cradle to grave which cover the complete stages spectrum, and gate to gate which focuses on the manufacturing stage, starting from the raw material at the factory gate until the product leave the gate of the factory (Hauschild, 2017).

LCA has been used to evaluate energy sources environmental impact. for renewable energy, the focus is determining the emission reduction and evaluating the energy green pathways. Numerous LCA studies were conducted in wind energy with various goals. Most studies were for specific locations and farms sizes, due to the direct effect on impact per the generated power, mainly the impact of the transportation, installation, and operation and maintenance (O&M) stages. More recent studies focused on the environmental impact of the new technological development on wind turbines such as offshore installations (Brussa et al., 2023; Garcia-Teruel et al., 2022; Yuan et al., 2023). Some studies adopted comparative life cycle analysis (Schreiber et al., 2019). (Ozoemena et al., 2018) compared the environmental impact of 4 different technological improvements opportunities on a 114 MW onshore wind farm located in UK with 1.5 MW, the improvement opportunities evaluated were using stiffer carbon



fiber to enlarge the rotor swept area without increasing the structural loads or equipment requirements, new tower concept using carbon fiber instead of metal allowing to increase the hub height from 65 meters to 100 meters without using higher cranes capacity, and permanent magnet generator using a lower rotational speed (150 rpm).

With expected increase in wind turbine capacities and installed numbers, wind turbine blades draw a significant interest in wind turbine LCA research area because of its high share on the total wind farm environmental impact (15-25) %, only exceeded by the tower (Mali and Garrett, 2022). In addition to the composite materials recycling challenges (glass and carbon fibers) which compromise around 80% of the total mass of the blade (Liu et al., 2019).

Considering the increased attention on carbon fibers and wind turbines blades LCA, this study introduces a scientific approach to evaluate and compare the environmental impact and cost of different options of three life cycle stages, taking a wind turbine blade as a case study.

## 2. METHODOLOGY

A case study has been made to describe the work done. The case study evaluates the alternative options of three stages of wind turbine life cycle namely (material acquisition stage, manufacturing stage and end of life treatment stage). The study blade is the National Renewable Energy Laboratory (NREL) WindPACT project reference turbine blade, 1.5 MW turbine with 33.35 m long and 4.335 tons in mass (Malcolm and Hansen, 2000). The case study assumes the turbine installation location near Eskilstuna Sweden. The study evaluates different options for each life cycle stage. The main evaluation criteria are the climate change impact represented by equivalent carbon dioxide emissions (kg CO<sub>2</sub> eq), energy intensity in mega joules (MJ) and cost in Euros. Figure 1 shows the options evaluated.

### 2.1 Material Acquisition Stage

Materials considered are composite fibers and resin, as it compromises approximately 75% of the total blade weight (Bortolotti et al., 2019). Recent studies proved that using carbon fibers as replacement of glass fibers is assumed to reduce weight due to the higher strength and stiffness. This study assumes a full replacement of glass fibers reinforced polymers (GFRP) by carbon fibers reinforced polymers (CFRP) with the assumption of 20% weight reduction based on (Corona et al., 2024) and (Ennis et al., 2019). Materials weight and cost are assumed based on (Bortolotti et al., 2019), environmental impacts are based on Environmental Footprint Database (Sala and Cerutti, 2018) used with OpenLCA software, (Jensen, 2019), (Rani et al., 2021) and (Strózyk et al., 2024).

The fiber glass environmental impact is assumed to be the same regardless of the directions and axials of the fiber. Materials Prices taken from USA market assumed to be the same for Europe, and 20% less for China due to the low cost of labor and energy. The environmental impact of China sourced materials is scaled up based on the difference of

energy mix impact between China and Europe. Material acquisition stage inputs are presented in Table 1.

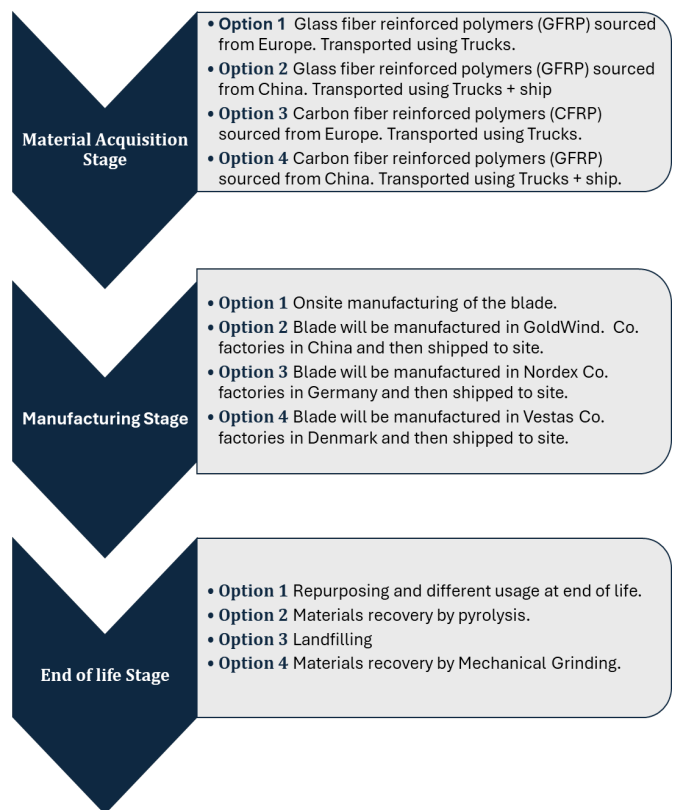


Fig. 1. Options per each stage.

Table 1. Material acquisition stage inputs

Description	Unit	Value
Glass fiber emissions	kg CO <sub>2</sub> eq/kg	4.79
Resin emissions	kg CO <sub>2</sub> eq/kg	6.59
Carbon fiber emissions	kg CO <sub>2</sub> eq/kg	11.2
Glass fiber energy intensity	MJ/kg	35.80
Resin energy intensity	MJ/kg	128.5
Carbon fiber energy intensity	MJ/kg	210
Glass fiber cost	€/kg	2.66
Resin cost	€/kg	3.38
Carbon fiber cost	€/kg	27.9
Fiber glass mass	kg	2453.79
Carbon fiber mass	kg	1963.03
Resin mass	kg	1292.16

## 2.2 Manufacturing Stage

Wind turbine blade manufacturing process involves various steps namely (material cutting, demold, infusion of the components, assembly, trim, overlay, posture, root cut and drill, root fastener installation, surface preparation, paint, surface finishing, weight and balance, inspection, and shipping preparation). These steps utilize simple equipment and machines in addition to a resin curing oven with a curing temperature approximately 70 °C. All equipment is assumed to be powered by electricity. Hence the main impact is caused by using electricity plus transporting the finished blade to the site. Manufacturing locations grid mix data are taken from Environmental Footprint Database (Sala and Cerutti, 2018) browsed using OpenLCA software. Manufacturing processes electricity demand and labor hours are based on (Bortolotti et al., 2019). Electricity average price and labor cost are based on (ILOSTAT, 2022). Manufacturing stage inputs are shown in Table 2.

## 2.3 End of life Stage

To improve wind power sustainability multiple academic and industrial parties are investigating several end-of-life options for wind turbine blades. The studies vary from adding secondary life to the wind turbine blade and using them as a construction material up to numerous ways to recover the fibers (Paulsen and Enevoldsen, 2021) (Rani et al., 2021) (Yousef et al., 2024). Main proposed options for the end-of-life stage are summarized below:

- Functional repurposing (cutting the wind turbine in pieces and using them for simple structures like bus stops and barns).
- Mechanical grinding (producing fiber rich powder to be used for new fibers production).
- Pyrolysis (obtaining pyrolysis gas and oil with other solid by products).
- Fluidized bed (reclaiming fibers through burning out the resin).
- Solvolysis (chemically decomposing the fibers matrix to get the fibers).
- High voltage pulse fragmentation (decomposing the fibers matrix by high voltage electrolysis process).
- Mechanical shredding and cement or asphalt co processing.

This study assumes four options for end-of-life stage, which are:

- Repurposing blade as a high voltage transmission pole based on (Henao et al., 2024).
- Fibers treatment through pyrolysis.
- Recovering fibers through mechanical grinding.
- Land filling at farm stie.

The inputs data for the end-of-life stage shown in Table 3 are taken from (Paulsen and Enevoldsen, 2021; Jensen, 2019; Liu et al., 2019; Sproul et al., 2023). The negative impact values represent the net gain acquired through the end-of-life treatment, it presents the difference between the recycled or

reused fibers and the production of virgin fibers or construction materials.

**Table 2. Manufacturing stage inputs**

Description	Unit	Value
Sweden electricity grid mix emissions	kg CO <sub>2</sub> eq/MW	0.0834
Denmark electricity grid mix emissions	kg CO <sub>2</sub> eq/MW	0.60768
Germany electricity grid mix emissions	kg CO <sub>2</sub> eq/MW	1.19462
China electricity grid mix emissions	kg CO <sub>2</sub> eq/MW	1.9158
Sweden electricity grid mix energy	MJ/MW	3.80815
Denmark electricity grid mix energy	MJ/MW	3.8283
Germany electricity grid mix energy	MJ/MW	7.57788
China electricity grid mix energy	MJ/MW	9.54177
Sweden electricity grid mix average price	€/MW	265.05
Denmark electricity grid mix average price	€/MW	325.5
Germany electricity grid mix average price	€/MW	372
China electricity grid mix average price	€/MW	74.4
One blade manufacturing labor hour	h	407.37
One blade manufacturing electric energy in MW	MW	1.5725
Labor cost in Sweden	€/h	47.8299
Labor cost in Denmark	€/h	53.3448
Labor cost in Germany	€/h	46.0908
Labor cost in China	€/h	4.464

## 2.4 Transportation and shipping

Transportation is considered for materials and manufactured blades. Land transportation is assumed to be by 7 Ton trucks for the materials and 30-ton trucks for the blade. A full 30 Ton truck is assumed for blade transportation, as it depends on the size required to fit the blade rather than the weight dependency. The China options sea transportation assume container shipping for materials and medium barge for blade. Distances assumed are 1000 km for material transportation in Europe and Google maps factory to site measured distance for the blade.

No road topology is considered in the study, the study assumes all roads are paved. Transportation main inputs are shown in Table 4. Emissions and energy intensity are based on Environmental Footprint Database (Sala and Cerutti, 2018)

browsed by OpenLCA software, transportation cost figures are based on (Sander van der Meulen et al., 2023).

**Table 3. End of life stage inputs.**

Description	Unit	Value
Repurposing emissions	kg CO <sub>2</sub> eq/kg	-1.2
Pyrolysis emissions	kg CO <sub>2</sub> eq/kg	-2.06
Landfill emission	kg CO <sub>2</sub> eq/kg	0.05477
Mechanical grinding emissions	kg CO <sub>2</sub> eq/kg	-1.29
Repurposing energy	MJ/kg	1.351
Pyrolysis energy	MJ/kg	30
Landfill energy	MJ/kg	0.35827
Mechanical grinding	MJ/kg	4.8
Repurposing Cost	€/kg	0
Pyrolysis Cost	€/kg	0.2556
Landfill Cost	€/kg	0.0882
Mechanical grinding	€/kg	0.0856

### 2.5 Model and calculations

To evaluate the options of each stage and calculate the total emissions, energy intensity and cost of the three stages, a python model is built to determine all scenarios of options combinations and calculate the total impacts and cost. Figure 2 demonstrates the model schematic diagram.

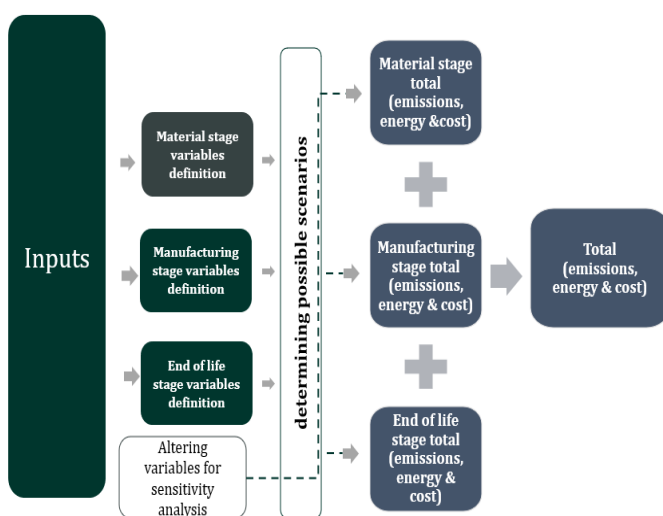


Fig. 2. Model structure.

Sensitivity analysis is made to explain the effect of (carbon fiber mass, glass fiber mass, materials transportation distance,

manufacturing location distance to site, manufacturing location grid mix, manufacturing location electricity price, manufacturing location labor cost, recycling method emissions, recycling method required energy and recycling method cost). The sensitivity analysis baseline scenario is GFRP as a material sourced from Europe, Germany as blade manufacturing location, and repurposing as an end-of-life treatment. The emissions and energy intensity of electricity are treated as independent variables, disregarding their mutual dependency due to the complexity of their relationship and reliance on electricity generation and grid operation technologies.

The study is conducted under the limitation of the data found in literature and Environmental Footprint Database, Industrial sources found was only for complete turbines, and the data source they use for LCA inventory was commercial databases. No consideration is made for the time value of money as the main future cost element is the end-of-life cost which is sourced based on literature estimation as most of the composite materials recycling methods are not mature enough yet.

**Table 4. Transportation inputs**

Description	Unit	Value
7-ton truck emissions	kg CO <sub>2</sub> eq/ton.km	0.2912
7-ton truck energy	MJ/ton.km	1.94286
7-ton truck cost	€/Ton.km	0.125
Containers ship emissions	kg CO <sub>2</sub> eq/ton.km	0.02954
Containers ship energy	MJ/ton.km	0.18034
Containers ship cost (€/Ton.km)	€/ton.km	0.0014
Site distance	km	0
Vestas factory distance to farm location	km	770
Nordx factory distance to farm location	km	930
Goldwind factory distance to farm location	km	900
China- Europe Sea distance	km	23000
Blade truck emissions	kg CO <sub>2</sub> eq/km	2.79529 9
Blade truck energy	MJ/km	18.6991 4
Blade cost	€/km	9.7565
Blade Barge emissions	kg CO <sub>2</sub> eq/ton.km	0.16447
Blade Barge energy	MJ/ton.km	1.04824 6
Blade Barge cost	€/Ton.km	0.09103 5



3. RESULTS AND DISCUSSION

Determining all combinations of options result in 64 possible scenarios combining one option for each stage. Figure 3 presents the total climate change impact, energy intensity and cost of the resulting 64 scenarios.

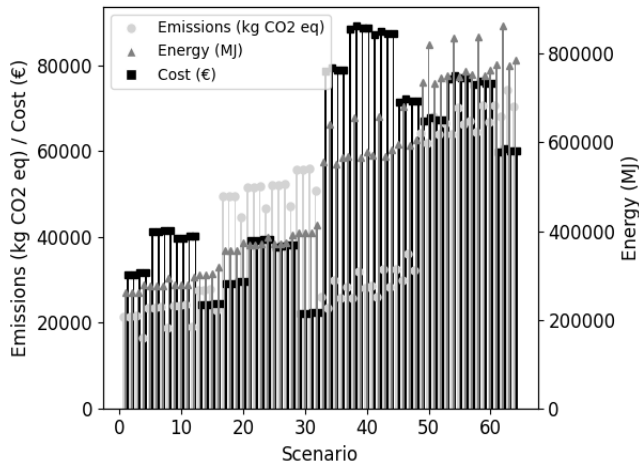


Fig. 3. Emissions, energy intensity and cost per scenario.

The total emissions vary between (16,525 to 74,384) kg CO<sub>2</sub> eq. The lowest emissions come with scenario 4, which represents GFRP as a material sourced from Europe with onsite blade manufacturing and mechanical grinding as a recycling option. The highest value represents CFRP as blade material, with China as material source and manufacturing location in addition to landfilling as end-of-life option.

Total energy intensity calculations fell in the range of (261,179 – 862,661 MJ). The highest energy intensity score is for scenario number 62 which represent a CFRP blade with China as material source and blade manufacturing location, and pyrolysis as end-of-life option. The lowest is for the GFRP blade with material sourced from Europe and site as a blade manufacturing location with repurposing as end-of-life option.

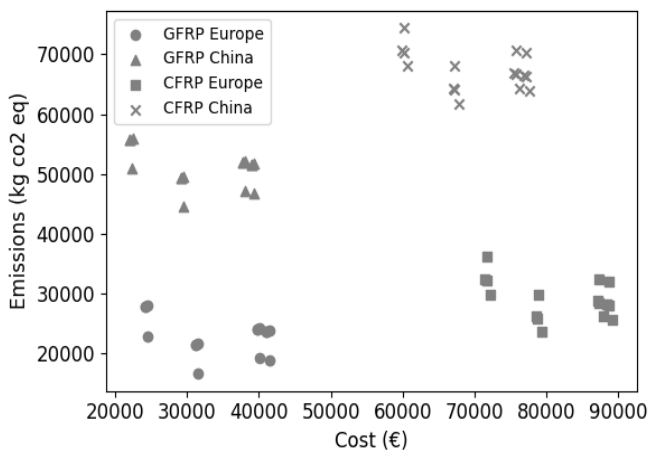


Fig. 4. Scenarios cost versus emissions.

Figure 3 shows that CFRP sourced from Europe with Denmark as a manufacturing location and pyrolysis as end-of-life option involve the highest cost blade (89,159 €), while the lowest cost

(22,115 €) represents a GFRP blade with material source and blade manufacturing location in China and repurposing as an end-of-life option.

The results show variation among the different scenarios with a general trend of high emissions and energy intensity for the CFRP blade where China set as material source, while excessive cost follows the carbon fiber sourced from Europe and blade manufactured in Europe. The high effect of location can be seen on all indicators, this can be attributed to the effect of transportation distance and type, and the effect of the energy mix in each location. The high energy demand, emissions and cost related to carbon fiber manufacturing made it less favorable compared to glass fiber.

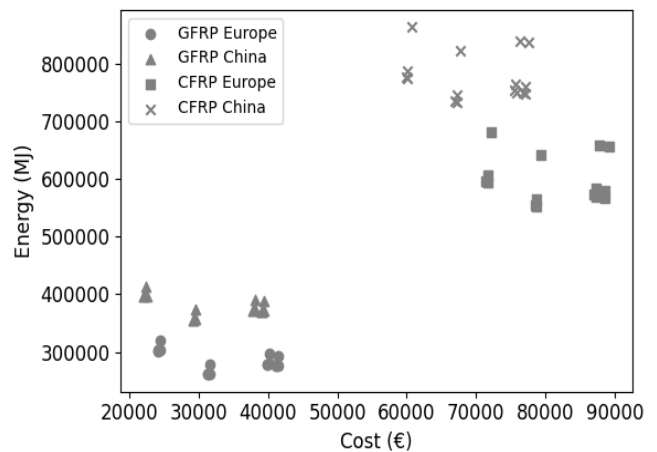


Fig. 5. Scenarios cost versus energy intensity.

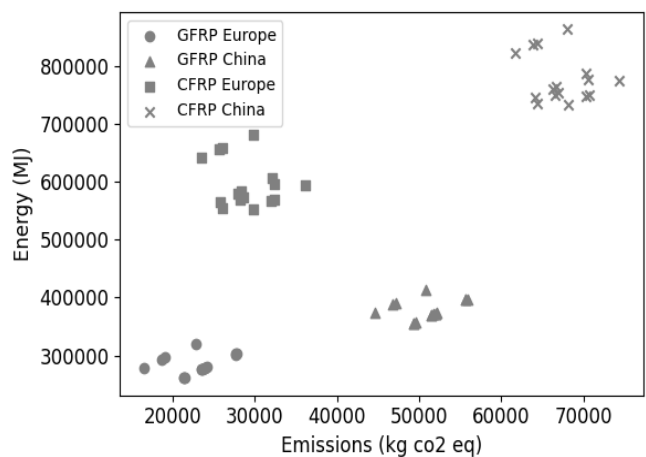


Fig. 6. Scenarios emissions versus energy intensity.

Correlation can be seen between the low cost and low emissions for the scenarios including European sourced GFRP and China as a blade manufacturing location, this can be attributed to the tradeoffs between the low emissions related to the glass fibers manufacturing in Europe and the low cost of labor and electricity in China. As the blade manufacturing processes requires small amount of power (1.57 MW) the effect of high emissions of China grid mix is not significantly affecting the results in this case (Fig. 4).

Correlation between low cost and low energy can be seen for the scenarios including GFRP as material, regardless of the manufacturing location or end of life treatment method, this result driven by the high energy intensity and inflated cost of the CFRP compared to the GFRP (Fig. 5).

Energy intensity and cost relations show higher sensitivity to the material source in the case of CFRP more than for GFRP case (Fig. 5). Europe CFRP represent the high-cost medium energy intensity and China CFRP represent the high energy intensity and medium cost. The high cost and high energy intensity comes with the scenarios linking carbon fiber and Europe manufacturing locations resulting from the high energy intensity of the carbon fibers and the high cost of labor in Europe.

reduce approximately 20% of the fiber glass emissions. The same applies for energy intensity which is affected mostly by the recycling method but with significantly minimal impact compared to the emissions. Most of the recycling methods require energy to perform the recycling and to produce the recovered materials. In addition to the recycling method, we can see the effect of changing the mass of carbon fiber and glass fiber affecting the energy intensity by 1.036 this effect is due to the high energy intensity of the carbon fibers compared to the glass fibers and the equivalent numbers caused by the interchangeability between the two materials as we reduce the glass fibers, we increase the carbon fibers and vice versa.

The minimal sensitivity of the emissions, energy intensity and cost for most variables with consideration of the wide range of

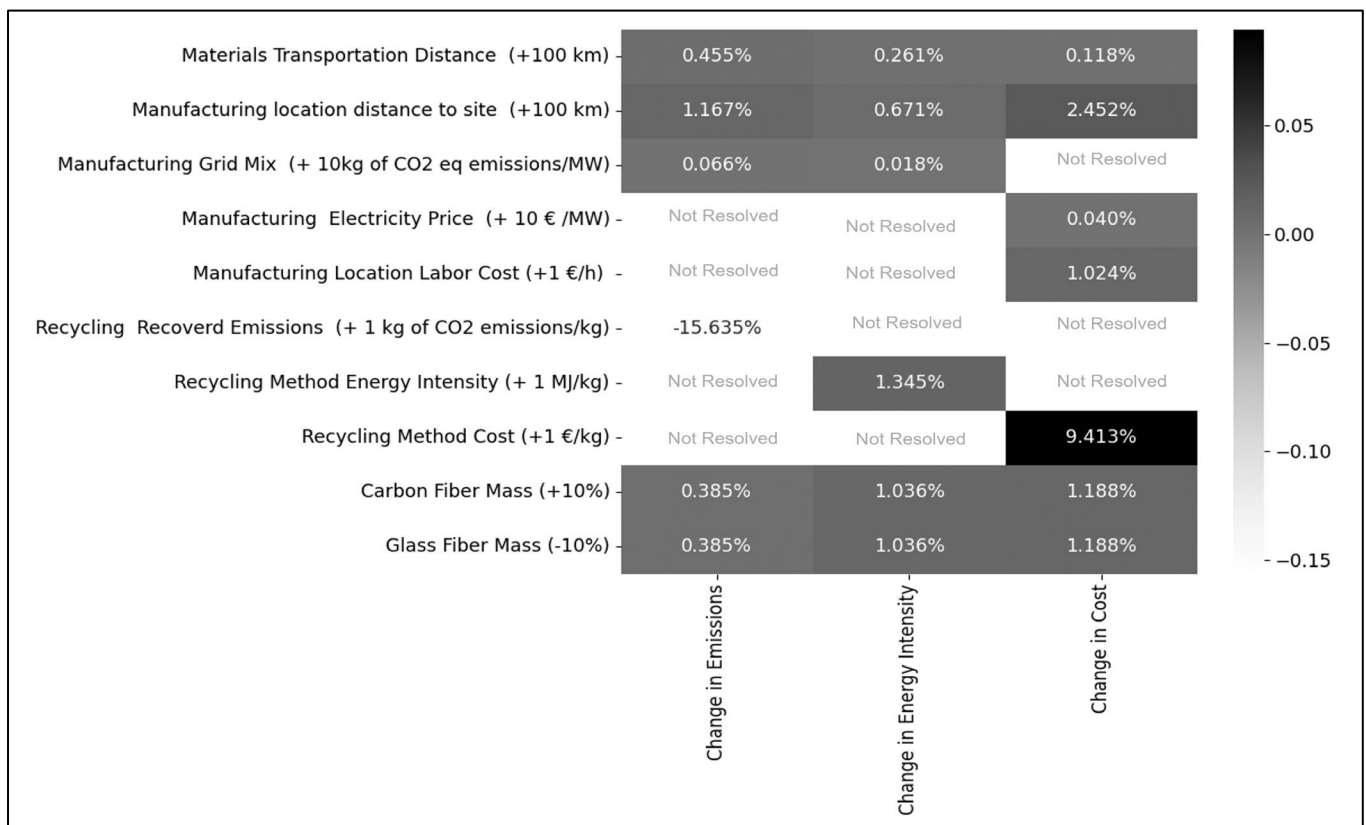


Fig. 7. Emissions, energy intensity and cost sensitivity to variables change.

Figure 6 presents the relation between the emissions and energy intensity of the scenarios. The effect of grid energy mix can be seen in the difference between Europe and China as a materials sources and manufacturing location. The GFRP of European source represent the lowest emissions and energy intensity, while the CFRP of the same origin imposed higher emissions and energy intensity compared to China sourced GFRP but lower emissions and energy compared to Chinese sourced CFRP.

Figure 7 represents the sensitivity analysis results, showing the percentage change in emissions, energy intensity and cost as result of changing one of the variables. Emissions exhibits high sensitivity to the end-of-life treatment method. This can be justified by the high materials emissions per weight compared to the other stages, recovery of 1 kg CO<sub>2</sub> eq /kg

the scenarios results prove the significance of joint effect of changing multiple variables at the same time as each scenario present a unique set of variables values.

Changes in electricity price, labor cost and recycling method cost only affect the total cost as no relation applied between the cost and the other impacts.

#### 4. CONCLUSION

The work presented has demonstrated a logical approach to evaluate several life stage options, which can improve the LCA studies. Furthermore, it has highlighted the importance of composite materials recycling. The study results have proved the magnitude of joint effect of changing several variables on the LCA and LCC studies.

The study results have shown the lowest climate change impact for scenarios 4 (16525 kg CO<sub>2</sub> eq), lowest energy intensity for scenario 1 (261179 MJ), and lowest cost for scenario 29 (22,115 €), while the highest impacts have been the results of scenario 63 (74,384 kg CO<sub>2</sub> eq), scenario 62 (862,661 MJ), and scenario 38 (89,159 €). This has proven that no single scenario can give the lowest or highest impact in all categories and gives room for optimization problem solution.

This work can be a valuable initial step in studying wind turbine blades material sourcing, manufacturing, and recycling.

Future work needs to include more life cycle stages, extra investigation on the interdependency of variables like the electricity mix relation with cost, and modeling different transportations mode and topography.

#### ACKNOWLEDGMENTS

This work has been supported by GEFWIN project (project number 2022-02407) funded by Vinnova under Eurostars/Eureka.

#### REFERENCES

- Bjørn, A., Moltesen, A., Laurent, A., Owsianiak, M., Corona, A., Birkved, M., and Hauschild, M. Z. (2017). Life cycle inventory analysis. In *Life Cycle Assessment: Theory and Practice* (pp. 117–165). Springer International Publishing. doi:10.1007/978-3-319-56475-3\_9
- Bortolotti, P., Berry, D., Murray, R., Gaertner, E., Jenne, D., Damiani, R., Barter, G., and Dykes, K. (2019). A Detailed Wind Turbine Blade Cost Model. [www.nrel.gov/publications](http://www.nrel.gov/publications).
- Brussa, G., Grosso, M., and Rigamonti, L. (2023). Life cycle assessment of a floating offshore wind farm in Italy. *Sustainable Production and Consumption*, 39, 134–144. doi:10.1016/j.spc.2023.05.006
- Cerutti, A., Pant, R., and Sala, S. (2018). Development of a weighting approach for the environmental footprint. European Commission. doi:10.2760/945290
- Corona, A., Markussen, C. M., Birkved, M., and Madsen, B. (2024). Comparative Environmental Sustainability Assessment of Bio-Based Fibre Reinforcement Materials for Wind Turbine Blades. In *Wind Engineering* (Vol. 39, Issue 1). APA.
- Costanzo, G., Brindley, G., Willems, G., Ramirez, L., Cole, P., Klonari, V., and Bickley, J. (2024). Wind energy in Europe 2023 Statistics and the outlook for 2024–2030.
- Ennis, B. L., Kelley, C. L., Naughton, B. T., Norris, R. E., Das, S., Lee, D., and Miller, D. A. (2019). Optimized Carbon Fiber Composites in Wind Turbine Blade Design. <https://classic.ntis.gov/help/order-methods/>
- Garcia-Teruel, A., Rinaldi, G., Thies, P. R., Johanning, L., and Jeffrey, H. (2022). Life cycle assessment of floating offshore wind farms: An evaluation of operation and maintenance. *Applied Energy*, 307. doi:10.1016/j.apenergy.2021.118067
- Hauschild, M. Z. (2017). Introduction to LCA methodology. In *Life Cycle Assessment: Theory and Practice* (pp. 59–66). Springer International Publishing. doi:10.1007/978-3-319-56475-3\_6
- Henao, Y., Grubert, E., Korey, M., Bank, L. C., and Gentry, R. (2024). Life Cycle Assessment and Life Cycle Cost Analysis of Repurposing Decommissioned Wind Turbine Blades as High-Voltage Transmission Poles. *Journal of Construction Engineering and Management*, 150(5). doi:10.1061/jcemd4.coeng-13718
- ILOSTAT, I. labor organization (2022, July 21). Hourly Labour Costs in Manufacturing.
- Jensen, J. P. (2019). Evaluating the environmental impacts of recycling wind turbines. *Wind Energy*, 22(2), 316–326. doi:10.1002/we.2287
- Liu, P., Meng, F., and Barlow, C. Y. (2019). Wind turbine blade end-of-life options: An eco-audit comparison. *Journal of Cleaner Production*, 212, 1268–1281. doi:10.1016/j.jclepro.2018.12.043
- Malcolm, D. J., and Hansen, A. C. (2000). WindPACT Turbine Rotor Design Study: June 2000--June 2002 (Revised). <http://www.osti.gov/bridge>
- Mali, S., and Garrett, P. (2022). Life Cycle Assessment of Electricity Production from an onshore V136-4.2 MW Wind Plant.
- Ozoemena, M., Cheung, W. M., and Hasan, R. (2018). Comparative LCA of technology improvement opportunities for a 1.5-MW wind turbine in the context of an onshore wind farm. *Clean Technologies and Environmental Policy*, 20(1), 173–190. doi:10.1007/s10098-017-1466-2
- Paulsen, E. B., and Enevoldsen, P. (2021). A multidisciplinary review of recycling methods for end-of-life wind turbine blades. *Energies*, 14(14). doi:10.3390/en14144247
- Rani, M., Choudhary, P., Krishnan, V., and Zafar, S. (2021). A review on recycling and reuse methods for carbon fiber/glass fiber composites waste from wind turbine blades. In *Composites Part B: Engineering* (Vol. 215). Elsevier Ltd. doi:10.1016/j.compositesb.2021.108768
- Sala, S., and Cerutti, A. K. (2018). Development of a weighting approach for the Environmental Footprint. doi:10.2760/446145
- Schreiber, A., Marx, J., and Zapp, P. (2019). Comparative life cycle assessment of electricity generation by different wind turbine types. *Journal of Cleaner Production*, 233, 561–572. doi:10.1016/j.jclepro.2019.06.058
- Sproul, E., Williams, M., Rencheck, M. L., Korey, M., and Ennis, B. L. (2023). Life cycle assessment of wind turbine blade recycling approaches in the United States. *IOP Conference Series: Materials Science and Engineering*, 1293(1), 012027. doi:10.1088/1757-899x/1293/1/012027
- Strózyk, M. A., Muddasar, M., Conroy, T. J., Hermansson, F., Janssen, M., Svanström, M., Frank, E., Culebras, M., and Collins, M. N. (2024). Decreasing the environmental impact of carbon fibre production via microwave carbonisation enabled by self-assembled nanostructured coatings. *Advanced Composites and Hybrid Materials*, 7(2). doi:10.1007/s42114-024-00853-2

- Swedish Wind Energy Association (2024). Wind Energy Statistics and forecast.
- van der Meulen, Sander, Grijspaardt, Tom, Mars, Wim, van der Geest, Wouter, Roest-Crollius, Adriaan, and Jan Kiel. (2023). Cost Figures for Freight Transport.
- Warners, J. J., Vouros, S., Kyprianidis, K., Benders, R., and Nienhuis, P. (2023). Future Potential Impact of Wind Energy in Sweden's bidding area SE3.
- Yousef, S., Eimontas, J., Stasiulaitiene, I., Zakarauskas, K., and Striūgas, N. (2024). Recovery of energy and carbon fibre from wind turbine blades waste (carbon fibre/unsaturated polyester resin) using pyrolysis process and its life-cycle assessment. *Environmental Research*, 245. doi:/10.1016/j.envres.2023.118016
- Yuan, W., Feng, J. C., Zhang, S., Sun, L., Cai, Y., Yang, Z., and Sheng, S. (2023). Floating wind power in deep-sea area: Life cycle assessment of environmental impacts. *Advances in Applied Energy*, 9. doi:/10.1016/j.adapen.2023.100122

# Dynamic simulation models in the planning of experiments for control development

Esko K. Juuso\* Luis J. Yebra\*\*

\* *Control Engineering, Environmental and Chemical Engineering, Faculty of Technology, P.O.Box 4300, FI-90014 University of Oulu, Finland, (e-mails: esko.juuso@oulu.fi)*

\*\* *Plataforma Solar de Almería, CIEMAT. 04200 Tabernas, Almería, Spain, (e-mails: luis.yebra@psa.es )*

---

**Abstract:** This paper focuses on the utilization of dynamic simulation models in the planning of experiments for control development. The simulation system is a set of models based on the first principles for system level simulation of the complete TCP-100 research facility at Plataforma Solar de Almería (CIEMAT). This new research facility replaced the 32-year-old ACUREX facility with which so many advances in Automatic Control were reached by the research community. The dynamic models are developed to speed up this research for the new field. The part for control development is the solar field whose parabolic trough collectors (PTCs) are modelled at module level and combined into PTCs and loops. The presented models of the parabolic trough field (PTC) will be validated with experimental data and the controllers will be tested under real conditions. The sequential loops have different operating conditions. This research uses the parameters based on the parameter selection from providers' data sheets and the engineering design project of the TCP-100. The system level model has been implemented in the Modelica language. All state variables are temperatures according to the modelling hypothesis applied, and the inputs of the model are: solar radiation, ambient temperature, setpoints for both circuits pumps, setpoints for two loops control valves, and setpoint for air cooling power. The simulation experiments are first focused on the modules, PTCs and loops of the solar field and the full model need to be extended with dynamic LE models before going to the full simulation tests. In the test campaigns with the new facility, the dynamic LE models are used for planning the test cases.

*Keywords:* nonlinear scaling, uncertainty, dynamic modelling, first principles, simulation, operation of solar PTC plants

---

## 1. INTRODUCTION

Modelling and Control of solar thermal power plants is among the research activities performed at Plataforma Solar de Almería (PSA, PSA-CIEMAT). In the past, active developments of mathematical models and control techniques were done with the ACUREX experimental research facility whose key unit was a parabolic trough collector (PTC) field.

The first modelling and control works were done by R. Carmona, Director of PSA center, in the period from 1985 to 1987. Carmona (1985) defended his dissertation presenting a non-linear distributed mathematical model of the ACUREX field and proposing an adaptive control temperature technique (Camacho et al., 1986). Many control strategies for solar systems have been tested in this facility in its 32 years of life (Camacho et al., 2007; L.Brus et al., 2010; Gallego et al., 2013). Nowadays the TCP-100 facility has replaced ACUREX field and it was specially designed to continue the research activities in Automatic Control, aimed at contributing to the enhancement of the efficiency of this plant technology.

Many parabolic trough collector (PTC) plants have been commissioned in the last 20 years. Only in Spain around 45 PTCs power plants have been setup and more than 26 abroad, built or under construction (PROTERMOSOLAR, 2024). As examples, we can mention the three 50 MW Solnova and the two 50 MW Heliogenery parabolic trough plants of Abengoa in Spain, and the SOLANA and Mojave Solar parabolic trough plant constructed in Arizona and California, each of 280 MW power production.

The main approach followed in the research activities developed so far was to define as control objective the regulation of the outlet temperature of the PTC field around a desired setpoint. These are complementary additional objectives dealing with the automatic start-up, different operating point operation changes and shutdowns of the plant. A previous simulation based analysis of the facility used the nonlinear distributed parameter model presented in (Gallego et al., 2016). A more recent system level dynamic model based on the first principles has been developed and presented in (Pérez et al., 2018). This model provides various possibilities for simulation experiments for developing and validating control solutions. It was used

in (Yebra et al., 2020) for the development of operation training techniques for the TCP-100 facility.

The nonlinear scaling approach has been earlier used for the ACUREX facility (Juuso and Yebra, 2013; Juuso, 2016). The TCP-100 plant has more detailed control possibilities (Fig. 1). This brings new control cases but also makes the tuning more complicated. Simulation models will be used as a replica of the process in the development of controllers.

This paper is organized as follows: Section 2 summarizes the TCP-100 plant. Section 3 focuses on different possibilities to use the first principles simulation model in tuning. Section 4 presents the nonlinear data analysis methodology. Section 5 presents a planning of simulation experiments to be performed for typical operation days. Finally, Section 6 provides some concluding remarks and future works.

## 2. TCP-100 FACILITY

The TCP-100 facility consists of two thermofluid circuits thermally connected by a heat exchanger. This research focuses on the solar field is formed by three PTC loops, each of them composed by two PTCs in a North-South orientation (Fig. 1). Each PTC is 100 meter long, formed by eight modules and all in parallel. Figure 2 shows the first PTC in the first loop.

The solar field is in the primary circuit (Fig. 1). In each loop, the PTCs are connected in the South extreme, and the *colder* PTC will always be the first in the row, placed at the right part of each loop. Each circuit has one tank: the primary tank T-2 with  $10m^3$  volume and the storage tank T-1 in the secondary with a volume of  $115 m^3$ . The pumps for each circuit are placed after both tanks and can be controlled. There is an oil cooler in the secondary circuit.

The other loop, including a storage tank, a cooler and the connecting heat exchanger, may be bypassed during the daily operation to let the control system to choose the operational mode at each time. Operating conditions are chosen with different operation modes:

- (1) Stopped facility. In this mode, both circuits are in stand-by. Both pumps are stopped and the solar field unfocused.
- (2) Both pumps are working and the solar field is unfocused.
- (3) The storage tank is charging with the cooler stopped.
- (4) The storage tank is charging with the cooler working (variable charge).
- (5) The storage tank is discharging.
- (6) The solar field is cooling.

The new solar field provides new remarkable features to its predecessor ACUREX. The main differences among both facilities could be summarized as follows: The ACUREX solar field consisted of 480 East-West aligned single axis tracking PTC forming 10 parallel loops. Each loops was 172 m long, and formed by four 12-module collectors suitably connected in series. The active part of the loop (those parts receiving beam irradiance) measuring 142 m and the passive part (those not receiving beam irradiance)



Fig. 1. Top view of the TCP-100 field at Plataforma Solar de Almería (PSA-CIEMAT). The three loops are shown, with two PTCs in each of them, numbered from 1 (rightmost) to 6 (leftmost). The first loop is formed by the connected pair 1-2 (right loop), the second loop by 3-4 (center loop) and the third by 5-6 (left loop).



Fig. 2. Lateral view of the first PTC in the first loop at Pataforma Solar de Almería (PSA-CIEMAT). It is composed of 8 modules of 12 meters length.

30 m. The HTF used was Therminol 55 thermal oil, capable of supporting temperatures of up to  $300^{\circ}C$ . There were temperature sensors and the inlet and outlet of each loop and the solar field, and the flow rate could be controlled with the pump field. The experimentation of advanced control techniques can utilize the new sensors and actuators installed in the TCP-100 facility summed up in the next. Temperatures are measured in the inlet and outlet of the solar field, the inlet and outlet of each loop, in the inlet and outlet of each PTC and the middle point of each PTC. Volumetric flow rates are measured for each loop. Control valves are used to regulate the mass flow rates in each loop.

## 3. TCP-100 FACILITY MODEL

The simulation studies can use a hybrid (continuous and discrete) system level model based on the first principles model (Pérez et al., 2018). The parameters for that model were obtained from the plant engineering design project data and are also used in this paper. The system level model has been implemented in the Modelica language with the modelling tool Dymola (DassaultSystems, 2018), which applies special algorithms for the manipulation of hybrid models (Mattsson et al., 1999).



After the symbolic manipulations performed by Dymola, the model can be expressed as a general nonlinear state space system in the form

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u}) \tag{1}$$

$$\mathbf{y} = \mathbf{G}(\mathbf{x}, \mathbf{u}) \tag{2}$$

where  $\mathbf{x} \in \mathbf{R}^{28}$ ,  $\mathbf{u} = \{(\mathbf{u}_c, \mathbf{u}_d) \in \mathbf{R}^7 \times \{0, 1\}^3\}$ , and  $\mathbf{y} \in \mathbf{R}^{N_y}$ , where  $N_y$  could be arbitrary chosen from the variables computed in the model detailed in (Pérez et al., 2018). The variables are used in two ways:

**28 state variables ( $\mathbf{x} \in \mathbf{R}^{28}$ ) .**

Each one means a temperature for: each PTC medium control volume (CV), see (Patankar, 1980), each PTC absorber tube, each PTC glass envelope, each tank medium CV, each tank metal walls, each tank isolation layer, each of both medium CVs in the HEX, the HEX metal wall and the air cooler medium CV. All state variables are temperatures according to the modelling hypothesis applied.

**10 input variables ( $\mathbf{u} = \{(\mathbf{u}_c, \mathbf{u}_d) \in \mathbf{R}^7 \times \{0, 1\}^3\}$ ) .**

Seven real input variables ( $\mathbf{u}_c \in \mathbf{R}^7$ ): solar radiation, ambient temperature, setpoints for both circuits pumps, setpoints for two loops control valves, setpoint for air cooling power; and 3 boolean input variables ( $\mathbf{u}_d \in \{0, 1\}^3$ ): bypass activation for the storage tank, for the HEX, and solar field defocusing activation signal.

The Dymola model is capturing the thermal dynamics for the validation of the facility operation modes and operation training purposes as a causal block because of the representation of the inputs and outputs. All the manipulable inputs are shown with the `RealInput` interface component:

- The volumetric flow rates ( $l/s$ ) in control loops for pumps in primary circuit (Syltherm800 medium) and secondary circuit (Therminol55 medium).
- The setpoint for the air cooler cooling power that modulates forced convection.
- The setpoints for both control valves apertures that vary the mass flow rate ( $kg/s$ ) through 1st and 2nd loops.
- The Boolean control input is used in commanding the bypass of the storage tank in the secondary circuit.
- The Boolean control input is used to bypass the HEX, simultaneously in both circuits: primary and secondary.
- `SF_Defocus` is the boolean control input to defocus the solar field. The whole solar field is not reached by any solar irradiance when this signal is activated.

The non-controllable or disturbance inputs are the solar irradiance and the ambient temperature.

The output of the model is a generic output vector `y[:]` that represents in a general form any arbitrary output computed by the model and that could vary from one to another simulation experiment.

Figure 3 shows the summarized model of the TCP-100 facility, that is being acted by a discrete controller implemented with the StateGraph formalism implemented in the Modelica Standard Library. More details about this experiment can be found in (Yebrá et al., 2020).

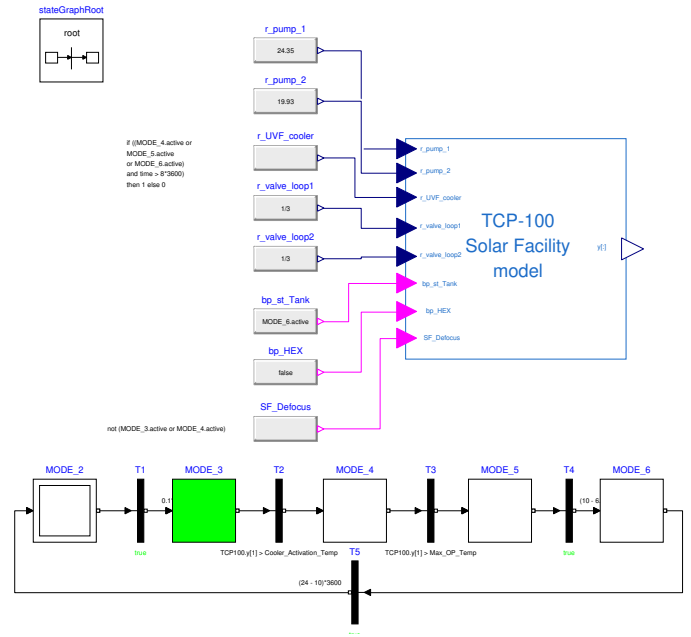


Fig. 3. Modelica model of the TCP-100 facility commanded by a discrete controller implemented with the StateGraph library.

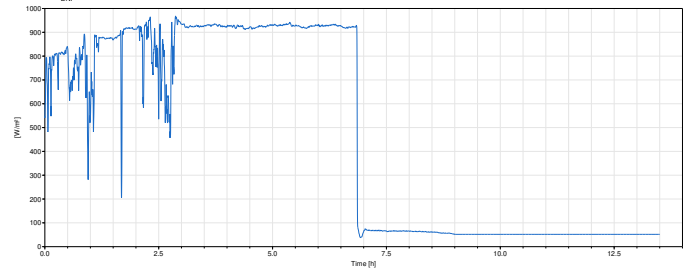


Fig. 4. Direct Normal Irradiance applied in the simulation experiment.

Figures 4 and 5 show respectively the Direct Normal Irradiance (DNI) applied to the model in Figure 3, and the simulated results for the storage tanks in both subcircuits. In the primary circuit with Syltherm800 HTF, the mean temperature of tank T1 is shown when the controller forces the plant to pass through different operation modes. After 5.2 hours from the beginning of the experiment, the T1 temperature rises to its maximum daily value of 316°C. Then, the solar field is defocused and the primary circuit exchanges energy with the secondary, still keeping on charging the T2 tank until it reaches its maximum mean temperature of 219°C at 5.8 hours. During some hours, the system is evolving thermally coupled in the absence of incoming DNI, and at time 8 hours the whole system begins to be cooled. At time 9 hours the bypass of tank T2 is activated, so the cooler is acting over a lower thermal load, which makes the primary circuit cool down to ambient temperature (16°C) at 13.5 hours. In this simulation experiment, the model of the facility has passed through most of the operational modes indicated in Section 2.

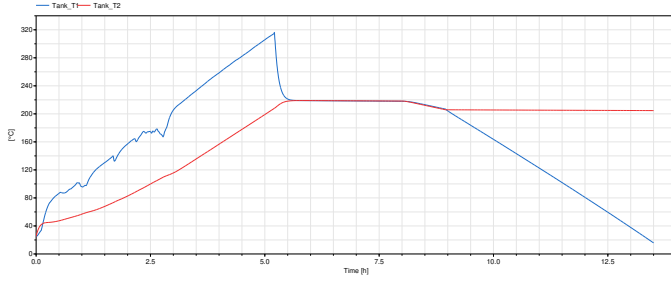


Fig. 5. Simulated temperature profiles for the tanks T1 and T2 in the TCP-100 facility, under the DNI profile in Fig. 4 and StateGraph controller in Fig. 3.

#### 4. NONLINEAR DATA-BASED MODELS

Tests with the previous collector system have shown clear nonlinear behaviour in the normal operating range. The directions of interactions remain constant but the meanings of the variables depend strongly on the operating conditions. In many cases, the nonlinear systems can be implemented with nonlinear scaling and linear interaction models. In the beginning of tuning, the uncertainties need to be taken into account. The representation with natural language is beneficial for understanding and comparing with expert knowledge.

The *energy balance* of the collector field can be represented by expression (Juuso, 2009):

$$I_{eff}A_{eff} = (1 - \eta_p)F\rho cT_{diff}, \quad (3)$$

where  $I_{eff}$  is effective irradiance ( $Wm^{-2}$ ),  $A_{eff}$  effective collector area ( $m^2$ ),  $\eta_p$  a general loss factor,  $F$  flow rate of the oil ( $m^3s^{-1}$ ),  $\rho$  oil density  $kgm^{-3}$ ,  $c$  specific heat of oil ( $Jkg^{-1}K^{-1}$ ) and  $T_{diff}$  temperature difference between the inlet and the outlet ( $^{\circ}C$ ). The effective irradiance is the direct irradiance modified by taking into account the solar time, declination and azimuth. The density decreases and the specific heat increases resulting a nonlinear increase of the term. In the start-up, the flow is limited by the high viscosity.

##### 4.1 Nonlinear scaling

The nonlinear scaling was presented as a methodology for improving membership functions of fuzzy set systems already in (Juuso and Leiviskä, 1992; Juuso, 1992). Nonlinear scaling functions (*NSFs*) are monotonously increasing functions  $x_j = f(X_j)$  where  $x_j$  is the variable and  $X_j$  the corresponding scaled variable in the range  $[-2, 2]$ . The function  $f()$  consist of two second order polynomials, one for the negative values of  $X_j \in [-2, 0]$  and one for the positive values  $X_j \in [0, 2]$ , respectively. Five parameters are needed to define these functions since the overall functions are continuous (Fig. 6). The core area  $[(c_l)_j, (c_h)_j]$ , corresponding  $[-1, 1]$ , is within the support area defined by the minimum and maximum values (Juuso, 2004). The corresponding inverse functions  $X_j = f^{-1}(x_j)$  based on square root functions are used for scaling to the scaled range.

Everything can be defined manually, but it is important to obtain the variable specific parameters of the scaling functions by data-based methodologies. Arithmetical means

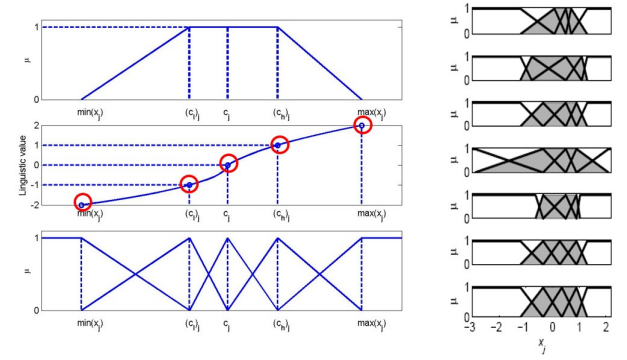


Fig. 6. Nonlinear scaling and membership functions.

and medians were used in Juuso (2004). The current solution uses the central tendency values based on generalised norms (Juuso and Lahdelma, 2010). The generalised norm is defined by

$$\|{}^{\tau}M_j^p\|_p = (M_j^p)^{1/p} = \left[ \frac{1}{N} \sum_{i=1}^N (x_j)_i^{p/1/p} \right], \quad (4)$$

where the order of the moment  $p \in R$  is non-zero, and  $N$  is the number of data values obtained in each sample time  $\tau$ . The norm (4) calculated for variables  $x_j$ ,  $j = 1, \dots, n$ , have the same dimensions as the corresponding variables. The norm  $\|{}^{\tau}M_j^p\|_p$  can be used as a central tendency value if all values  $x_j > 0$ , i.e.  $\|{}^{\tau}M_j^p\|_p \in R$ .

The analysis divides the measurement values into two parts by the point where the skewness changes from positive to negative, i.e.  $\gamma_3^p = 0$ . Then the data set is divided into two parts: a lower part and an upper part. The same analysis is done for these two data sets. The estimates of the corner points,  $(c_l)_j$  and  $(c_h)_j$ , are the points where  $\gamma_3^p = 0$  for the lower and upper data sets, respectively. Since the search of these points is performed by using the order of the moment, the resulting orders  $(p_l)_j$ ,  $(p_0)_j$  and  $(p_h)_j$  are good estimates when additional data sets are used. The orders of the norms help in changes in operating conditions.

##### 4.2 Steady state LE models

Linguistic equation (LE) models consist of two parts: *interactions* are handled with linear equations, and nonlinearities are taken into account by *nonlinear scaling* (Juuso, 1999). In the LE models, the nonlinear scaling is performed twice: first scaling from real values to the interval  $[-2, 2]$  before applying linguistic equations, and then scaling from the interval  $[-2, 2]$  to real values after applying equations. The linguistic level of the input variable  $x_j$  is calculated the inverse functions of the polynomials (Juuso, 2004). More inputs can be included with a steady state LE model represented by

$$x_{out} = f_{out} \left( \frac{\sum_{j=1, j \neq out}^m A_{ij} f_j^{-1}(x_j) + B_i}{A_{i out}} \right) \quad (5)$$

where the functions  $f_j$  and  $f_{out}$  are nonlinear scaling functions of the input variables  $x_j$ ,  $j = 1, \dots, m$  and the output  $x_{out}$ , respectively.



The LE model includes linguistification and delinguistification blocks for the nonlinear scaling of variables. The linear interaction model is used in the equation block which can include a set of equations as well. These blocks are shown in Fig. 7.

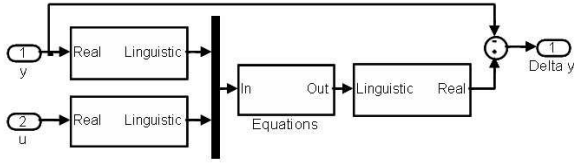


Fig. 7. Blocks of dynamic LE models for calculating  $\Delta y$ .

#### 4.3 Dynamic LE models

*Dynamic LE models* are rather simple input-output models, where the old value of the simulated variable and the current value of the control variable as inputs and the new value of the simulated variable as an output, can be used since nonlinearities are taken into account by nonlinear scaling functions (Fig. 7). For the default LE model, all the degrees of the polynomials in parametric models become very low, i.e. all the parametric models become the same:

$$y(t) + a_1 y(t-1) = b_1 u(t-n_k) + e(t). \quad (6)$$

This model is a special case with three variables,  $y(t)$ ,  $y(t-1)$  and  $u(t-n_k)$ , and a zero bias.

The output, the derivative of the variable  $y$ , is integrated with numerical integration methods:

$$y = \int_0^{T_I} F(t, y, u) dt + y_0, \quad (7)$$

where  $T_I$  is the time period for integration, and  $y_0$  the initial condition. Usually, several values from the integration step or the previous steps are used in evaluating the new value. Step size control adapts the simulation to changing operating conditions.

*Effective time delays* depend on the working conditions (process case), e.g. the delays are closely related to the production rate in many industrial processes. In the block shown in Figure 8, the delay of the variable  $Var1$  depends on the variable  $Var2$ : the linguistic level of the variable  $Var2$  is multiplied by 1 or -1 to get the linguistic level of the delay for the variable  $Var1$ , coefficient 1 means that the delay increases when the variable  $Var2$  increases. The real value of the delay is obtained by the delinguistification block.

Conventional mechanistic models do not work since there are problems with oscillations and irradiation disturbances. In dynamic LE models, the new temperature difference  $\tilde{T}_{diff}(t + \Delta t)$  between the inlet and outlet depends on the irradiance, oil flow and previous temperature difference:

$$\tilde{T}_{diff}(t + \Delta t) = a_1 \tilde{T}_{diff}(t) + a_2 \tilde{I}_{eff}(t) + a_3 \tilde{F}(t), \quad (8)$$

where coefficients  $a_1$ ,  $a_2$  and  $a_3$  depend on operating conditions, i.e. each submodel has different coefficients.

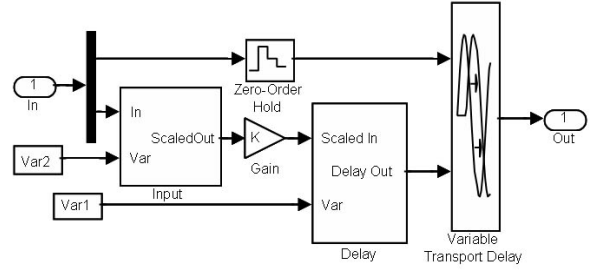


Fig. 8. Time delay of  $Var1$  depends on  $Var2$ .

The nonlinear scaling functions of the outlet temperature does not depend on time. Model coefficients and the scaling functions for  $T_{diff}$ ,  $I_{eff}$  and  $F$  are all model specific. For the ACUREX field, the fuzzy LE system with four operating areas is clearly the best overall model (Juuso, 2003, 2009): the simulator moves smoothly from the start-up mode via the low mode to the normal mode and later visits shortly in the high mode and the low mode before returning to the low mode in the afternoon. Even oscillatory conditions, including irradiation disturbances, are handled correctly. The dynamic LE simulator predicts well the average behaviour but requires improvements for predicting the maximum temperature since the process changes considerably during the first hour. For handling special situations, additional fuzzy models have been developed on the basis of the Fuzzy-ROSA method (Juuso et al., 2000).

#### 4.4 Working point model

The volumetric heat capacity increases very fast in the start-up stage but later remains almost constant because the normal operating temperature range is fairly narrow. This nonlinear effect is handled with the working point LE model

$$wp(i) = \tilde{I}_{eff}(i) - \tilde{T}_{diff}(i), \quad (9)$$

where  $\tilde{I}_{eff}(i)$  and  $\tilde{T}_{diff}(i)$ , which are obtained by nonlinear scaling of variables: efficient irradiance  $I_{eff}$  and temperature difference between the inlet and outlet,  $T_{diff} = T_{out} - T_{in}$ , correspondingly. The outlet temperature  $T_{out}(i)$  is the outlet temperature of the module  $i$ . Since each loop consists of 16 modules, there are 48 sequential modules in the solar field. The outlet temperatures at modules 16, 32 and 48 are controlled.

The working point,  $wp$ , represents a fluctuation from the normal operation. In the normal working point,  $wp = 0$ : the irradiance  $\tilde{I}_{eff}$  and the temperature difference,  $\tilde{T}_{diff}$ , are on the same level. A high working point ( $wp > 0$ ) means low  $\tilde{T}_{diff}$  compared with the irradiance level  $\tilde{I}_{eff}$ . Correspondingly, a low working point ( $wp < 0$ ) means high  $\tilde{T}_{diff}$  compared to the irradiance level  $\tilde{I}_{eff}$ . The normal limit ( $wp_{min} = 0$ ) reduces oscillations by using slightly lower setpoints during heavy cloudy periods. This is not sufficient when the irradiance is high between cloudy periods. Higher limits ( $wp_{min} = 1$ ) shorten the oscillation periods after clouds more efficiently.

## 5. PLANNING OF EXPERIMENTS

The TCP-100 solar thermal power plant replaces the ACUREX experimental research facility. Therefore, the scaling functions of the irradiance ( $W/m^2$ ) do not change which means that also the indicator of the cloudiness remains the same. This is a good starting point for the planning of the experiments. All the other variables are in totally different value ranges.

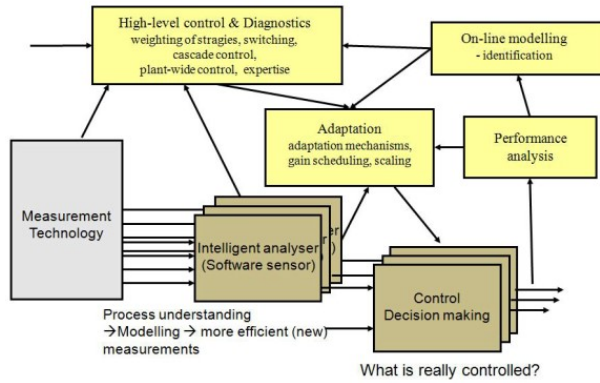


Fig. 9. Modules of the intelligent analyzers and control.

In the LE systems, all variables  $x_j$  are handled with variable specific nonlinear scaling functions:  $x_j = f_j(X_j)$  and  $X_j = f_j^{-1}(x_j)$ , see Section 4. The interactions of the scaled values  $X_j$  are presented with linear equations, like (8). In the starting phase, the scaling functions are available only for the irradiance. The scaling of the temperature differences  $T_{diff}$  is estimated for the loops from the configuration of the field.

The *feedback controller* is a PI-type LE controller with one manipulating variable, oil flow, and one controlled variable, the maximum outlet temperatures of the loops. The PI-type means that the change of control is the sum of the error and the change of error. The acronym LE means that dimensionless scaled values are used in the control equation. The very compact basic controller provides a good basis for advanced extensions: the scaling functions can be to versatile operating conditions, and control equations can be extended from the PI type with different algorithms, e.g. all different types of PID controllers could be used. The blocks are the same for different modules and loops.

The LE controller contains several parametric scaling functions for variables, errors, changes and corrections. Since there is no actual test data available from the new research plant, the parameters are chosen before the test campaign by using previous test results from the ACUREX plant and adjusting or scaling them to correspond better to the specifications of the new TCP-100 collector field or using the simulator of the new field. These properties are tuned after the test campaign.

*Intelligent analysers* have been used for detecting changes in operating conditions to activate adaptation and model-based control and to provide indirect measurements for the high-level control (Fig. 9). There are many improvements, which are planned to be introduced to the new TCP-100 facility. For the first tests, the intelligent analysers are not used. The correction factors based on the working

point value  $wp$  are utilized in the adaptation of the LE control. The fine-tuning with the predictive braking and asymmetrical actions are left for later studies.

The *working point*  $wp$  is important in both in this study and the final system since the model-based control limits the acceptable range of the temperature setpoint by using the chosen working point (Fig. 9). The fluctuation indicators are used for modifying these limits to react better to cloudiness and other disturbances. The manual setpoints are used only within these limits. Dynamic models developed for the TCP-100 facility could be used for development in this task.

*High-level control* is aimed for manual activating, weighting and closing different actions. As there are many actions, this is needed to run the tests efficiently. These ideas will be developed interactively during the test campaigns to provide a basis for the performance analysis and integration of expertise (Fig. 9).

The full *first principles model* is highly complicated and a lot of tuning work is needed before it can be used in tuning the controller for the special cases listed above. Actually, the simulator would already need adaptive parts. A better way is to focus first on the PTC loops (Fig. 2). There are three loops in the solar field (Fig. 1) which all consist of two PTCs both having eight similar modules. These can be handled with the same parametric LE model. The modules are working in different operating conditions: the input and output temperatures depend on the sequence of the modules. The control of the loops introduces additional differences between the loops.

The project will then continue first with the full dynamic models enhanced with the new LE models. Then the full set of the experiments can be started in the real new TCP-100 facility which finally provides the data which can be used in the tuning of the plant and the control system. The simulation studies provide a starting point for the test campaigns with the new field. The parameters will be updated offline during the test days by using the recursive approach. The TCP-100 facility includes more units, loops and connections. There are more sensors for the temperatures and volumetric flows. The control is available in each of the loops. The dynamic simulation model includes 28 state variables and seven input variables (Section 3).

The dynamic simulation model is used as a plant in this research. This is a flexible solution for analysing different weather conditions and disturbances. The strongly fluctuating situations are difficult to handle reliably with models. However, they can be taken as scenarios in this model based analysis. The idea of the nonlinear scaling is that the algorithms remain unchanged.

## 6. CONCLUSIONS AND FUTURE RESEARCH

This research focuses on starting to apply the intelligent models and control algorithms for the new TCP-100 solar thermal plant. The scaling functions of the ACUREX facility remain unchanged for the irradiance which also means that the earlier indicators of the cloudiness can be used. The new facility includes more units, loops and connections. The algorithms are not changed and the

data analysis can be done by using the dynamic models for a limited set of measurements and subsystems. This research means preliminary simulation experiments. The first principles models would require adaptive parts before going the full simulation experiments. The work can be started with the loops and modules by using parametric linguistic equation models. The simulation studies will be extended with these models before going to the test campaigns with the new facility.

#### ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support by the Ministerio de Ciencia, Innovación y Universidades of Spain for the grants TED2021-129189B-C21/TED2021-129189B-C22 (MODIAG-PTC) and PCI2022-134974-2 (DISOPED), funded by MCIN/AEI/10.13939/501100011033 and by the European Union "NextGenerationEU/PRTR", and the Co-Innovation joint project Highly Optimized Energy Systems (HOPE) funded by Business Finland (grant number 3364/31/2020).

#### REFERENCES

- Camacho, E.F., Carmona, R., and Rubio, F.R. (1986). *Adaptive control of the ACUREX field*. Springer Verlag, London.
- Camacho, E., Rubio, F., Berenguel, M., and Valenzuela, L. (2007). A survey on control schemes for distributed solar collector fields. part I: Modeling and basic control approaches. *Solar Energy*, 81, 1240–1251. doi:10.1016/j.solener.2007.01.002.
- Carmona, R. (1985). *Análisis, Modelado y control de un campo de colectores solares distribuidos con sistema de seguimiento en un eje*. Ph.D. thesis, Universidad de Sevilla.
- DassaultSystems (2018). Dymola User Manual.
- Gallego, A.J., Fele, F., Camacho, E.F., and Yebra, L.J. (2013). Observer-based model predictive control of a solar trough plant. *Solar Energy*, 97, 426–435. doi:10.1016/j.solener.2013.09.002.
- Gallego, A.J., Yebra, L.J., Camacho, E.F., and Sánchez, A.J. (2016). Mathematical modeling of the parabolic trough collector field of the tcp-100 research plant. In *9th EUROSIM Congress on Modelling and Simulation*. 12-16 September (Finland).
- Juuso, E. and Lahdelma, S. (2010). Intelligent scaling of features in fault diagnosis. In *7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2010 - MFPT 2010, 22-24 June 2010, Stratford-upon-Avon, UK*, volume 2, 1358–1372. BINDT. ISBN=978-1-61839-013-4.
- Juuso, E.K. (1992). Linguistic equation framework for adaptive expert systems. In J. Stephenson (ed.), *Modelling and Simulation 1992, Proceedings of the 1992 European Simulation Multiconference, York, UK, June 1-3, 1992*, 99–103. SCS International, San Diego, USA.
- Juuso, E.K. (1999). Fuzzy control in process industry: The linguistic equation approach. In H.B. Verbruggen, H.J. Zimmermann, and R. Babuška (eds.), *Fuzzy Algorithms for Control, International Series in Intelligent Technologies*, volume 14 of *International Series in Intelligent Technologies*, 243–300. Kluwer, Boston. doi:10.1007/978-94-011-4405-6\_10.
- Juuso, E.K. (2003). Intelligent dynamic simulation of a solar collector field. In A. Verbraeck and V. Hlupic (eds.), *Simulation in Industry, 15th European Simulation Symposium ESS 2003*, 443–449. SCS, Gruner Druck, Erlangen, Germany.
- Juuso, E.K. (2004). Integration of intelligent systems in development of smart adaptive systems. *International Journal of Approximate Reasoning*, 35(3), 307–337. doi:10.1016/j.ijar.2003.08.008.
- Juuso, E.K. (2009). Dynamic simulation of solar collector fields in changing operating conditions. In B. Elmegaard, C. Veje, M.P. Nielsen, and T. Mølbak (eds.), *Proceedings of SIMS 50 - the 50th International Conference of Scandinavian Simulation Society, October 7-8, Fredericia, Denmark*, 341–348. DTU, Lungby, Denmark.
- Juuso, E.K. (2016). Intelligent control of a solar thermal power plant - adaptation in varying conditions. In *2016 5th International Conference on Power Science and Engineering, Venice, Italy, 14-17 December 2016*.
- Juuso, E.K. and Leiviskä, K. (1992). Adaptive expert systems for metallurgical processes. *IFAC Proceedings Volumes*, 25(17), 119–124. doi:10.1016/B978-0-08-041704-2.50027-3.
- Juuso, E.K., Schauten, D., Slawinski, T., and Kiendl, H. (2000). Combination of linguistic equations and the fuzzy-ROSA method in dynamic simulation of a solar collector field. In L. Yliniemi and E. Juuso (eds.), *Proceedings of TOOLMET 2000 Symposium - Tool Environments and Development Methods for Intelligent Systems, Oulu, April 13-14, 2000*, 63–77. Oulun yliopistopaino, Oulu.
- Juuso, E.K. and Yebra, L.J. (2013). Optimisation of solar energy collection with smart adaptive control. In *IECON Proceedings (Industrial Electronics Conference), 10-14 November, 2013, Vienna, Austria*, 7938–7943. doi:10.1109/IECON.2013.6700459.
- L.Brus, T.Wigren, and D.Zambrano (2010). Feedforward model predictive control of a non-linear solar collector plant with varying delays. *IET Journal of Control Theory and Applications*, 4 (8), 1421–1435. doi:10.1049/iet-cta.2009-0316.
- Mattsson, S.E., Otter, M., and Elmqvist, H. (1999). Modelica hybrid modeling and efficient simulation. In *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No.99CH36304)*, volume 4, 3502–3507.
- Patankar, S.V. (1980). *Numerical Heat Transfer and Fluid Flow. Series in Computational and Physical Processes in Mechanics and Thermal Sciences*. Taylor & Francis, London, UK.
- Pérez, J., Yebra, L.J., Dormido, S., and Zarza, E. (2018). First Principles System Level Modelling of TCP-100 Facility for Simulation of Operation Modes. *IFAC-PapersOnLine*, 51(2), 481–486. doi:10.1016/j.ifacol.2018.03.081.
- PROTERMOSOLAR (2024). Spanish association for the promotion of the thermosolar industry. URL <http://www.protermosolar.com/>.
- Yebra, L.J., Marquez, F.M., and Zufria, P.J. (2020). Simulation of tcp-100 facility system level model for operation training purposes. 1348–1353. IEEE. doi:10.1109/CSCI51800.2020.00251. URL <https://ieeexplore.ieee.org/document/9457861/>.

# Experimental and Numerical Testing of a Multi-Modular Floating Structure with Varying Connection Stiffness

Trine Aas-Hansen\* Vegard Njøten Fagerbakke\*\*  
Trygve Kristiansen\*\* Svein Sævik\*\* Dong Trong Nguyen\*\*\*

\* Norwegian University of Science and Technology, Trondheim, Norway (e-mail: trine.aas-hansen@ntnu.no).

\*\* Norwegian University of Science and Technology, Trondheim, Norway

\*\*\* Norwegian University of Science and Technology, Trondheim, Norway (e-mail: dong.t.nguyen@ntnu.no).

---

## Abstract:

This work is a step towards conceptualizing a smart multi-modular structure, whose main application is solar energy harvest, with the innovative idea of connectors that can be controlled to mitigate motions and loads in a changing environment. The paper presents selected preliminary results from experimental tests of an array of floating column-based modules exposed to regular waves of different periods. Each pair of neighboring modules was connected by two spring connectors with both tension and compression stiffness. The paper presents an investigation of motion responses versus load frequencies corresponding to four tested spring stiffnesses.

The model test results serve as a basis for validating a numerical model that is implemented for control design and simulation purposes. Wave-, mooring- and connector forces are considered in the simulations. The proposed method will act as a tool for further evaluation of the effect of changing the connection stiffness according to the incoming waves and the investigation of whether it is beneficial to apply a smart connector that can adapt to varying sea states.

*Keywords:* Renewable Energy, Floating Solar, Model Testing, Numerical Modelling, Control

---

## 1. INTRODUCTION

### 1.1 Background

The world demands more green energy to reduce the global carbon emissions. To replace polluting energy sources, a large variety of sustainable solutions is essential. According to the Intergovernmental Panel on Climate Change, solar photovoltaic (PV) has become a competitive energy source as the cost has decreased by 85 % between 2010 and 2019 (IPCC, 2023). In areas where solar irradiation is abundant, but available land areas are in high demand, building floating solar plants could be a valuable contribution towards the energy transition. In addition to the advantage of not using land areas, Kumar et al. (2021) summarized benefits of floating solar, including better cooling effect, reduction of water evaporation and less accumulation of dust.

If floating large-area structures can be designed to sustain higher environmental loads in an exposed or offshore environment, new solutions for floating photovoltaic (FPV) power plants can be investigated. DNV (2022) listed the

\* The PhD position of the first author is funded by Department of Marine Technology at NTNU. This work was partly supported by the Research Council of Norway through SFI BLUES, grant number 309281.

main opportunities for FPV as; making use of abundant solar energy in more areas; maximizing the use of space and existing infrastructure by combining FPV with for instance offshore wind installations; provide offshore charging for electric marine vessels or for production of alternative fuels; supplying green energy to islands or maritime industry; and finally, the reuse of competence from other marine industries.

Placing floating structures in harsher offshore environments with wind, current and changing wave states is not a new research topic. The oil and gas industry has designed and built offshore installations for decades. Ideas for very large floating structures (VLFSs) explores making large areas available on water surfaces for e.g. buildings, floating ports, airports and agriculture (Lamas-Pardo et al., 2015). To reduce bending moments on these large-area structures, connecting several smaller modules together to make a flexible structure that is allowed to move with the changing sea surface could be a solution. Such multi-modular structures, in a smaller scale than the VLFSs, will in this paper be investigated for the purpose of solar harvest in exposed areas.

A nonlinear model of connected floating modules using network theory has been used to analyze the response and connection loads of a multi-modular structure (Zhang

et al., 2015). Shi et al. (2018) validated the network modelling method by experimental testing of three modules in an array, and Ding et al. (2021) by full scale results from the Scientific Research and Demonstration Platform in the South China Sea.

Multi-modular structures introduce connection points which present weak links in terms of fatigue life and durability. This motivates research on how to reduce relative module motion and loads in connections when such structures are exposed to changing environmental conditions. The motion response and oscillations as well as connector loads of a multi-modular structure were shown by Jiang et al. (2021) to be affected by the connector stiffness between modules. By adapting this stiffness to different wave periods, studies have shown that it is possible to retain the structure in a state where oscillations are kept at a minimum, also known as amplitude death state (Xu et al., 2014; Xia et al., 2016; Zhang et al., 2017). By actively controlling their stiffness, the connectors themselves can be used as actuators.

This paper investigates a type of multi-modular structure with modules larger than the components of typical floating solar rafts, but smaller than the modules in very large floating structures. Further, this paper contributes to research by including varying connector stiffness in both numerical and experimental tests. The main objectives are to analyze the behavior of different configurations of multi-modular structures, investigate the effect of different connector stiffness and to validate a numerical simulator by comparable experimental results. The motivation for the work is to prepare for development of a control system that changes the dynamics of the structure as a response to the current sea state by using active control of the stiffness in connection points. The work utilizes existing numerical and experimental modelling techniques on a new type of structure and, by creating a modeling framework, contributes towards an evaluation of the idea of creating smart multi-modular structures that adapt to changing environmental conditions.

## 1.2 Paper Outline

The paper starts with Sect. 2 presenting a test case with necessary parameters for both the experimental and simulator setup. It describes how the test case is adapted for model testing and introduces the mathematical modelling of the dynamical system. Results are presented in Sect. 3 and concluding remarks are given in Sect. 4.

## 2. CASE STUDY

### 2.1 Test Case

A common test case is used in both the experimental and numerical investigation for easy comparison. This case is based on an offshore floating solar power concept, utilizing the same modules as used in previous model tests (Onsrud, 2019). Square rigid platforms with four cylindrical floaters or columns (Fig. 1) are flexibly connected to form an interconnected large-area structure. Arrays of 1, 2, 3 and 5 modules are subjected to regular waves in the longitudinal direction and are tested with varying stiffness

in the connection points. Table 1 presents the full-scale parameters of the test case.

Table 1. Test case parameters (full scale)

Parameter			Unit
Module size	$L \times W$	12×12	[m]
Module mass	$m$	9088	[kg]
Draft	$d$	1.13	[m]
Column radius	$r$	0.8	[m]
Column height	$h$	2.63	[m]
CG	$(x_g, y_g, z_g)$	(0,0,-1.3)	[m]
CF	$(x_f, y_f, z_f)$	(0,0,0)	[m]
CB	$(x_b, y_b, z_b)$	(0,0,d/2)	[m]
Module distance	$\Delta x$	1.3	[m]
No. of modules	$N$	1-5	[-]
No. of mooring lines	$N_m$	4	[-]
No. of connectors between neighbors	$N_c$	2	[-]
Mooring line stiffness	$K_m$	2800	[N/m]
Connector stiffnesses:	$K_c$	6-29	[kN/m]
- Conn. stiffness S1		29000	[N/m]
- Conn. stiffness S2		26000	[N/m]
- Conn. stiffness S3		15000	[N/m]
- Conn. stiffness S4		6000	[N/m]
Wave period	$T_w$	2.0-8.0	[s]
Wave steepness	$H/\lambda$	1/116	[-]
Water density	$\rho$	1000	[kg/m <sup>3</sup> ]

Table 2. Run overview

Parameter	Variations
Module configurations:	1×1, 2×1, 3×1, 5×1
Connector stiffness:	S1, S2, S3, S4, Hinge*
Wave period:	2.0-8.0 s

\*Only model test with configuration 2×1 and 5×1

### 2.2 Model Test

The model tests were performed in a towing tank at NTNU in Trondheim, Norway. The tank is 2.8 by 25 meters, with a water depth of 0.7 meters. It has a wave maker and a wave beach to absorb energy from the waves. The model was Froude scaled by 1:20 according to the parameters presented in Table 1. Figure 2 shows a photo of the model test setup for the three-module configuration in the towing tank where the experiments were performed.

The first and last module in each test configuration were moored in four corners by pre-tensioned springs (Figs. 1 and 2). All modules are thus kept in a neutral position in calm water. Each pair of modules was connected by two spring connectors (Figs. 3 and 4) functioning with both tension and compression stiffness. Tests were performed according to Table 2. Each combination of connector type and number of modules was exposed to regular waves travelling in the negative  $x$ -direction. The wave period was limited by the water depth in the tank, ensuring linear wave behavior.

The response of each module was tracked in six degrees of freedom (DOFs) using a camera-based motion capture system. Load cells were used to measure the force in local  $x$ -direction in one spring of each connection pair, see Fig. 3. The wave height was measured by wave probes at 8 different locations throughout the tank.



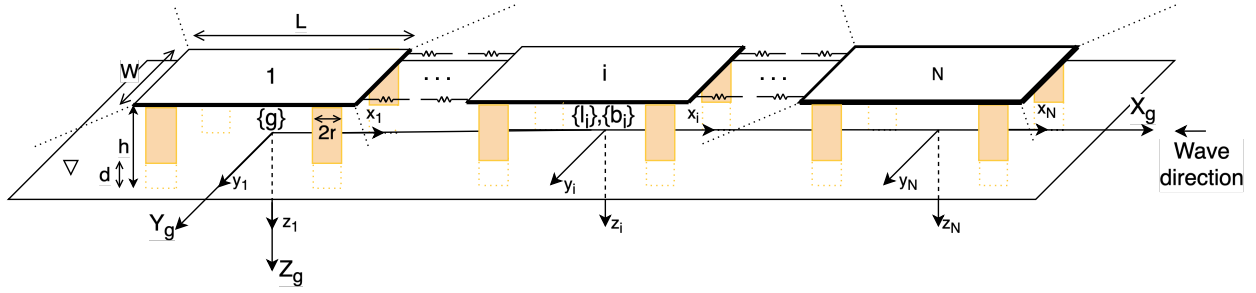


Fig. 1. General overview of the test case. Module numbering and global reference frame  $\{g\} = (X_g, Y_g, Z_g)$ , as well as selected parameters from Table 1 are presented. The local  $\{l_i\}$  and body-fixed  $\{b_i\}$  reference frames coincides when the modules are in neutral position. Further, connection springs are indicated between modules, and mooring lines are indicated as dotted lines in the corners of the end modules.

The results from this model test have also been evaluated in a master thesis from NTNU by Fagerbakke (2023).

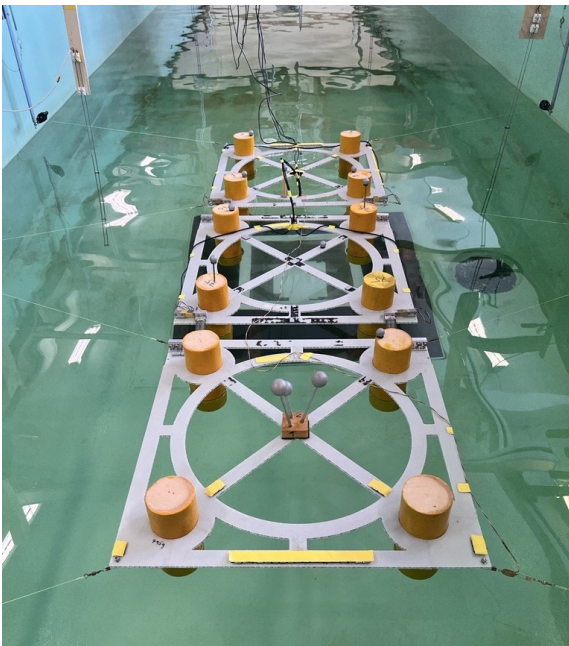


Fig. 2. Model test setup,  $N=3$ .

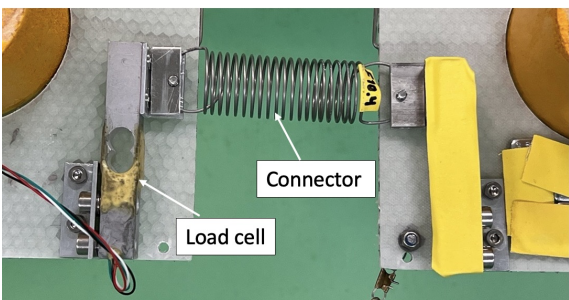


Fig. 3. Model test connector and load cell.

### 2.3 Modelling of the Test Case

The numerical model is based on a mathematical description of rigid modules with flexible connections. The main assumptions for the numerical model are:

*Assumption 1.* The modules are assumed rigid.

*Assumption 2.* Waves are modelled as linear deep-water waves, unaffected by shallow-water effects

*Assumption 3.* Motions in  $xy$ -plane are assumed small, thus reducing the analysis to a 2D-problem.

*Assumption 4.* Connector springs between the modules are limited to axial forces.

*Assumption 5.* The effect of finite column length is neglected.

*Assumption 6.* Modules are allowed to overlap if motions are large, collisions are not modelled.

*Reference Frames:* Three types of reference frames were used to model the multi-modular island, see Fig. 1. A global inertial reference frame defined by  $\{g\} = (X_g, Y_g, Z_g)$ , where  $X_g$  and  $Y_g$  is zero in the initial position of the geometrical center of the first module, and  $Z_g$  is zero at the water surface, with positive axis downwards. A local coordinate system  $\{l_i\} = (x_i^l, y_i^l, z_i^l)$  has its origin in the neutral position of each module. In addition, each module has its own body-fixed reference frame,  $\{b_i\} = (x_i^b, y_i^b, z_i^b)$ , also with the positive  $z$ -axis down from the sea surface.

The position and orientation vector for a 6 DOF module  $i$ , in its local reference frame is defined as  $\eta_i = [\mathbf{p}_i^T, \Theta_i^T]^T$ , where  $\mathbf{p}_i = [x_i, y_i, z_i]^T$  represent the module's  $x$ -,  $y$ - and  $z$ -position in  $\{l_i\}$ , and  $\Theta_i = [\phi_i, \theta_i, \psi_i]^T$  represent the orientation of  $\{b_i\}$  in  $\{l_i\}$  in roll, pitch and yaw respectively. The velocity in the body-fixed reference frame of each module is  $\nu_i = [u_i, v_i, w_i, p_i, q_i, r_i]^T$  representing translational velocities in  $x$ ,  $y$  and  $z$  ( $u_i, v_i, w_i$ ), and rotational velocities around these axes ( $p_i, q_i, r_i$ ).

The transformation between two reference frames is described by  $\tilde{\eta}_i = \mathbf{J}(\Theta_i)\nu_i$  with the transformation matrix  $\mathbf{J}(\Theta)$  being a block diagonal matrix defined by the transformation of translations,  $\mathbf{R}(\Theta)$ , and rotations,  $\mathbf{T}(\Theta)$ , according to Fossen (2021).

*Equation of Motion:* The equation of motion for module  $i$  in its body-fixed frame  $\{b_i\}$  is given by

$$\mathbf{M}_i(\omega)\dot{\nu}_i + \mathbf{D}_i(\omega)\nu_i + \mathbf{J}^T(\Theta_i)\mathbf{C}_i\eta_i = \mathbf{F}_i^b(\omega), \quad i = 1, \dots, N \quad (1)$$

$\mathbf{M}_i(\omega)$  includes the rigid body mass matrix  $\mathbf{M}_{RB,i}$  and hydrodynamic added mass  $\mathbf{A}_i(\omega)$ . The force vector  $\mathbf{F}_i^b$  is the sum of connector forces  $\mathbf{F}_{c,i}^b$ , mooring forces  $\mathbf{F}_{m,i}^b$ , and wave loads  $\mathbf{F}_{w,i}^b(\omega)$ .  $\mathbf{D}(\omega)_i$  and  $\mathbf{C}_i$  represents potential damping and restoring matrices respectively.

*Wave Loads:* The wave loads are modeled as potential linear excitation loads,  $\mathbf{F}_{pot.,i}^b(\omega)$ , and hydrodynamic drag,  $\mathbf{F}_{morison,i}^b(\omega)$ . The potential wave loads can be expressed as

$$\mathbf{F}_{pot.,i}^b = \bar{\mathbf{F}}_{pot.,i}(\omega) \cos(\omega t - kx_{lg,i}^g + \alpha_i) \quad (2)$$

The amplitude,  $\bar{\mathbf{F}}_{pot.,i}(\omega)$ , and phase angle,  $\alpha_i(\omega)$ , are calculated with help from WADAM (DNV-GL, 2017).  $k$  is the wave number, defined by  $k = \omega^2/g$ , where  $g$  is the gravity acceleration.  $x_{lg,i}^g$  is the position of the local reference frame of the module in the global reference frame. To account for viscous effects, drag loads are added through Morison's equation in vertical and horizontal directions. Drag forces and moments are calculated using strip theory and the crossflow principle for each column  $i,p$ :

$$\begin{aligned} \mathbf{f}_{morison,i,p}^b &= \int \frac{1}{2} \rho C_D 2r |\mathbf{u}_{r,i,p}| \mathbf{u}_{r,i,p} dz \\ \mathbf{m}_{morison,i,p}^b &= \mathbf{r}_{i,p} \times \mathbf{f}_{morison,i,p}^b \end{aligned} \quad (3)$$

where  $C_D$  is the drag coefficient,  $\mathbf{u}_{r,i}$ , is the relative velocity vector on the column and  $\mathbf{r}_{i,p}$  is the distance between the body frame origin and the force application point. The total drag load is the sum of loads on all  $N_p$  columns on the module:

$$\mathbf{F}_{morison,i}^b = \sum_{p=1}^{N_p} \begin{bmatrix} \mathbf{f}_{morison,i,p}^b \\ \mathbf{m}_{morison,i,p}^b \end{bmatrix} \quad (4)$$

*Connector Loads:* The moving neighboring modules will impose forces on each other based on the distance between connection points and the spring stiffness of the connector. Following Assumptions 3 and 4, forces occurring from movement along the transversal and vertical direction (i.e. along  $y$  and  $z$ ) are neglected, thus the connector forces are a result of movement in  $x$ -direction. The numbering of the four connectors of module  $i$  is presented in Fig. 4.

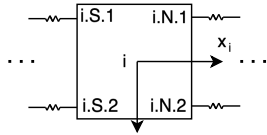


Fig. 4. Connectors 1-4 on module  $i$ . N indicates the north face of the module, S indicates the south face.

The total connection force on module  $i$  is given by

$$\mathbf{F}_{c,i}^b = \sum_{j=1}^N \sum_{\substack{m \in \mathcal{M} \\ o \in \mathcal{O}}} \Phi_{m,i,j} \mathbf{K}_c(\mathbf{p}_i, \mathbf{p}_j)_{m,o}, \quad \begin{cases} \mathcal{M} & = [\text{N}, \text{S}] \\ \mathcal{O} & = [\text{S}, \text{N}] \end{cases} \quad (5)$$

where  $\Phi_m \in \mathbb{R}^{N \times N}$  represent the topology matrices for the  $m$  face of the modules.  $\Phi_{N,i,j}$  is 1 if the north face of module  $i$  is connected to module  $j$ , and 0 if not. Further,  $\mathbf{K}_c(\mathbf{p}_i, \mathbf{p}_j)_{m,o}$  is given by

$$\mathbf{K}_c(\mathbf{p}_i, \mathbf{p}_j)_{m,o} = \sum_{k=1}^{N_c} \begin{bmatrix} \mathbf{f}_c^b(\mathbf{p}_i, \mathbf{p}_j)_{m,o,k} \\ \mathbf{r}_{c,i,m,k} \times \mathbf{f}_c^b(\mathbf{p}_i, \mathbf{p}_j)_{m,o,k} \end{bmatrix} \quad (6)$$

where

$$\mathbf{f}_c(\mathbf{p}_i, \mathbf{p}_j)_{m,o,k} = -K_c(\mathbf{p}_{i,m,k}^l - \mathbf{p}_{j,o,k}^l - \mathbf{p}_0) \quad (7)$$

Simplified to consider motion in  $x$ -direction only:

$$\mathbf{K}_c(\mathbf{p}_i, \mathbf{p}_j)_{m,o} = -K_c \sum_{k=1}^{N_c} [(x_{i,m,k}^g - x_{j,o,k}^g - \delta x), 0, 0, 0, 0, 0]^\top \quad (8)$$

where  $\delta x$  is the neutral distance between modules in  $x$ -direction.

*Mooring Loads:* Mooring stiffness is implemented with only one connection point moving, and the other fixed in a simulated anchoring point. The force and moment from mooring line  $k_m$  on module  $i$  are given by

$$\mathbf{f}_{m,i,k_m}^l = (K_m(\|\mathbf{l}_{i,k_m}\| - L_{k_m}) + f_{pt,k_m}) \frac{\mathbf{l}_{i,k_m}}{\|\mathbf{l}_{i,k_m}\|} \quad (9)$$

$$\mathbf{m}_{m,i,k_m}^b = \mathbf{r}_{m,i,k_m}^b \times \mathbf{J}^\top(\Theta_i) \mathbf{f}_{m,i,k_m}^l$$

where  $f_{pt,k_m}$  is the pretension in the mooring line,  $\mathbf{l}_{i,k_m}$  is the vector describing the relative position of mooring point and the anchoring point,  $L_{k_m}$  is the initial length of mooring line  $k_m$  and  $\mathbf{r}_{m,i,k_m}^b$  is the lever arm to the point where mooring line  $k_m$  is connected to module  $i$ .

The total mooring load on module  $i$  is the sum of forces and moments from all  $N_m$  mooring lines connected to it:

$$\mathbf{F}_{m,i}^b = \sum_{k_m=1}^{N_m} \Phi_i \begin{bmatrix} \mathbf{f}_{m,i,k_m}^b \\ \mathbf{m}_{m,i,k_m}^b \end{bmatrix}, \quad i = 1, \dots, N \quad (10)$$

where  $\Phi \in \mathbb{R}^N$  and  $\Phi_i$  is equal to 1 if module  $i$  is moored, and 0 if it is not.

### 3. RESULTS AND DISCUSSION

#### 3.1 Natural Periods

Table 3.  $N = 2$ , Calculated undamped natural periods (mass + added mass  $1.5e4$ )

	$K_c = S1$	$K_c = S2$	$K_c = S3$	$K_c = S4$
$T_{n1}$ [s]	11.8	11.8	11.8	11.8
$T_{n2}$ [s]	2.2	2.3	3.0	4.5

Table 4.  $N = 3$ , Calculated undamped natural periods in surge (mass + added mass  $1.5e4$ )

	$K_c = S1$	$K_c = S2$	$K_c = S3$	$K_c = S4$
$T_{n1}$ [s]	14.5	14.5	14.5	14.7
$T_{n2}$ [s]	3.1	3.2	4.2	5.9
$T_{n3}$ [s]	1.8	1.9	2.6	3.8

Table 3 presents the estimated full scale undamped natural periods in surge of a 1 DOF system with 2 modules and with different stiffness of the connector spring.  $T_{n1}$  corresponds to the mode shape where both modules are moving in the same direction, while  $T_{n2}$  corresponds to motion in opposite directions.

Natural periods for a system with 3 modules are presented in Table 4.  $T_{n1}$  corresponds to all modules moving in the same direction.  $T_{n2}$  corresponds to the middle module at a standstill, while the two end modules move in opposite directions causing large relative motion in both connection pairs.  $T_{n3}$  is the natural period of the mode shape where the first and third modules move in the same direction, but opposite the middle. Based on the nature of these mode

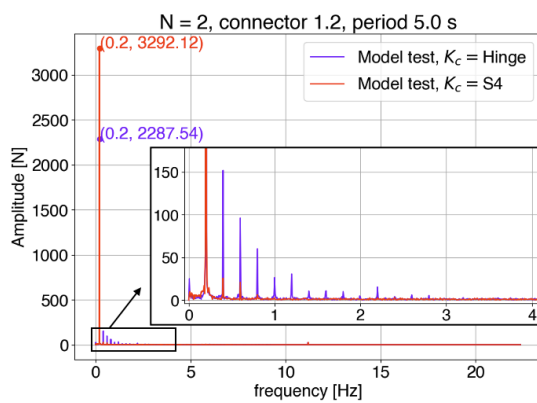
shapes, the relative motion, and thus the connector force, is expected to be largest at  $T_{n2}$  and  $T_{n3}$ .

Finally, the mode shapes for the five-module configuration are similar to the three-module case. With more modules, there are more mode shapes and possible resonance periods, and it becomes less evident which wave periods will lead to the largest mean connection load amplitudes.

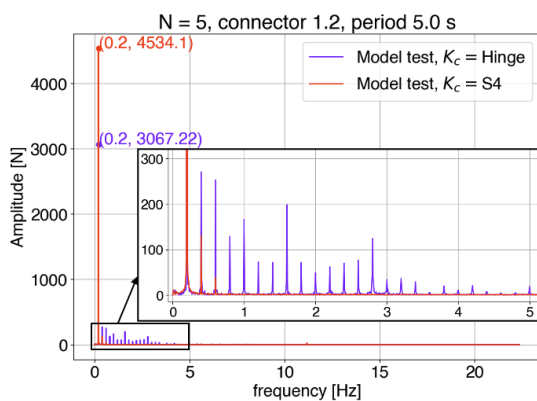
The first natural periods are corresponding to global surge motion of the island and are not expected to lead to significant relative motion. These were calculated to be outside the interval of wave periods that have been investigated, and  $T_{n1}$  will thus not be discussed further in this work.

Heave natural period for a single module due to hydrostatic restoring force is estimated to be 2.7 s.

### 3.2 Stiffness Dependence in Model Test



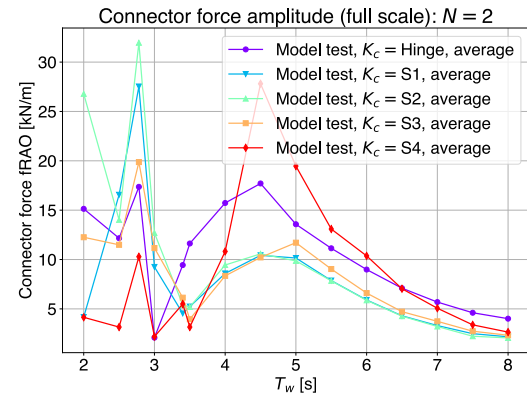
(a) N=2



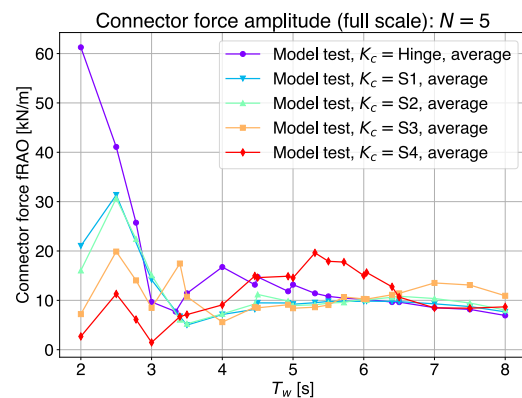
(b) N=5

Fig. 5. Measured higher harmonic forces in one connector between module 1 and 2. Hinged connection compared to the softest spring connection, stiffness  $K_c = S4$ .

Connector loads are plotted in frequency domain in Figs. 5(a) (N=2) and 5(b) (N=5) for wave period  $T_w = 5.0$  s. To highlight the effect of the connector stiffness, these figures present two extreme cases: the hinged connector and the softest spring connector,  $S4$ . Although amplitudes at the wave frequency are higher, the softer spring appears to transfer little to no forces at higher frequencies.



(a) N=2



(b) N=5

Fig. 6. Average connector force all connectors.

Figures 6(a) and 6(b) present connector forces in the local  $x$ -direction for N=2 and N=5 configurations respectively. The hinged connector generally yields higher connector loads than the tests where modules are connected by springs. It appears beneficial with a softer connection at lower wave periods, while for longer periods the stiffer springs are preferable to avoid the resonance peak related to  $T_{n2}$ .

As mentioned in Sect. 2.2, the first and last module of each array were moored in all tests. The connector stiffness is therefore expected to have less impact on module response of a two-module array, where both modules are moored, than for a five-module array with three middle modules held in place only by connectors. This can explain the similar surge motion at  $K_c = S1, S2$  and  $S3$  for two-module array, seen in Fig. 7(a). The softest spring,  $S4$ , yields a resonance peak at  $T_w = 4.5$  s, coinciding with the calculated second natural period of the system,  $T_{n2}$ . Similar resonance peaks for the stiffer connection stiffnesses are visible, though less prominent.

Collisions between modules, as well as larger sway motion, were observed for some tests. The occurrence was most severe for the five-module configuration with the two softest springs  $S3$  and  $S4$ . Ideally, the symmetrical model should show little to no sway motion when subjected to regular waves in the longitudinal direction. However, due to model asymmetries and low bending stiffness of the springs, collisions lead to transversal motion of some



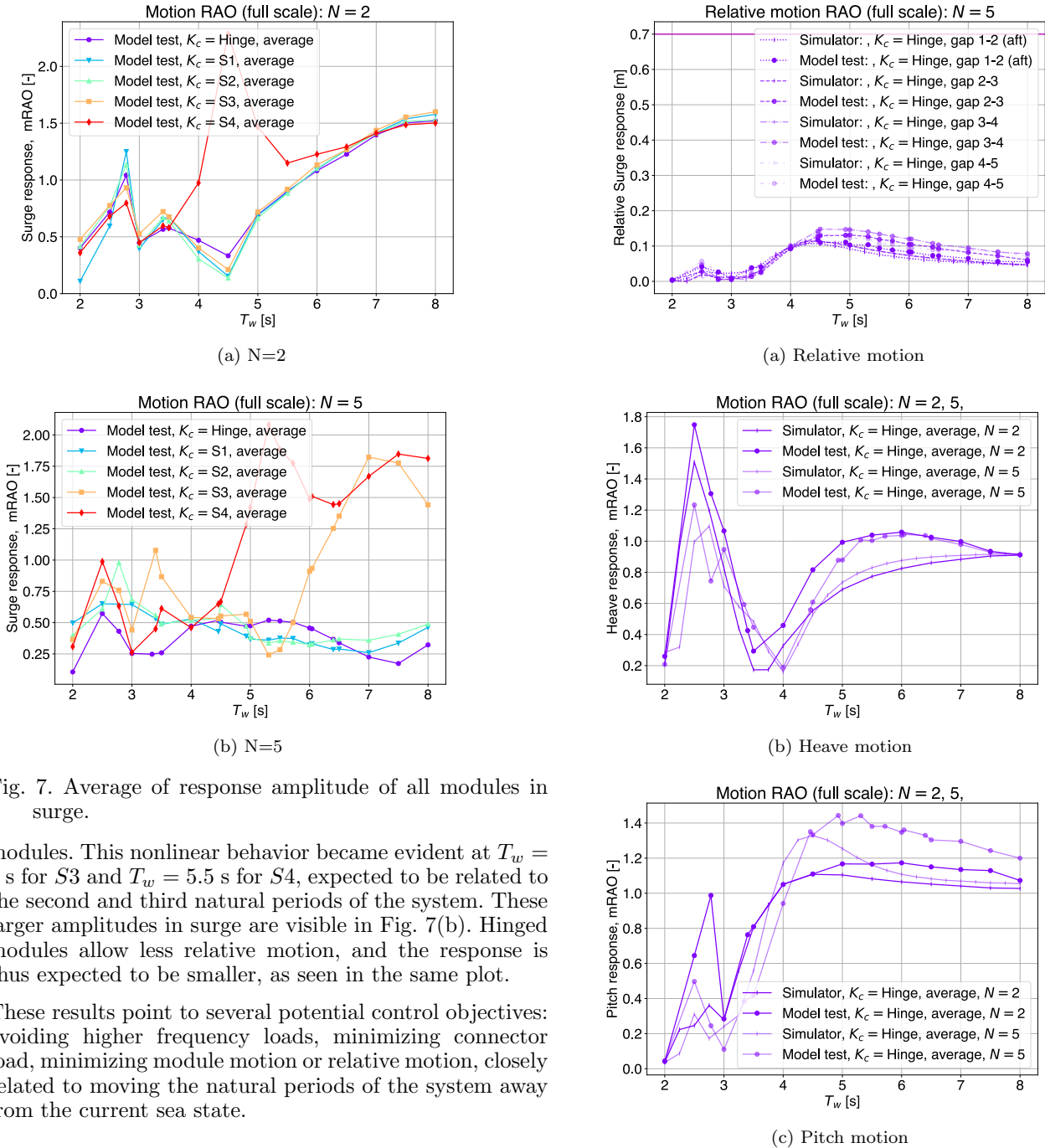


Fig. 7. Average of response amplitude of all modules in surge.

modules. This nonlinear behavior became evident at  $T_w = 6$  s for S3 and  $T_w = 5.5$  s for S4, expected to be related to the second and third natural periods of the system. These larger amplitudes in surge are visible in Fig. 7(b). Hinged modules allow less relative motion, and the response is thus expected to be smaller, as seen in the same plot.

These results point to several potential control objectives: avoiding higher frequency loads, minimizing connector load, minimizing module motion or relative motion, closely related to moving the natural periods of the system away from the current sea state.

### 3.3 Comparison with Numerical Model

The hinged configuration is simulated by setting the connector stiffness high,  $K_c = 4e5$ . As seen in Fig. 8(a), the simulator is able to give a satisfactory estimation of the relative motion between hinged modules, a motion mainly occurring due to pitch motion. Further, as the connection springs get softer, there is a larger discrepancy between ex-perimental and numerical results. Fig. 9 shows the relative motion between modules for different connector stiffnesses. The peaks can be seen in relation with Tab. 4 as resonance peaks. The plots include a solid horizontal line indicating the collision limit.

The heave and pitch response for different connector stiffnesses are shown in Figs. 11(a) and 11(b) respectively. An

FIG. 8. Comparison of results with hinged modules.

average response is calculated from the motion amplitude of all modules. The heave response is mostly governed by potential forces, and not affected as much by the connections to other modules. The resonance peak corresponds to the calculated eigenperiod in heave for one module. The heave and pitch plots show a closer correspondence with the model test than the estimated surge motion and connector force shown in Fig. 10.

### 3.4 Error Sources

The main sources of discrepancy between the model test and the simulations are considered to be:

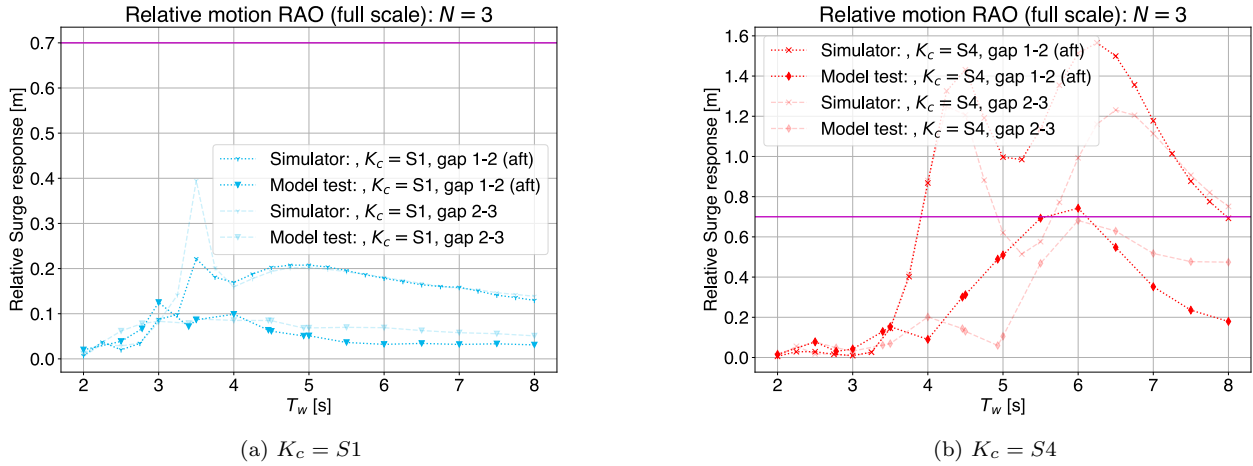


Fig. 9. Relative motion in surge between modules,  $N=3$ .

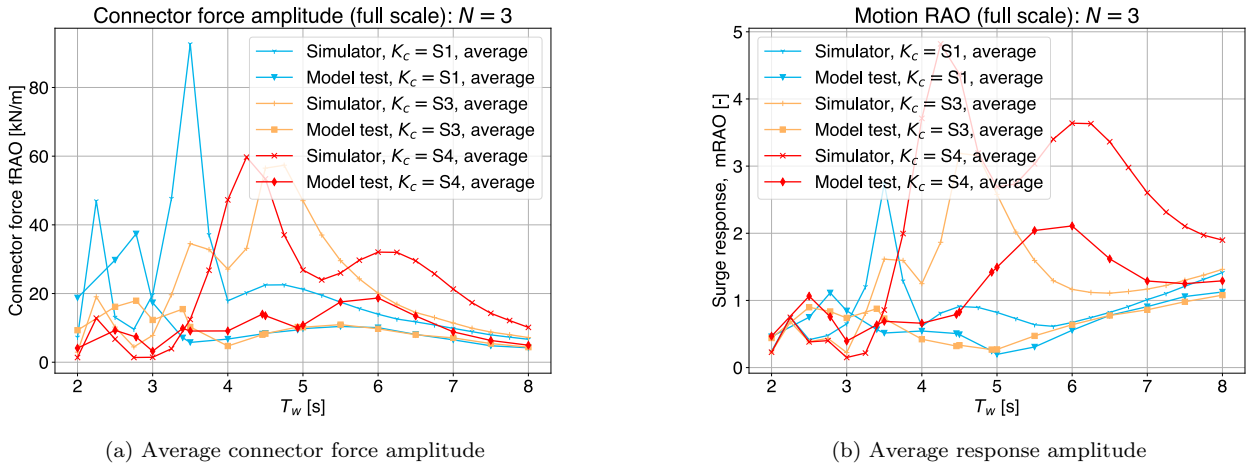


Fig. 10. Comparison between experimental and numerical results,  $N = 3$ , at different connector stiffness  $K_c$ .

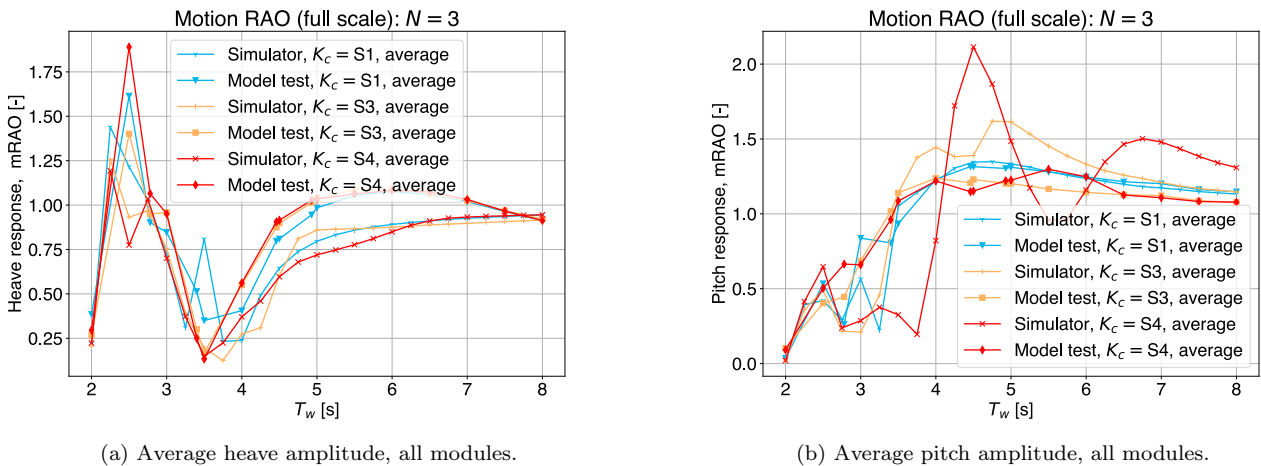


Fig. 11. Comparison between experimental and numerical results,  $N=3$ , heave and pitch response at different connector stiffness  $K_c$ .

- *Mooring line friction:* The mooring line in the model test included pulleys resulting in unmeasured friction and considerable added damping to the system. A corresponding force has not been included in the numerical model.
- *Nonlinear connector behavior:* The design of the connection springs only allowed for a known stiffness in the axial direction. The transversal and vertical, as well as the rotational stiffnesses are unknown. This leads to nonlinear behavior, especially for larger

waves and softer springs, that is not captured by the numerical model. The simulator does not seem to recreate the connector force measured in the model test well.

- *Collisions*: Collisions between modules and sway motion forced by the springs were observed for some runs. When collisions occur, forces in connections may be transferred to the module structure or to transversal motion without being measured, meaning that in these cases the load cells do not capture all interaction between neighboring modules.
- *Module overlap*: Related to collisions. Modules are allowed to overlap in the numerical simulations, meaning that response from the simulations can become large at resonance since the relative motion between modules is not limited by collisions.
- *Wake effects*: The design with several columns in close proximity of each other is expected to experience significant wake interaction, which is not considered in the simulator. These interactions will affect the hydrodynamic loads on the module columns, especially higher harmonic loads.

#### 4. CONCLUSION

This paper has presented experimental and numerical analyses of a multi-modular floating structure, to investigate three main objectives: 1) Analyze the behavior of a multi-modular structure in changing wave conditions, 2) Investigate the effect of changing the stiffness of connectors between neighboring modules and 3) Validate a numerical simulator by comparable experimental results. First, choosing a spring connection with a proper stiffness instead of a hinged connection between modules appears to be beneficial to reduce connection loads between neighboring modules. When the array of modules is subjected to longer waves, a stiffer connection leads to a lower connection load, while at shorter waves a softer spring seems preferable. A softer spring shows fewer higher harmonic frequencies in the connection load than a hinge type connection. Choosing a softer spring in the connections could thus be beneficial if higher frequency loads are undesirable. The optimal stiffness for different sea states must be found as a balance between minimizing loads and staying within acceptable limits for module motion.

In general, the simulator proposed in this work overestimates load and response compared to model tests. The simulator does not capture the full dynamics shown by the model test and needs further investigation, especially considering the error sources mentioned in Sect. 3.4.

There is a significant amount of uncertainty related to the model tests, particularly related to connector design and mooring configuration. An improved connector design in future tests is necessary, and a mooring system without pulleys is preferable. Other possible topics for future experimental testing include investigating a 2D matrix of modules, changing wave angle, and including irregular sea states.

#### ACKNOWLEDGEMENTS

We would like to thank the engineers from NTNU IMT for all assistance during model tests.

#### REFERENCES

- Ding, R., Yan, D.L., Zhang, H.C., Lu, Y., Shi, Q.J., Tian, C., Zhang, J.L., Ni, X.U., Xu, D.L., and Wu, Y.S. (2021). An application of network modeling method to scientific research and demonstration platform-connector load analysis. *Journal of Hydrodynamics* 2021 33:1, 33, 33–42. doi:10.1007/S42241-021-0011-2.
- DNV (2022). The future of floating solar: Drivers and barriers to growth.
- DNV-GL (2017). *SESAM USER MANUAL - WADAM - Wave Analysis by Diffraction and Morison theory*. DNV-GL, 9.3 edition.
- Fagerbakke, V.N. (2023). *Enabling Intelligent Multi-Modular Concept for Solar Energy Harvest by Local Control of Connectors*. Master's thesis, NTNU.
- Fossen, T. (2021). *Handbook of Marine Craft Hydrodynamics and Motion Control*. Wiley.
- IPCC (2023). Sections. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. 35–115. doi: 10.59327/IPCC/AR6-9789291691647.
- Jiang, D., Tan, K.H., Wang, C.M., and Dai, J. (2021). Research and development in connector systems for very large floating structures. *Ocean Engineering*, 232, 109150. doi:10.1016/J.OCEANENG.2021.109150.
- Kumar, M., Niyaz, H.M., and Gupta, R. (2021). Challenges and opportunities towards the development of floating photovoltaic systems. *Solar Energy Materials and Solar Cells*, 233, 111408. doi: 10.1016/J.SOLMAT.2021.111408.
- Lamas-Pardo, M., Iglesias, G., and Carral, L. (2015). A review of very large floating structures (vlfs) for coastal and offshore uses. *Ocean Engineering*, 109, 677–690. doi: 10.1016/J.OCEANENG.2015.09.012.
- Onsrud, M. (2019). *An Experimental Study on the Wave-Induced Vertical Response of an Articulated Multi-Module Floating Solar Island*. Master's thesis, NTNU.
- Shi, Q.J., Zhang, H.C., Xu, D.L., Qi, E.R., Tian, C., Ding, J., Wu, Y.S., Lu, Y., and Li, Z.W. (2018). Experimental validation of network modeling method on a three-modular floating platform model. *Coastal Engineering*, 137, 92–102. doi:10.1016/j.coastaleng.2018.04.001.
- Xia, S.Y., Xu, D.L., Zhang, H.C., Qi, E.R., Hu, J.J., and Wu, Y.S. (2016). On retaining a multi-module floating structure in an amplitude death state. *Ocean Engineering*, 121, 134–142. doi:10.1016/j.oceaneng.2016.05.024.
- Xu, D.L., Zhang, H.C., Qi, E.R., Hu, J.J., and Wu, Y.S. (2014). On study of nonlinear network dynamics of flexibly connected multi-module very large floating structures. doi:10.1061/9780784413609.181.
- Zhang, H.C., Xu, D.L., Lu, C., Qi, E.R., Tian, C., and Wu, Y.S. (2017). Connection effect on amplitude death stability of multi-module floating airport. *Ocean Engineering*, 129, 46–56. doi:10.1016/j.oceaneng.2016.11.011.
- Zhang, H.C., Xu, D.L., Xia, S.Y., Lu, C., Qi, E.R., Tian, C., and Wu, Y.S. (2015). Nonlinear network modeling of multi-module floating structures with arbitrary flexible connections. *Journal of Fluids and Structures*, 59, 270–284. doi:10.1016/j.jfluidstructs.2015.09.012.

# Life Cycle Assessment of Floating Offshore Wind Farms: The Case of Hywind Tampen in Norway

Omid Lotfizadeh\*, Zahir Barahmand\*\*, Hadi Amlashi\*\*\*

\* *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway (Tel: +47 96728466; e-mail: [omid.lotfizade@gmail.com](mailto:omid.lotfizade@gmail.com)).*

\*\* *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway (e-mail: [zahir.barahmand@usn.no](mailto:zahir.barahmand@usn.no)).*

\*\*\* *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway (e-mail: [hadi.amlashi@usn.no](mailto:hadi.amlashi@usn.no)).*

---

**Abstract:** To address climate change and energy security issues from fossil fuels, wind power is a promising renewable energy source, projected to grow significantly by 2050. Offshore wind energy, especially floating offshore wind farms shows great potential due to higher and more consistent wind speeds at sea. However, these turbines have negative environmental burdens throughout their life cycle. This The present study focuses on a comprehensive cradle-to-grave life cycle assessment of the Hywind Tampen floating offshore wind farm in Norway. The assessment covers all stages from manufacturing, transportation, installation, operation, and maintenance to decommissioning, utilizing openLCA® software and ecoinvent 3.9 database with the ReCiPe 2016 impact assessment method. Key findings indicate that manufacturing is the primary contributor to total emissions, followed by operation and maintenance. The study emphasizes the necessity of developing more sustainable manufacturing methods, designing turbines that are more efficient and versatile, and better maintenance forecasting and planning in order to minimize the environmental impact of these turbines.

**Keywords** life cycle assessment, offshore wind, floating, openLCA®, wind energy, renewable energy, climate change

---

## 1. INTRODUCTION

Rapid urbanization and population growth are driving a 50% increase in global energy demand in the coming years (Skår, 2022). This surge is primarily met by fossil fuels, leading to resource depletion, global warming, and other environmental impacts. Among renewable energy resources wind power stands out as one of the most accessible and environmentally friendly options (Narayanan, 2023). Additionally, offshore wind energy is emerging as one of the most promising options for the coming years and decades, thanks to the higher and more consistent wind speeds found in open seas (Kaltenborn et al., 2023).

Despite producing clean electricity, offshore wind turbines have environmental impacts across their life cycle, including manufacturing, installation, and decommissioning. The environmental impact and energy performance of offshore wind technology are commonly assessed using Life Cycle Assessment (LCA) (Bhandari et al., 2020) which is the most commonly employed method to simulate and assess the environmental impacts of products and processes (Barahmand and Eikeland, 2022).

Offshore wind turbines are categorized into two types based on their foundations (Bhattacharya, 2019).

**Grounded (bottom-fixed)** The wind turbine is securely bolted or driven into the seabed, like a giant anchor.

**Floating:** The wind turbine sits on a special platform that floats on the water, held in place by mooring lines.

Floating wind is an emerging technology, thus there is a limited availability of studies on the subject. After conducting a literature review, to date, the authors identified only 9 LCA studies on floating offshore wind. Although some studies like (Alsubal et al., 2021) were performed for life cycle cost assessment (LCCA). Among these studies only (Bang et al., 2019; Brussa et al., 2023; Garcia-Teruel et al., 2022; Struthers et al., 2023; Yildiz et al., 2021), were focused only on the floating platforms while the rest of them were more interested in bottom-fixed platforms.

Yildiz et al. (2021) conducted LCA on only one wind turbine. On the other hand, Bhandari (2020) conducted LCA on both farms and wind turbines, the rest of the studies was conducted on wind farms. The rest of the previous LCA studies were conducted on all life cycle stages of the wind farm including manufacturing, transportation, operation and maintenance (O&M) and decommissioning. On the other hand, Skår (2022) considered only the decommissioning stage. In all founded research, the manufacturing stage is regarded as the most important stage due to its highest contribution to the total emissions. Only a few studies conducted LCA on real-world wind farm case studies.

## 2. BACKGROUND

In recent years, the offshore wind industry has seen notable expansion, with offshore wind capacity growing by approximately 30% annually since 2010. Moreover, the size of the largest wind turbines has risen from 3 MW in 2010 to 8 MW in 2016, with projected ratings reaching up to 15–20 MW by 2030 (Garcia-Teruel et al., 2022).

While most deployed technologies utilize bottom-fixed structures such as monopiles or jackets, the utilization of floating turbines is rising as the industry explores locations with deeper sea depths. There's ongoing debate and research to determine the economic viability of floating platforms compared to bottom-fixed turbines, typically within the transition depth range of 50 to 100 meters. This threshold may be affected by factors such as the type of floater and the site conditions. However, for depths more than 100 meters, floating concepts are widely regarded as the most cost-effective approach (Karimirad, 2014). The floating wind turbine foundations can be categorized into three main types, as illustrated in Fig. 1 adopted from (Bhattacharya, 2019):

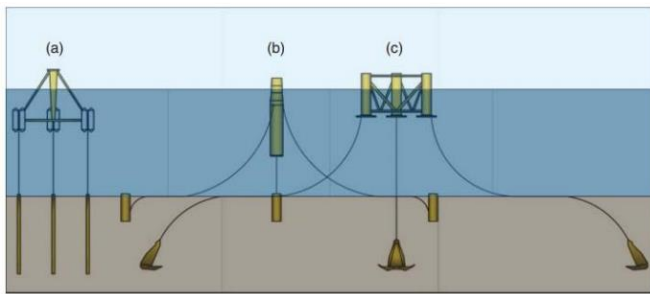


Fig. 1. The main Types of floating wind turbine adopted from (Bhattacharya, 2019).

1-TLP (Tension Leg Platform) with mooring stabilization: This system utilizes tensioned mooring for stability and is firmly anchored to the seabed to maintain buoyancy and stability.

2-Spar buoy with ballast stabilization, optionally equipped with motion control stabilizers: this system features a deep cylindrical base for ballast, with the lower section significantly heavier than the upper section, ensuring the center of buoyancy is higher than the center of gravity. While cost-effective initially, these structures require greater water depths and are not suitable for shallow environments.

3-Semisubmersible buoyancy stabilization: This design combines ballasting and tensioning principles, requiring substantial steel components.

## 3. METHODOLOGY

As per ISO 14040 and 14044 standards, the Life Cycle Assessment (LCA) framework comprises four stages (Lotfzadeh, 2024):

- Defining goals and scope
- Conducting a Life Cycle Inventory (LCI) analysis
- Performing a Life Cycle Impact Assessment (LCIA)

- Interpreting the results

### 3.1 Goal and scope

The initial step in an LCA, defining goals and scope, is widely regarded as crucial as it sets the research context, defines modelling requirements, and outlines project planning (Hesan, 2023).

The goals of this study were to: 1) Assessing the environmental impact of all life cycle stages of the Hywind Tampen wind farm. 2) Identifying the key elements affecting the environmental impact of offshore wind projects. 3) Learning about potential opportunities for environmental optimization throughout the life cycle and 4) Identifying relevant areas for further studies.

A cradle-to-grave method is chosen, and the boundaries of the system are shown Fig. 2. The defined functional unit (FU) in this study is 1 MWh of electricity generated by the wind farm during its life cycle and then delivered to the grid. Recycling was not included in the current study's end of life (EOL) stage due to uncertainties and data availability issues. As shown in Fig. 2, recycling falls outside the system boundaries.

### 3.2 Life cycle inventory analysis (LCI)

In this section the data collection and calculations will be briefly discussed. The Hywind Tampen is chosen as base case scenario because this wind farm uses the most recent technologies and largest turbine sizes in floating offshore wind (Lotfzadeh, 2024).

Table 1. Specifications of the base case (Lotfzadeh, 2024).

Wind Farm Name	Hywind Tampen
Distance to port	140 km
Power of each turbine	8 MW
Number of turbines	11
Wind Farm Capacity Factor	54 %
Generator type	Direct drive
Lifetime	20 years
Foundations	Concrete SPAR-type
Tower Length	92 m
Rotor Diameter	167 m
Total Height	175 m
Distance between the turbines	1.5 km
Water depth	200 m

Inventory analysis involves collecting data and performing calculations to identify the inputs and outputs of a product system. Inputs consist of energy, raw materials, and other products, while outputs encompass waste, water and air pollution, and various byproducts (Garcia-Teruel et al., 2022). These inputs and outputs were utilized as flows in each unit process and modelled using the openLCA® software. The inventory data were gathered from the following sources: 1) literature 2) reference wind turbines 3) environmental product declarations (EPDs).

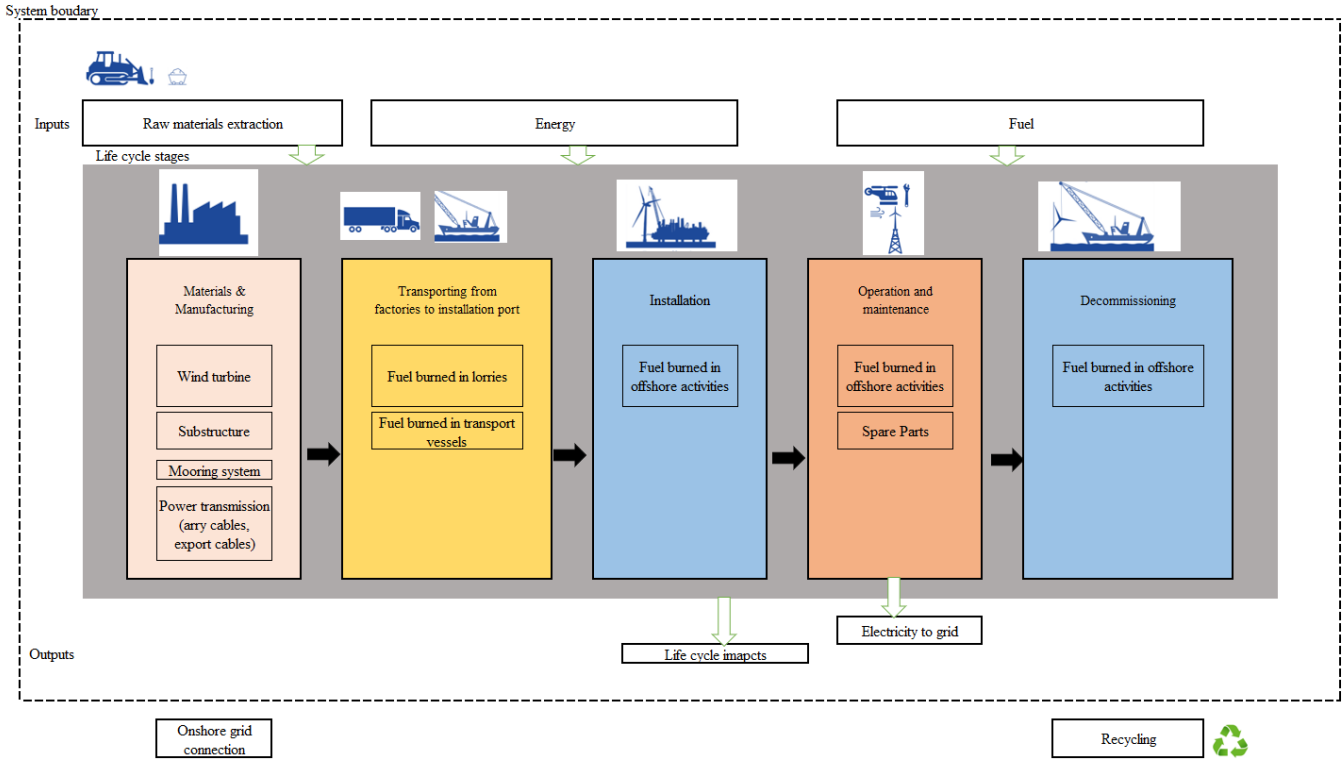


Fig. 2. System boundaries adopted from Lotfizadeh,2024)

It is important to note that access to specific details about wind turbines and wind farms is restricted due to commercial sensitivity. This lack of full transparency requires making certain assumptions when conducting LCA of offshore wind farms (Lotfizadeh, 2024). In the following a brief description of inventories and calculations will be given. Detailed inventories and calculations are available in open access (“Supplementary materials-life cycle assessment of offshore wind Farms, Lotfizadeh,” 2024).

3.2.1 Materials and manufacturing

Simulating the raw material supply is done by using market datasets from the ecoinvent database, including material procurement and transit to Europe (Brussa et al., 2023). Previous research either focused on smaller wind turbines or lacked details about the materials used. Some studies like (Bang et al., 2019) and (Garcia-Teruel et al., 2022) estimated missing information by using regression. This study assumes a linear connection between the size of a turbine and material weight distribution. To determine the materials and weight for the 8 MW turbines, we used interpolation method based on a 6 MW turbine and a 15 MW reference turbine (Gaertner et al., 2020).

3.2.1.1 Tower and Nacelle

The main component of the 8 MW tower is low-alloy steel (Brussa et al., 2023). Siemens Gamesa EPD specifies the tower's length 92 meters, but information about its diameter and wall thickness is missing. The estimation of the weight was done using a linear interpolation method. The paint on the tower is negligible compared to the weight of other materials and was therefore excluded from the calculations.

For welding the processes “welding, arc, steel” in Ecoinvent was applied. In some other studies, the welding length was regarded as a continuous weld along the tower height. However, this study assumes that the tower is composed of welded segments, each with a height of 2 meters, and takes into account the peripheral length of these welded segments. Figure 3 illustrates the welding process, and Equation (1) demonstrates the calculation method.

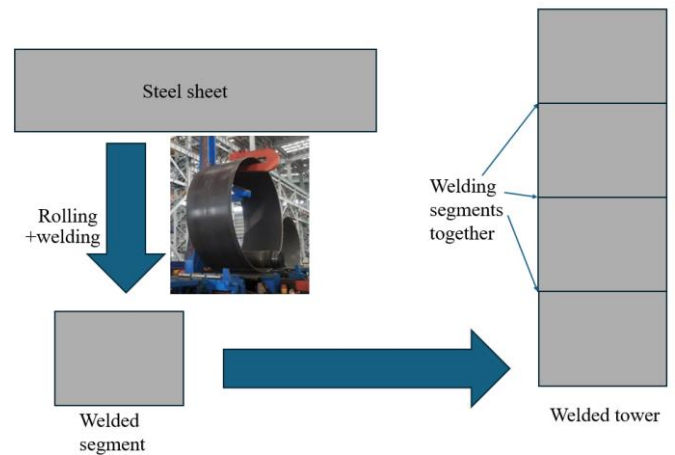


Fig. 3. Tower manufacturing process (Lotfizadeh, 2024).

$$L_W = L_T + N_s \times P \tag{1}$$

Where  $L_W$  represents the total welding length of the tower,  $L_T$  denotes the length of the wind turbine tower,  $N_s$  is the number of segments in the tower, and  $P$  is the perimeter of each segment. For an 8 MW wind turbine with a diameter of 10 m,



the total welding length  $L_W$  is calculated as follows. The tower length is 92 m, made up of 46 segments, each 2 m in height and 10 m in diameter.

$$L_W = 92 + 46 \times \pi \times 10 = 1537 \text{ m}$$

### 3.2.1.2 Substructure

The material and weight of the 8 MW turbine substructure were taken from the environmental product declaration (EPD) of Siemens Gamesa 8 MW wind turbine. The substructure comprises two main components: the spar structure and ballast. The welding length of the spar structure was also calculated using Equation (1).

### 3.2.1.3 Mooring System

The mooring system data for the Hywind Tampen project was unavailable, however the weight and material data for the Hywind Scotland project were obtained from the project's manufacturing factsheets ("Manufacturing Factsheets," 2024). As a result, it was assumed that the mooring chains and anchors for the two projects were identical.

### 3.2.1.4 Power Transmission

The power transmission category includes inter-array cables, export cables, and substations. As Hywind Tampen wind farm distributes electricity to the nearby oil platforms, no substation was used in this study's base case scenario. Hywind Tampen inter-array and export cables were made by JDR company, which also manufactured cables for Hywind Scotland project, hence this study relied on the manufacturing factsheets of the Hywind Scotland project to get data on cable specifications. The Hywind Tampen Inter-array cables are 2.5 kilometres long, 66kV dynamic array cables (Lotfizadeh, 2024). The length of the export cable for the Hywind Tampen wind farm was determined to be 45.4 kilometers based on the relative distances of the five nearby platforms.

### 3.2.2 Transportation

Two modes of transportation are covered within the study's boundaries. To begin, as previously stated, this study models the raw material supply chain by using market datasets from the Ecoinvent database, which includes both material acquisition and transit to Europe (Brussa et al., 2023). Second, transportation from the factory to the installation port. These transports are carried out by truck or vessel. It was assumed that some parts of the turbine components were transported by truck within Denmark to the Siemens Gamesa factory and after assembling there were transported by ship to Norway to be installed at the Hywind Tampen site.

### 3.2.2 Installation

Most prior research used the "transport, freight, sea, ferry - GLO" process in ecoinvent to model the emissions from vessel installation activities; however, this study chose ecoinvent's "diesel, burned in diesel-electric generating set" process.

The energy demand of all vessels in installation activities including, installing foundations, turbine tower, rotor, nacelle, cables and mooring system was calculated and set as "diesel, burned in diesel-electric generating set" process in the openLCA® software.

### 3.2.3 Operation and Maintenance (O&M)

This stage quantifies emissions from operations and maintenance (O&M) activities, including unexpected repairs due to failures, routine preventative maintenance, and spare parts. It is important to note that due to the lack of data on remotely operated vehicles (ROVs) in the Ecoinvent database, their activities and emissions were excluded from this study.

#### 3.2.3.1 Unexpected Maintenance

For unexpected maintenance the failure rates are categorized into major replacement, major repair, and minor repair. To calculate the overall number of turbine failures over their lifetime, the annual failure rates (Fig. 4) were multiplied by the number of turbines of the farm ( $\times 11 \times 20$ ). The time needed to fix each component within each operation and maintenance category were obtained from (Centeno-Telleria et al., 2024). With available energy consumption data for the vessels, the energy consumed for transport to the site and O&M operation was calculated in MWh using repair hours for each component. These figures were employed in ecoinvent's "diesel, burned in diesel-electric generating set" process in openLCA® software, following the same approach as the installation phase.

#### 3.2.3.2 Regular Maintenance

The same method applied to regular maintenance, assuming once-a-year visit of the wind farm for preventative maintenance (PM) of the wind turbines' components.

#### 3.2.3.3 Spare parts

There is limited publicly available data on wind turbine component replacement rates. This study adopted the same exchange rate as (Arvesen et al., 2013). The rate of annual replacement for large wind turbine components is assumed to be 0.075 per wind turbine, and for generators and blades 0.333 per wind turbine.

Table 2. Spare parts replacement rates (Lotfizadeh, 2024).

Spare Parts	Annual replacement Per Wind Turbine	Annual replacement Per Wind Farm	Lifetime replacement Per Wind Farm
Replacement large parts <sup>1</sup>	0.075	0.825	16.5
Blades	0.333	3.667	73.3
Generators	0.333	3.667	73.3

<sup>1</sup> Turret / Nose, Bedplate, Flange, Shaft Bearings, Yaw System

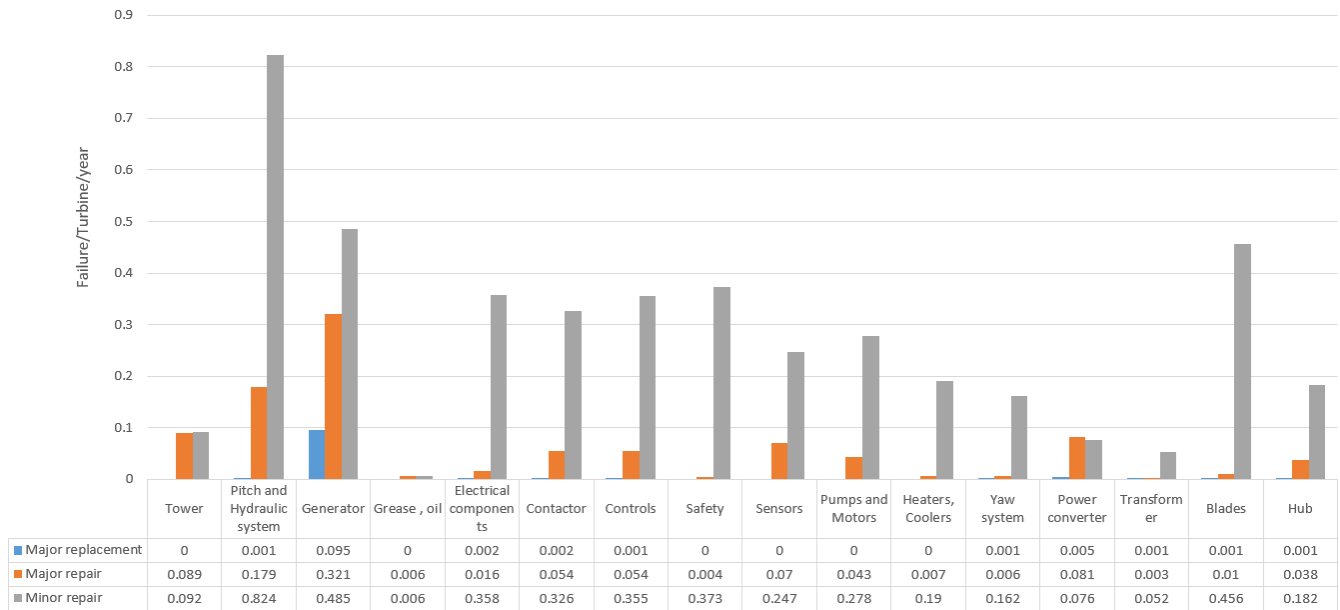


Fig. 4. Annual failure rates (Lotfizadeh, 2024).

3.2.4 Decommissioning

In this study it is assumed that the emissions from decommissioning stage are the reverse and equivalent to the installation stage.

3.2.4 Electricity delivered to the grid by the wind farm

The lifetime electricity production of the wind farm was calculated using the Equation 2.

$$E_{F,L,R} = C \times C_F \times L \times N_T - E_{Loss} \quad (2)$$

where, each term is described in Table 3.

Table 3. Different terms of electricity calculation equation

Term	Description	Unit
$E_{F,L,R}$	Real power production of the farm after losses	MWh
$C$	Capacity of each turbine	MWh
$C_F$	Capacity factor	-
$L$	Lifetime of the wind farm	hour
$N_T$	Number of turbines in the farm	-
$E_{Loss}$	Electrical loss due to downtime	MWh

The loss due down time was calculated to be :  $E_{Loss} = 53,508$

Then

$$E_{F,L,R} = 8 \times 0.54 \times 20 \times 365 \times 24 \times 11 - 53508 = 8,256,878 \text{ MWh}$$

3.3 Life Cycle Impact Assessment (LCIA)

The openLCA® version 2.1 and ecoinvent 3.9 databases were utilized to perform LCIA. The ReCiPe 2016 v1.03 midpoint (H) method was selected to ensure that the results are comparable with previous studies.

4. RESULTS AND DISCUSSION

The impacts of the base case scenario were measured using the ReCiPe Midpoint (H) 2016 approach, which included 18 impact categories. The results were normalized by dividing by  $E_{F,L,R}$  (the lifetime electrical power delivery of the farm after all losses in MWh). The results of the 18 impact categories of the base case scenario are shown in Table 5.

Some heatmaps were created using Microsoft Excel® software to help visualizing the data. These heatmaps employ three colors to depict varying levels of influence. Green colors indicate lesser impact values, yellow indicates the 50th percentile (the midpoint), and red intensifies when values exceed the middle and approach maximum impact. Fig. 5 illustrates the rule for creating heatmaps in Microsoft Excel®.

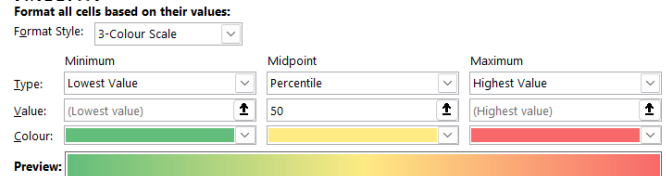


Fig. 5. The rule for creating heatmaps with Microsoft Excel®.



Table 4.A heatmap of the contribution of each life cycle stage to the total GWP for the base case scenario.

Stage	Contribution (%)	GWP (kg CO <sub>2</sub> -Eq/ MWh)
Wind Turbine Manufacturing	26.79%	9.85
Substructure Manufacturing	26.73%	9.83
Mooring system Manufacturing	2.82%	1.04
Power Transmission Manufacturing	1.32%	0.49
Transportation	0.07%	0.03
Installation	5.91%	2.17
O & M vessel	16.27%	5.98
O & M spare parts	14.18%	5.22
Decommissioning	5.91%	2.17
Total	100.00%	36.78

Transportation has a very low share, with 0.03 Kg CO<sub>2</sub>-Eq per MWh. Manufacturing contributes the most to overall GWP, and the floating farm's substructure had significant emissions due to the use of concrete to manufacture the spar substructure. Figure 6 illustrates the contribution of the main five life cycle stages to the total GWP. The second contributor to the total GWP emissions was operation and maintenance stage.

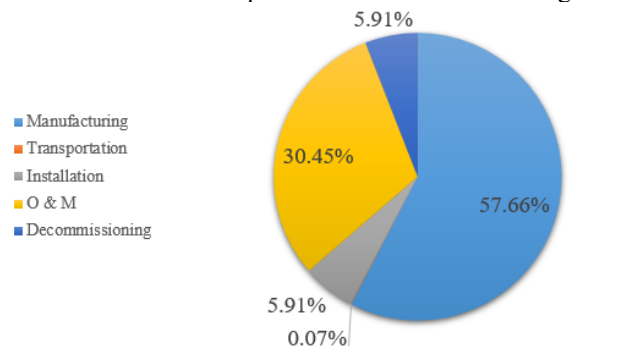


Fig. 6. Contribution of the main five life cycle stages to the total GWP.

4.1 Sensitivity Analysis

This section examines how variations in critical characteristics during the life cycle stages of the base case scenario impacts the overall results of the life cycle assessment.

As can be seen in Fig. 7 by decreasing the capacity factor (CF), it was expected that the global warming potential (GWP) and other environmental impacts would increase, which the results confirmed. Conversely, increasing the CF was expected to reduce GWP and other environmental impacts, and extending the farm's operational lifespan was anticipated to further decrease these impacts. Both hypotheses were validated by the results.

As the distance to the shore increases, the fuel consumption for vessel activities rises, leading to an increase in the GWP

amount. However, the increase in GWP due to changes in the capacity factor (CF) and lifetime was significantly greater than the increase resulting from changes in distance to shore.

The strategy of towing to the shore was assumed to be used only for major replacements. While major repairs and minor repairs were conducted at the wind farm location. Results indicated that GWP increased by 11.5% when this strategy was implemented. Therefore, the optimal O&M approach for major replacements is to perform operations at the wind farm site rather than towing the wind turbines back to shore.

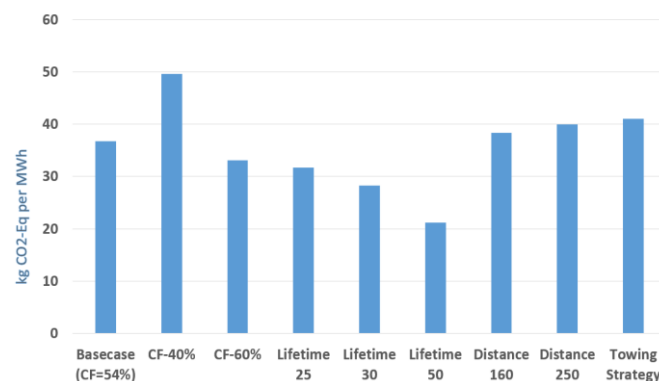


Fig. 7. An overview of GWP value in all scenarios.

5. CONCLUSIONS AND FURTHER RESEARCH

This paper provides a detailed assessment of the environmental implications associated with the Hywind Tampen floating offshore wind farm. The LCA findings indicated that, for the base case scenario, the GWP was calculated to be 36.78 kg CO<sub>2</sub>-Eq per MWh.

It was also discovered that the manufacturing stage was accounted for nearly 57% of total GWP emissions, followed closely by the operation and maintenance (O&M) stage. Wind turbine failures accounted for approximately 90% of emissions throughout the operation and maintenance stage. To address these challenged wind turbine component manufacturers ought to develop and implement more sustainable production practices. For example, design strategies that maximize generation capacity per unit of material used could significantly reduce emissions associated with the manufacturing stage. Furthermore, improving wind turbine reliability can lower the environmental impact of the operation and maintenance stage.

Additionally, The sensitivity analysis explored how various parameters impact the results. Notably, the capacity factor and lifetime of the wind farm significantly influence overall environmental impacts.

For further studies, it is recommended that:

- Using eco-friendly vessels during installation, operation and maintenance and decommissioning
- The O&M stage was shown to be the second-largest contributor to overall emissions in the evaluated wind farm. This emphasizes the significance of performing a sensitivity analysis for failure rates.

- Emissions from decommissioning were assumed to be equal to those from the installation stage. Further investigation of the decommissioning stage, as well as a sensitivity analysis using various decommissioning strategies, is recommended.
- The study did not include recycling in the end of life stage due to uncertainties and data availability issues. Further investigations on this stage, such as performing a cradle-to-

cradle LCA, could provide useful insights into the materials used to manufacture offshore wind turbines.

Table 5. The results of 18 impact categories of the base case scenario

Impact category	Reference unit/MWh	Value
acidification: terrestrial - terrestrial acidification potential (TAP)	kg SO <sub>2</sub> -Eq	0.15
climate change - global warming potential (GWP100)	kg CO <sub>2</sub> -Eq	36.78
ecotoxicity: freshwater - freshwater ecotoxicity potential (FETP)	kg 1,4-DCB-Eq	2.93
ecotoxicity: marine - marine ecotoxicity potential (METP)	kg 1,4-DCB-Eq	3.90
ecotoxicity: terrestrial - terrestrial ecotoxicity potential (TETP)	kg 1,4-DCB-Eq	305.75
energy resources: non-renewable, fossil - fossil fuel potential (FFP)	kg oil-Eq	8.75
eutrophication: freshwater - freshwater eutrophication potential (FEP)	kg P-Eq	0.01
eutrophication: marine - marine eutrophication potential (MEP)	kg N-Eq	0.01
human toxicity: carcinogenic - human toxicity potential (HTPc)	kg 1,4-DCB-Eq	15.89
human toxicity: non-carcinogenic - human toxicity potential (HTPnc)	kg 1,4-DCB-Eq	46.71
ionising radiation - ionising radiation potential (IRP)	kBq Co-60-Eq	1.01
land use - agricultural land occupation (LOP)	m <sup>2</sup> *a crop-Eq	0.72
material resources: metals/minerals - surplus ore potential (SOP)	kg Cu-Eq	71.09
ozone depletion - ozone depletion potential (ODP <sub>infinite</sub> )	kg CFC-11-Eq	0.00
particulate matter formation - particulate matter formation potential (PMFP)	kg PM <sub>2.5</sub> -Eq	0.08
photochemical oxidant formation: human health - photochemical oxidant formation potential: humans (HOFP)	kg NO <sub>x</sub> -Eq	0.21
photochemical oxidant formation: terrestrial ecosystems - photochemical oxidant formation potential: ecosystems (EOFP)	kg NO <sub>x</sub> -Eq	0.21
water use - water consumption potential (WCP)	m <sup>3</sup>	0.23

## REFERENCES

- Alsubal, S., Alaloul, W.S., Musarat, M.A., Shawn, E.L., Liew, M.S., and Palaniappan, P. (2021). Life cycle cost assessment of offshore wind farm: Kudat malaysia case. *Sustainability* (Switzerland) 13. doi:10.3390/su13147943
- Arvesen, A., Birkeland, C., and Hertwich, E.G. (2013). The Importance of Ships and Spare Parts in LCAs of Offshore Wind Power. *Environ. Sci. Technol.* 47, 2948–2956. doi:10.1021/es304509r
- Bang, J.-I., Ma, C., Tarantino, E., Vela, A., and Yamane, D. (2019). Life Cycle Assessment of Greenhouse Gas Emissions for Floating Offshore Wind Energy in California.
- Barahmand, Z., and Eikeland, M.S. (2022). Life Cycle Assessment under Uncertainty: A Scoping Review. *World* 3, 692–717. doi:10.3390/world3030039
- Bhandari, R., Kumar, B., and Mayer, F. (2020). Life cycle greenhouse gas emission from wind farms in reference to turbine sizes and capacity factors. *Journal of Cleaner Production* 277, 123385. doi:10.1016/j.jclepro.2020.123385
- Bhattacharya, S. (2019). *Design of Foundations for Offshore Wind Turbines*, 1st edition. ed. Wiley, Hoboken, NJ, USA.
- Brussa, G., Grosso, M., and Rigamonti, L. (2023). Life cycle assessment of a floating offshore wind farm in Italy. *Sustainable Production and Consumption* 39, 134–144. doi:10.1016/j.spc.2023.05.006
- Centeno-Telleria, M., Yue, H., Carrol, J., Penalba, M., and Aizpurua, J.I. (2024). Impact of operations and maintenance on the energy production of floating offshore wind farms across the North Sea and the Iberian Peninsula. *Renewable Energy* 224, 120217. doi:10.1016/j.renene.2024.120217
- Gaertner, E., Rinker, J., Sethuraman, L., Zahle, F., Anderson, B., Barter, G., Abbas, N., Meng, F., Bortolotti, P., Skrzypinski, W., Scott, G., Feil, R., Bredmose, H., Dykes, K., Shields, M., Allen, C., and Viselli, A. (2020). Definition of the IEA 15-Megawatt Offshore Reference Wind Turbine (Report), Definition of the IEA 15-Megawatt Offshore Reference Wind Turbine. National Renewable Energy Laboratory (NREL).
- Garcia-Teruel, A., Rinaldi, G., Thies, P.R., Johanning, L., and Jeffrey, H. (2022). Life cycle assessment of floating

- offshore wind farms: An evaluation of operation and maintenance. *Applied Energy* 307, 118067. doi:10.1016/j.apenergy.2021.118067
- Hesan, M. (2023). Life Cycle Assessment of an NPK Fertilizer Production with the Focus on Principal Harmful Substances (Master thesis). University of South-Eastern Norway.
- Kaltenborn, B.P., Keller, R., and Krange, O. (2023). Attitudes toward Wind Power in Norway—Solution or Problem in Renewable Energy Development? *Environmental Management* 72, 922–931. doi:10.1007/s00267-023-01870-5
- Karimirad, M. (2014). *Offshore Energy Structures: For Wind Power, Wave Energy and Hybrid Marine Platforms*, 2014th edition. ed. Springer, Cham Heidelberg.
- Lotfizadeh, O. (2024). Life Cycle Assessment of Offshore Wind Farms – A Comparative Study of Floating Vs. Fixed Offshore Wind Turbines (Master thesis). University of South-Eastern Norway, Porsgrunn.
- Manufacturing Factsheets [WWW Document]. (2024). . HIE. URL <https://www.offshorewindscotland.org.uk/the-scottish-offshore-wind-industry/manufacturing-factsheets/www.offshorewindscotland.org.uk> (accessed 3.16.24).
- Narayanan, V.L. (2023). Reinforcement learning in wind energy - a review. *International Journal of Green Energy*. doi:10.1080/15435075.2023.2281329
- Skår, E.H. (2022). An Assessment of the Global Warming Potential of Marine Operations Related to Decommissioning of Offshore Wind Farms (Master thesis). NTNU.
- Struthers, I.A., Avanessova, N., Gray, A., Noonan, M., Thomson, R.C., and Harrison, G.P. (2023). Life Cycle Assessment of Four Floating Wind Farms around Scotland Using a Site-Specific Operation and Maintenance Model with SOVs. *Energies* 16. doi:10.3390/en16237739
- Supplementary Materials-Life Cycle Assessment of Offshore Wind Farms – A Comparative Study of Floating Vs. Fixed Offshore Wind Turbines-Master Thesis-Omid Lotfizadeh, (2024). doi:10.23642/usn.25818493.v1
- Yildiz, N., Hemida, H., and Baniotopoulos, C. (2021). Life cycle assessment of a barge-type floating wind turbine and comparison with other types of wind turbines. *Energies* 14. doi:10.3390/en14185656

# Computationally Efficient Optimization of Long Term Energy Storage Using Machine Learning<sup>\*</sup>

Simon Karlsson, Stavros Vouros, Kristian Sandström,  
Konstantinos Kyprianidis

Mälardalen University, Västerås, CO 72123 Sweden (e-mail:  
simon.karlsson@mdu.se

---

## Abstract:

Energy storage can be charged when energy is cheap and discharged when it is expensive to make an energy system more profitable or used to make the plant operation more efficient to reduce CO<sub>2</sub> emissions. To optimize long term energy storage with conventional methods a long time horizon must be used. When the long term energy storage is combined with a complex energy system the computational cost becomes large when using conventional methods. To reduce the time horizon, an algorithm will be used to decide the state of charge of the long term energy storage at the end of the day. This algorithm is trained using machine learning with data of the optimal state of charge obtained by running computationally heavy long time mixed integer linear programming ahead of time. Then a one-day or week mixed integer linear programming optimization will be done for the production planning. The seasonal patterns of the long term energy storage can then be captured while giving the plant operator a simple one-day or week production plan. A case study will be done with a combined heat and power plant system with 4 boilers, a long-term thermal storage, and a hydrogen storage system. Using this method the complexities of a multi energy system with long term energy storage can be captured while doing day ahead production planning.

*Keywords:* Energy, Optimization, Energy Storage, Machine Learning, Unit Commitment, Production Planning

---

## 1. INTRODUCTION

Energy storage is an important technology in the transition to more sustainable energy system since the energy generated from variable renewable energy sources will not match up with demand. This leads to energy having to be stored to meet demand without oversizing the energy generation and curtailing energy. Some types of renewable energy generation such as solar or wind also have seasonal patterns which can require long term energy storage (LTES) for efficient operation of the energy system International Energy Agency (2024). Because of this, the optimization of LTES is important to help the efficient transition towards a more sustainable energy system. For example, Brey et al. investigate how hydrogen could be used as seasonal energy storage in Spain and conclude that it could be used to smooth out seasonal imbalances Brey (2021).

There are different kinds of electricity markets, in some of these markets like Nordpool in northern Europe. Trading is done with both electricity users and suppliers placing bids and then a price is decided depending on where these bids meet Nordpool (2024). In this system, the bidding period

is 1 hour and because of this, there are requirements on the computational speed of the optimization process for electricity suppliers. To optimize LTES with conventional methods like mixed integer linear programming (MILP) a long time horizon must be used which can make the optimization computationally expensive. This time can be too long to make bids on the electricity market especially if the optimization has to be run several times to run different uncertainty scenarios.

Saletti et al. use linear programming (LP) for the long time horizon (LTH) while MILP is used for the short time horizon (STH) Saletti et al. (2022). This method has a fast solution time, however, the solving time of the (LTH still depends on system complexity. The objective of the optimization is to meet the heat and electricity demand of a hospital and not maximize profit by selling to the electricity market. Marzi et al. use MILP to do day ahead scheduling of a multi energy system with LTES considering uncertainty Marzi et al. (2023). The computation time for their method is, however, too long to do bidding in less than one hour.

In a study by Bischi et al. a rolling horizon is used together with typical weeks to optimize plant operation with MILP considering the entire year. The goal of this optimization is however not to consider how the state of charge (SOC)

---

<sup>\*</sup> This work is supported by project ProPlan under the RESILIENT competence center, financed by the Swedish Energy Agency (2021-90273) and co-financed by Mälardalen University and industrial partners. The support received by Mälarenergi AB and Eskilstuna Strängnäs Energi och Miljö AB is appreciated.

of the storage will change over the year but to make sure that yearly emission constraints are met optimally.

Bruninx et al. optimize an energy system with energy storage using unit commitment by considering reserve capacity in a computationally efficient way Bruninx and Delarue (2017). The time horizon in this study is however 24 hours so the focus is not on LTES. Optimization of a compressed air energy storage is done by Ghaljehei et al. by using using stochastic programming and mixed integer nonlinear programming. Here the time horizon is also 24 hours so it is not fit for LTES Ghaljehei and Golkar (2017).

System states are used to optimize medium and long term energy storage in a study by Worgin et al. Worgin et al. (2016). Here some states of the system are defined and clustered and based on what cluster the system is in the storage is operated accordingly.

A life cycle analysis with energy storage optimization is done by Dong et al. where the energy storage is optimized using a representative day for each season Dong et al. (2023). This representative day is used to calculate how much the storages will be charged or discharged during each season to store energy over seasons. This method will however not work when doing operational optimization since it will just have the same operation every day and not ex discharge the storages more for a day with high electricity prices

Mi et al. use multi timescale optimization to do generation and expansion planning where the longest timescale is one month Mi et al. (2021). Here the longer timescale is however used to optimize capacity credits and not to optimize LTES. Zhang et al. also use multi timescale optimization to optimize the operation of an energy system with hydrogen energy storage Zhang et al. (2023). Here a rolling horizon optimization is used where different kinds of energy have different time resolutions. Here two days ahead is used to optimize the energy storage using MILP. Su et al. use multiple timescales and add a flexibility requirement to make the energy system more prepared for uncertain future disruptions Su et al. (2023). Here a short, medium, and long time horizon is used where the long time horizon is one week.

In a study by Bahlwan et al. the design and operational operation of an energy system with long term thermal energy storage is optimized Bahlwan et al. (2022). Here switch on priority is used to do the operational optimization where one energy conversion technology is used first and only if this technology can not supply the demand the next conversion technology is used. This method will however not work well for a system where the operational cost of different technologies changes and there is no electricity demand but instead electricity is sold to the grid.

Reinforcement learning is used by Alabi et al. to control an energy system with energy storage and carbon capture Alabi et al. (2023). Here reinforcement learning is used to control the power output of the energy units and not to optimize any kind of LTES. Sleptchenko et al. use LP to optimize multiple different energy storage technologies as a part of an energy system Sleptchenko and Sgouridis (2019). In this study, the focus is not on computational speed but on the seasonal patterns of the storage operation.

Water value is an optimization method to optimize how hydropower reservoirs are used. Here a value of the water in the reservoir is calculated and used to determine if the reservoir should be discharged Helseth et al. (2017); Jahns et al. (2020). This method is quite computationally costly if used for daily production planning with the optimization by Helseth et al. taking between 28 and 40 hours.

### 1.1 Current Work

In this paper day ahead planning will be done for an energy system with LTES where the goal is to make as much profit as possible by selling electricity to the grid while supplying the required district heating (DH) demand. Instead of using LP or MILP to optimize the long term behavior of the system a machine learning (ML) model will be used to predict the end of day SOC of storages and then MILP will be used to optimize the daily operation with this SOC as a constraint.

The contribution of this work will be (i) to develop a new faster method to optimize LTES which allows for scenario analysis in production planning or be used in studies where the optimization has to be used many times. (ii) Analyze the effect of system complexity and optimization horizon. (iii) Test which input features give the best prediction.

## 2. METHODOLOGY

Because MILP is slow over long time horizons, a ML algorithm is used to predict the SOC of the storage's at the end of the day or week so the MILP can run for one day or week instead of a longer time. The ML algorithm is trained using optimal SOC data obtained by running the MILP on historical electricity price and DH demand data. Because only a few years of DH demand data was available some synthetic electricity price and DH demand data were also generated for training. This was done by using the probability density function (PDF) which can be seen in Eq. (1) to decide how much the scenarios should deviate from the real data like in Marzi et al. (2023) but with some changes. These changes are, instead of using the PDF to decide the deviation from the real data the PDF is used to decide the change in deviation at each timestep. Some of the spikes in the electricity price were also randomly removed and new ones were added so the spikes in the electricity price would not occur at the same time of year in all the generated scenarios. This synthetic data was run through the MILP to get optimal SOC data for training the ML algorithm. A flowchart of how the training and optimization are done can be seen in **Fig. 1**.

$$PDF(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

A case study based on the system seen in **Fig. 2** where the full system has four combined heat and power (CHP) plants, one TES (thermal energy storage) which uses water to store heat. There is also a hydrogen energy storage (HES) with an electrolyzer to convert electricity to hydrogen, a hydrogen storage tank, and then a fuel cell (FC) to convert the hydrogen back to electricity For both the electrolyzer and FC there are some losses in the form of heat which is

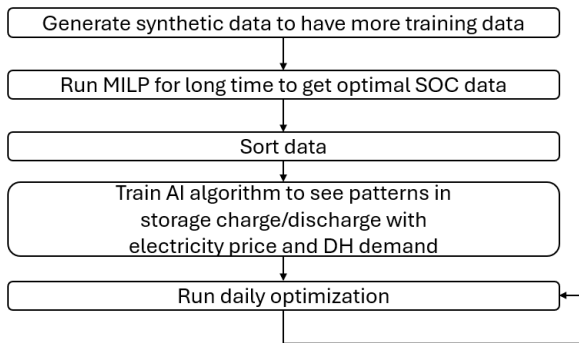


Fig. 1. Flowchart of method

used to both charge the TES and provide heat to the DH network. The system is used to provide the DH demand to the district heating network and sell electricity to the electricity grid. To evaluate how system complexity affects the current methods performance some other cases were also evaluated, these are one case with one boiler, the HES, and the TES, one case with one boiler and the TES, and one case with one boiler and the HES.

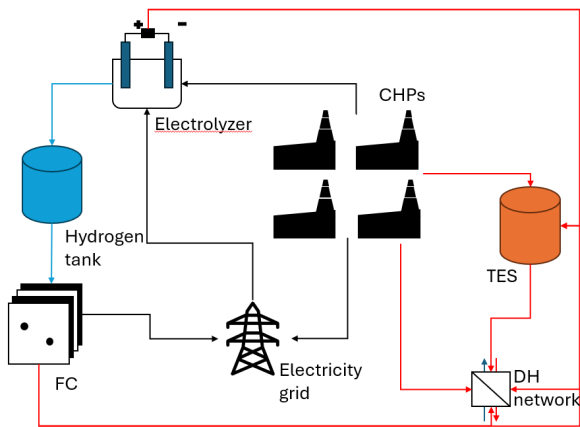


Fig. 2. Full system layout

### 2.1 Mixed Integer Linear Programming (MILP)

The general MILP formulation can be seen in Eq. (2), where  $x$  is a vector of the decision variables,  $c$  is a vector with the relationship between the decision variables, and  $A$  and  $b$  make the constraints where  $A$  is a matrix and  $b$  a vector. When running the MILP the binary constraints were relaxed to increase computational speed since generating training data without relaxing binaries was too computationally costly with the used hardware. However, if the model is simple enough or there is enough computing power the full model could be run with binary constraints. Other methods could also be used to increase the computational speed of the MILP.

$$\begin{aligned} & \min(c^T x) \\ & \text{st. } Ax \leq b \end{aligned} \quad (2)$$

The objective function can be seen in Eq. (3) where  $C_{eco}$  is the economic cost and  $C_{change}$  is a penalty to punish

uneven operation of the storages and boilers. In the results when profit is referred to it refers to  $C_{eco}$ . The constraints that are considered in the MILP model can be seen in Table 1 with what constraints apply to each unit.

$$C = C_{eco} + C_{change} \quad (3)$$

Table 1. List of constraints for MILP model

Constraint	CHP	HES	TES
Max/min power	✓	✓	✓
Ramp up/down	✓	✓	✓
SOC	×	✓	✓
Min up/down time	✓	×	×
on/off status	✓	✓	✓
electricity to heat ratio	✓	×	×
SOC start and end of time horizon	×	✓	✓
Heat loss to environment	×	×	✓
Startup status	✓	×	×
DH demand met	-	-	-
Transmission capacity out of plant	-	-	-

The MILP was tested in three different ways, the first is to just run the MILP for 1 year to get the optimal behaviour of the system. The second way is to give the MILP a constraint at the end of day SOC and then run the MILP for 36 hours but only taking the operation from the first 24. The third option is to use a rolling horizon optimization Bischi et al. (2019); Marquant et al. (2015) where the optimization is done daily with a one week time horizon. For this method the constraint on the SOC on the storages was also set to happen after one week. A optimality gap of 1% was used for the MILP optimization

### 2.2 Machine Learning (ML) Algorithm

Some different ML algorithms were tested these are deep neural network (DNN), random forest (RF), historic gradient boosting (HGB), and Gaussian regression (GR). For all of these hyper parameter optimization was done and for the DNN different architectures of the network were also tested. The variables being predicted are the optimal daily or weekly charge and discharge from the HES and TES where training data is retrieved by running the MILP with a long time horizon. Some different input features were tested to get the lowest prediction error possible. The training and testing data were split by having training data be the data generated based on the first year and the testing data be the real data from the second year.

Preprocessing of the data was done before passing it to the ML algorithm. This preprocessing consisted of calculating the mean, max, and minimum daily electricity price and DH demand and monthly and weekly mean electricity price and DH demand. The data was also scaled with the electricity price, DH demand, SOC of the storages, day of the year, and weekday being scaled between 0 and 1 and the charge/discharge of the storages being scaled between -1 and 1 where -1 is fully discharging and 1 is fully charging. The loss metric used during the training of the ML models is mean square error.

A lot of the charge and discharge data of the storages is distributed around 0 to avoid any bias in the model training weights were used in the loss function to make

all charge and discharge amounts be equally represented. This was done using DenseWeight which applies weights to different values of the training data based on kernel density estimation Steininger et al. (2021).

### 3. RESULTS

#### 3.1 Long term Mixed Integer Linear Programming MILP

**Figure 3** shows how the SOC of the HES and TES change when using MILP to optimize the system. The data used for optimization is from 2017, the year the ML algorithm makes its prediction. The storages does not start and end at the same SOC since the MILP optimization was done over 3 (2016-2018) with the SOC being constrained to be the same at the beginning of 2016 and the end of 2018.

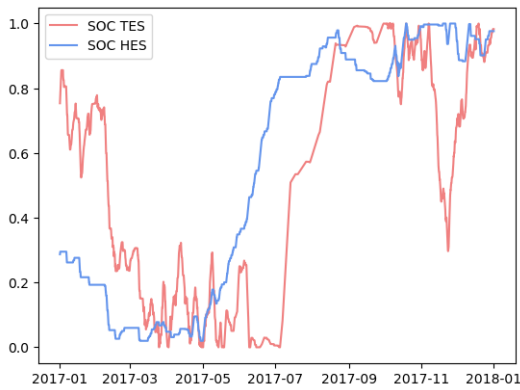


Fig. 3. SOC of storages based on MILP over predicted year

#### 3.2 Full System One Day Prediction Horizon

The MAE as a percentage of the maximum occurred charge/discharge of the storages for different ML algorithms can be seen in **Table 2**. As can be seen, the MAE of HGB is the lowest, however, this MAE is achieved by having the charge/discharge around zero all the time which does not lead to a good operation of the storages. The DNN on the other hand makes predictions that are based on the features and most of the time the decision to charge or discharge the storage is correct. The amount charged or discharged is however often wrong. This leads to the DNN operating the storages in a better way than HGB even though the MAE is higher. RF operates the storages in a similar way as HGB in that it tries to keep the charge/discharge around zero. GR operates the storages in a way that is somewhere between the strategy of the DNN and HGB. Because of this, the DNN is used as the ML algorithm for the rest of the results.

Table 2. Prediction performance of ML algorithms

ML method	MAE HES	MAE TES
DNN	25%	20%
RF	22%	19%
HGB	14%	16%
GR	23%	19%

The ML algorithm predicts the optimal SOC of the storages at the end of the day or week based on the features that

can be seen in **Table 3**. Different combinations of features were tested but these were chosen since they gave the lowest mean absolute error (MAE). The data from the long term MILP and ML prediction using a DNN (deep neural network) can be seen in **Figs. 4** and **5**, here a one-day prediction horizon was used. As can be seen, the prediction error is evenly spread except for predicting too low values when the HES is charged at maximum power. The HES has a MAE of 4500 kWh and the TES has a MAE of 332 000 kWh. This MAE is quite high, around 25% and 20% of the maximum daily charge/discharge power that occurred for the HES and TES. This error is however not important as long as the ML algorithm can give predictions that have a good operation of the storages in the daily MILP optimization.

In future research, this error could be reduced either by using a more complicated method such as first classifying if the storage will be charged, discharged, or not used, and then after that having 2 different specialized models for charging and discharging for each of the storages. Different ML models could also be used for different periods of the year. Another way to improve the results could be to use reinforcement learning and a one-day or week MILP model to more directly optimize based on the objective function.

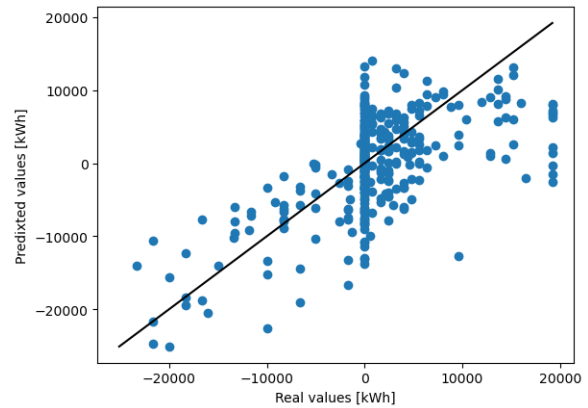


Fig. 4. Correlation between predicted and real charge/dis-charge for HES using DNN with a one-day prediction horizon

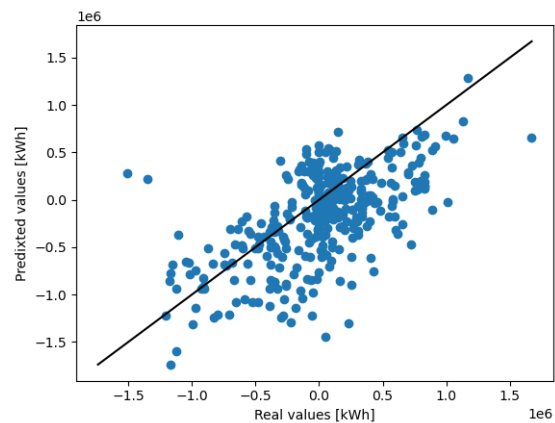


Fig. 5. Correlation between predicted and real charge/dis-charge for TES using DNN with a one-day prediction horizon



Table 3. Features used for ML

Feature	HES	TES
Current day mean electricity price and DH heat demand	✓	✓
Current day max electricity price and DH heat demand	✓	✓
Current day min electricity price and DH heat demand	✓	✓
Two weeks of mean electricity price	✓	✓
Two weeks of mean DH demand	✗	✓
Two months of mean electricity price and DH demand	✓	✓
Time of year	✓	✓
Day of week	✓	✓
SOC of storages	✓	✓

The DNN performs better in some parts of the year and worse in others as can be seen in **Figs. 6** and **7**. The accuracy might be able to be improved if multiple ML models were trained for different parts of the year. The algorithm does however still mostly charge and discharge the storages at the correct time but the amount charged or discharged is often wrong. For both figures the DNN was trained using data generated based on data from 2016 and then tested using real data from 2017. The optimality gap used for the MILP for both training and testing is 1%.

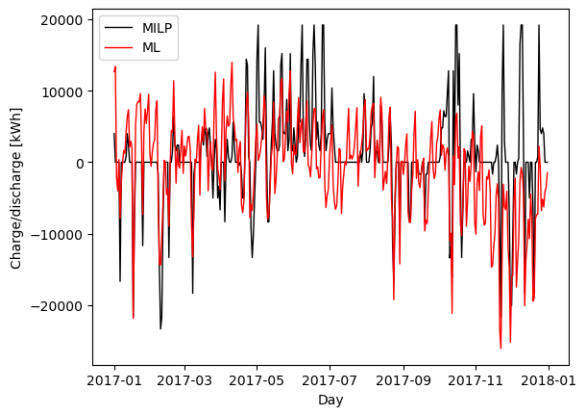


Fig. 6. Comparison charge/discharge HES MILP and DNN

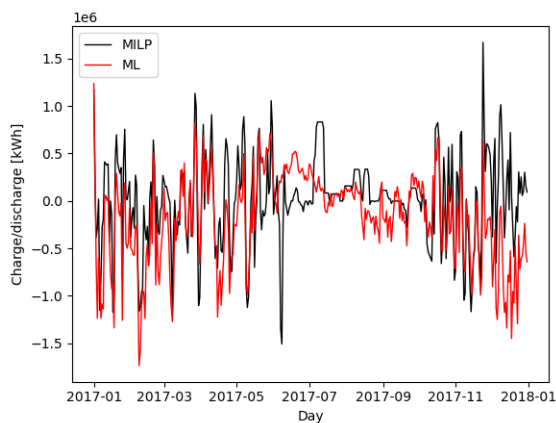


Fig. 7. Comparison charge/discharge TES MILP and DNN

The SOC of the HES for MILP and with constraints from the ML model can be seen in **Fig. 8** and the same for the TES in **Fig. 9**. Both the HES and TES SOC are quite different between using MILP and using DNN model constraints. The important thing here however is not that the SOC of the storages are the same but how profitable the operation of the entire energy system is in both scenarios. This will be discussed in the next section. The SOC pattern for the TES is however similar between the MILP and DNN model constraint with it discharging during the winter and charging during the summer. There is a difference in when the storages is being charged/discharged between the MILP and DNN. The reason for this could be that the DNN gets a low electricity price as an input and therefore charges the storage while the MILP does not charge the storage since it has all the data and knows that there will be an even cheaper electricity price in the future. In reality, a forecast for the electricity price would have to be used to operate the MILP in this way which could make the results of the MILP and DNN more similar.

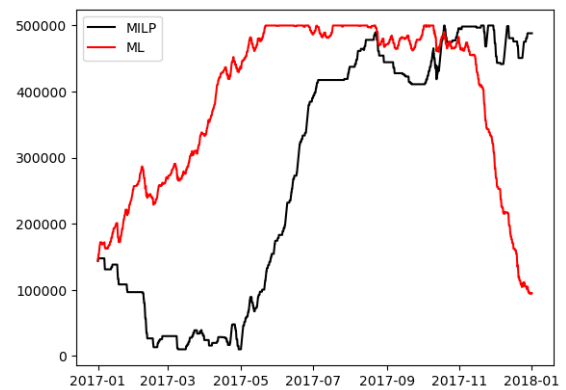


Fig. 8. Comparison SOC HES MILP and DNN

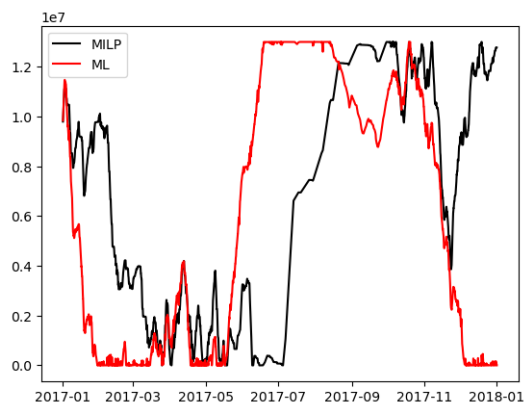


Fig. 9. Comparison SOC TES MILP and DNN

### 3.3 Comparison Cases

**Table 4** compares the profitability of the daily MILP model with the ML constraints and when constraints are taken from the optimal operation of the previous year. The comparison is made as a ratio profitability compared to the long time MILP optimization results. When a one day time horizon is used the daily MILP model with constraints



from DNN outperforms the other models with constraints from the optimal operation of the last year and other ML methods. All of them are also close to the long time horizon MILP optimization being 1, 4, and 6 percent away. Of note here is that the energy storages is only one part of the system so it is not the only factor effecting profitability.

When a one week time horizon is used, both the models with constraints from ML algorithms and last year's optimal operation have the same profitability as the long time MILP optimization. This is because a rolling time horizon is used with the constraint placed at the end of the week but only the operation from the first day is used, then the next day the optimization is run again. This means that even if there is some error in the prediction the operation of the first day can be good since there is no constraint for the SOC at the end of the first day.

Even though the MAE of the prediction is high the profitability is not greatly impacted. This is because the prediction is based on the electricity price DH demand and time of year. This gives a good operation even if the prediction is different from the value from the long time MILP.

Another benefit to using ML to give constraints to the MILP compared to using values from last year is that if a change in the electricity price or DH demand were to occur the operational plan can change. This makes the constraints given by the ML model more robust. The diversity and amount of training data generated and the number of years used to generate training data will also affect how robust the optimization is.

One thing to note when looking at **Table 4** is that the SOC the storages is not constrained to be the same at the end of the year which affects the profitability. The final SOC with the DNN constraints can be seen in **Figs. 8** and **9**. The final SOC when using the optimal results from the last year can be seen in the same figures but looking at the beginning of the year.

Table 4. Profitability comparison with constraints from ML models and taking SOC values from last year MILP optimization

Method	Time horizon	
	1 day	1 week
DNN	0.97	1
Last year MILP	0.96	1
GR	0.94	1
HGB	0.94	1
RF	0.94	1

In **Table 5** the MAE and computational speed of the method using daily and weekly MILP with DNN constraints and the long term MILP model can be seen. Here the MAE is a percentage of the maximum occurred charge/discharge of the storages. As can be seen, the optimization is fast both for the one day and one week time horizon when using constraints from the DNN. When running the MILP for a year with the full model the optimization time is over 30 hours which does not allow for day ahead planning. Another problem when running the model in this way is that a forecast for electricity price and DH demand is needed for the entire year. When using the ML algorithm

only a forecast for the average electricity price for the next 2 weeks and next 2 months is needed. Even this can be removed with some increase in the prediction error.

The MAE of the HES prediction increases when doing a one week prediction while the MAE of the TES prediction decreases. This is likely because the seasonal patterns of the TES are stronger which makes a one week prediction easier since any irregular spikes in temperature will have a lower effect. The HES is more driven by the electricity price which has a less seasonal pattern so in this case the increase in features for the DNN only increase the MAE.

Table 5. MAE for different time resolutions and horizons

Time horizon	MAE HES	MAE TES	Computational time
1 day	25%	20%	1.78 s
1 week	30%	18%	12.33 s
Long term MILP	-	-	180.2 s*

\*With binary constraints relaxed, the full model takes over 30 hours to run

**Table 6** shows the error when doing predictions based on data from a simpler system. The prediction is slightly better for the system with only one boiler. The prediction is better on simpler systems since the behavior of the system becomes less complex and therefore easier to predict.

Table 6. Comparison different systems

System	MAE HES	MAE TES
One boiler only TES	-	20%
One boiler only HES	22%	-
One boiler HES and TES	22%	18%
Four boiler HES and TES	25%	20%

#### 4. DISCUSSION

This method is fast enough to implement in real-time, when doing so retraining of the ML model should be done to catch any new patterns in electricity price or DH demand. The period between retraining will have to be decided based on testing different periods. When retraining the algorithm data could be generated again to increase the training data since the generated data is created based on real data and will therefore have some similar patterns.

Some things are required for it to be possible to use this method, the first is some historical data that can be used for training and creation of synthetic data or a way of creating realistic synthetic data without any real data. Some long-term energy storage is also needed for this method to be effective, if no long-term energy storage exists conventional methods are more suitable for optimization.

The use case for this kind of optimization method is in cases where the optimization has to be done in a short time or where the optimization has to be done a lot of times, for both of these cases LTES should also be a part of the energy system. For the case where optimization has to be done fast it could be at a powerplant where the MILP model is too complex to optimize over a long time horizon, then this method can be used to speed up the optimization. For a case where optimization has to be done many times, there could be a case where the MILP optimization is part of an

inner loop where it has to run many times per iteration of some other optimization layer.

Since the synthetic data is only used for training the ML algorithms and is generated without using any of the testing data the use of synthetic data should not have any negative impact on the results. The use of synthetic data might also not be needed if enough historical data is available, other methods of generating synthetic data could also be used.

## 5. CONCLUSIONS

Using ML to reduce the time horizon of a MILP model by constraining the SOC of LTES gives a similar economic operation to letting the MILP run over a long time horizon. The MAE of the prediction is large but the economic operation is still good with this method. This method outperforms using the past years storage operation when running the MILP daily and has an equal performance when running the MILP weekly.

The ML method that gives the best operation of the storages is a DNN.

## ACKNOWLEDGEMENTS

This work is supported by project ProPlan under the RESILIENT competence center, financed by the Swedish Energy Agency (2021-90273) and co-financed by Mälardalen University and industrial partners. The support received by Mälarenergi AB and Eskilstuna Strängnäs Energi och Miljö AB is appreciated.

## REFERENCES

- Alabi, T.M., Lawrence, N.P., Lu, L., Yang, Z., and Gopaluni, R.B. (2023). Automated deep reinforcement learning for real-time scheduling strategy of multi-energy system integrated with post-carbon and direct-air carbon captured system. *Applied Energy*, 333. doi:10.1016/j.apenergy.2022.120633.
- Bahlawan, H., Losi, E., Manservigi, L., Morini, M., Pinelli, M., Spina, P.R., and Venturini, M. (2022). Optimization of a renewable energy plant with seasonal energy storage for the transition towards 100 *Renewable Energy*, 198, 1296–1306. doi:10.1016/j.renene.2022.08.126.
- Bischi, A., Taccari, L., Martelli, E., Amaldi, E., Manzolini, G., Silva, P., Campanari, S., and Macchi, E. (2019). A rolling-horizon optimization algorithm for the long term operational scheduling of cogeneration systems. *Energy*, 184, 73–90. doi:10.1016/j.energy.2017.12.022.
- Brey, J.J. (2021). Use of hydrogen as a seasonal energy storage system to manage renewable power deployment in Spain by 2030. *International Journal of Hydrogen Energy*, 46, 17447–17457. doi:10.1016/j.ijhydene.2020.04.089.
- Bruninx, K. and Delarue, E. (2017). Improved energy storage system unit commitment scheduling. doi:10.1109/PTC.2017.7981000.
- Dong, H., Fu, Y., Jia, Q., Zhang, T., and Meng, D. (2023). Low carbon optimization of integrated energy microgrid based on life cycle analysis method and multi time scale energy storage. *Renewable Energy*, 206, 60–71. doi:10.1016/j.renene.2023.02.034.
- Ghaljehei, M. and Golkar, M.A. (2017). Effect of optimal generation scheduling of compressed air energy storage and wind power generation on economic and technical issues. doi:10.1109/SGC.2017.8308854.
- Helseth, A., Fodstad, M., Askeland, M., Mo, B., Nilsen, O.B., Pérez-Díaz, J.I., Chazarra, M., and Guisández, I. (2017). Assessing hydropower operational profitability considering energy and reserve markets. *IET Renewable Power Generation*, 11, 1640–1647. doi:10.1049/iet-rpg.2017.0407.
- International Energy Agency (2024). Managing the seasonal variability of electricity demand and supply. URL <https://www.iea.org/reports/managing-the-seasonal-variability-of-electricity-demand-and-supply>.
- Jahns, C., Podewski, C., and Weber, C. (2020). Supply curves for hydro reservoirs – estimation and usage in large-scale electricity market models. *Energy Economics*, 87. doi:10.1016/j.eneco.2020.104696.
- Marquant, J.F., Evins, R., and Carmeliet, J. (2015). Reducing computation time with a rolling horizon approach applied to a milp formulation of multiple urban energy hub system. volume 51, 2137–2146. Elsevier B.V. doi:10.1016/j.procs.2015.05.486.
- Marzi, E., Morini, M., Saletti, C., Vouros, S., Zaccaria, V., Kyprianidis, K., and Gambarotta, A. (2023). Power-to-gas for energy system flexibility under uncertainty in demand, production and price. *Energy*, 284. doi:10.1016/j.energy.2023.129212.
- Mi, Y., Liu, C., Yang, J., Zhang, H., and Wu, Q. (2021). Low-carbon generation expansion planning considering uncertainty of renewable energy at multi-time scales. *Global Energy Interconnection*, 4, 261–272. doi:10.1016/j.gloi.2021.07.005.
- Nordpool (2024). Nordpool. URL <https://www.nordpoolgroup.com/en/>.
- Saletti, C., Morini, M., and Gambarotta, A. (2022). Smart management of integrated energy systems through co-optimization with long and short horizons. *Energy*, 250. doi:10.1016/j.energy.2022.123748.
- Sleptchenko, A. and Sgouridis, S. (2019). Joint optimization of energy production storage. doi:10.1109/IEA.2019.8714874.
- Steininger, M., Kobs, K., Davidson, P., Krause, A., and Hotho, A. (2021). Density-based weighting for imbalanced regression. *Machine Learning*, 110, 2187–2211. doi:10.1007/s10994-021-06023-5.
- Su, S., Tan, D., Li, X., Li, X., Chen, Z., Lian, R., and Soronga, B.K.L. (2023). Multi-time scale coordinated optimization of new energy high permeability power system considering flexibility requirements. *Journal of Electrical Engineering and Technology*, 18, 815–828. doi:10.1007/s42835-022-01244-7.
- Wogrin, S., Galbally, D., and Reneses, J. (2016). Optimizing storage operations in medium- and long-term power system models. *IEEE Transactions on Power Systems*, 31, 3129–3138. doi:10.1109/TPWRS.2015.2471099.
- Zhang, L., Dai, W., Zhao, B., Zhang, X., Liu, M., Wu, Q., and Chen, J. (2023). Multi-time-scale economic scheduling method for electro-hydrogen integrated energy system based on day-ahead long-time-scale and intra-day mpc hierarchical rolling optimization. *Frontiers in Energy Research*, 11. doi:10.3389/fenrg.2023.1132005.

# A Battery Model for Transportation and Stationary Applications

E. Dahlquist\* M. Azaza\* M. Shabani\*\*\*

J. R. Vazquez\*\* M. M. Nezhad\*\*\* A. Fattouh\*\*\*\* A. D. Fentaye\*

\*Mälardalen University (MDU), School of Business, Society and Engineering, Future Energy Center, Universitetsplan 1, 72123 Västerås, Sweden (e-mail: [erik.dahlquist@mdu.se](mailto:erik.dahlquist@mdu.se), [maher.azaza@mdu.se](mailto:maher.azaza@mdu.se), [amare.desalegn.fentaye@mdu.se](mailto:amare.desalegn.fentaye@mdu.se))

\*\*Hitachi Energy, Valhallavägen, 77131 Ludvika, Sweden (e-mail: [josefin.rojas.vazquez@hitachi.se](mailto:josefin.rojas.vazquez@hitachi.se))

\*\*\* Mälardalen University (MDU), Department of Sustainable Energy Systems, Universitetsplan 1, 72123 Västerås, Sweden (e-mail: [meysam.majidi.nezhad@mdu.se](mailto:meysam.majidi.nezhad@mdu.se))

\*\*\*\* Mälardalen University (MDU), School of Innovation, Design and Technology (IDT), Division of Product Realization (PR), 63105 Eskilstuna, Sweden (e-mail: [anas.fattouh@mdu.se](mailto:anas.fattouh@mdu.se))

**Abstract:** Batteries are used in electric vehicles as well as in stationary applications. In electric vehicles, high energy density, as kilowatt hour per kilogram (kWh/kg), is needed while stationary applications are less sensitive to the energy density. Principally, it may be a good idea first to use batteries for transportation applications and then when capacity has reached a certain level start using them for other applications in a “second life”. Both for optimizing the performance of operations in 1st and 2nd life and for making fair commercial agreements when selling used batteries for second life applications, there is a need to make predictions of Remaining Useful Life (RUL) and State of Health (SoH). For this purpose, battery models are needed. This paper presents a methodology for building a useful battery model based on our experiments and literature data. Single cells and cells in series of Nickel Manganese Cobalt (NMC) batteries and Lithium Iron Phosphate (LFP) batteries have been cycled. Electrochemical Impedance Spectra (EIS) and differential capacity (dQ/dV) have been measured for each cycle. These data were then used to develop SoH and RUL models using various regression methods. The developed models are described and discussed, and the results are presented in the paper.

**Keywords:** Battery model, electric vehicles (EV), stationary applications, kilowatt hour per kilogram (kWh/kg), Nickel Manganese Cobalt (NMC), Lithium Iron Phosphate (LFP), Electrochemical Impedance Spectra (EIS), differential capacity (dQ/dV), regression.

## 1. INTRODUCTION AND RELATED WORK

There is a strong interest in following the degradation of batteries during the first life to give the possibility to predict the remaining useful life (RUL) for the battery, especially for the second-life use of batteries. By following the state of health (SoH) continuously, correlating it to how the battery has been used, and following it until it is totally worn out, reasonably good prediction models can be determined and used. This information can be utilized for 2<sup>nd</sup> life use of batteries (Chirumalla et al., 2023, 2024).

In this paper, we want to extract experience from what has been done earlier and add to it our own battery modeling approach for the simulation of batteries used in Vehicles. For second life use of batteries there is a high interest to use for power grid balances as shown in e.g. Dahlquist et al. (2023).

Many researchers have modelled battery performance in different ways. Pelletier et al. (2017) focused on modelling cycle-life as a function to the depth of discharge (DOD) and discharged rate relative to the battery maximum capacity (C-rate). Ahmadiana et al. (2018) modeled battery resistance growth versus state of charge (SoC) and battery degradation as a function to DOD. Maheshwari et al. (2020) have modelled

cycling life vs C-rate using a non-linear model. O’Kane et al. (2022) have used the PyBaMM, an open source modelling environment for simulation of the effect of different variables important for degradation of Li-ion batteries. Four degradation mechanisms are coupled in PyBaMM. The most important ones are the loss of lithium inventory and the loss of active material. For the same cell, five different pathways have been evaluated. Lam and Bauer (2012) performed a lot of cycling of LFP batteries and from the experiments, an empirical model was constructed, which was capable of modeling the capacity fading in electric vehicles (EV) battery cells under most operating conditions. Ravali and Raju (2023) developed a Lithium-Ion Battery model for estimation of degradation capacity and SoC using Sigma Point Kalman filter. After one thousand cycles, the amount of capacity faded from 24Ah to 20.5Ah at 25°C.

de la Vega et al. (2023) have proposed to monitor battery performance by measuring instantaneous terminal cell voltage as a function to SoC during the first discharge cycle, as a reference cycle, and the  $n^{\text{th}}$  cycle. The SoC interval  $[\text{SoC}_{\text{min}}, \text{SoC}_{\text{max}}] = [0.55, 0.75]$  was chosen to calculate the  $\Delta\text{VRMS}$  indicator, which is terminal voltage difference between first and  $n^{\text{th}}$  cycle at  $\text{SoC}_{\text{min}}$  and  $\text{SoC}_{\text{max}}$ . This is an interesting

approach to give the possibility to follow the degradation continuously if the values at  $SoC_{min}$  and  $SoC_{max}$  are stored continuously.

Shamarova et al. (2022) have developed a method utilizing data from electrical circuit models (ECM) where dependency of ECM parameters on the electrochemical properties of the battery was examined in using a pseudo-two-dimensional (P2D) model. This is combining physical and statistical modelling approaches. Wildfeuer et al. (2023) made a set of experiments studying impact on SoH for capacity, resistance, Li-inventory, positive electrode losses, and negative electrode losses for SoC 10-100% and temp 20-60°C for NAC batteries. Panchal et al. (2017) did similar experimental studies for LFP batteries.

Drive cycles with different modes like acceleration, constant speed, and deceleration in both highway and city driving were implemented at -6°C, 2°C, 10°C, and 23°C ambient temperatures with all accessories on. 4.6% capacity fade occurred over 3 months of driving. The empirical degradation model was fitted to these data, and an extrapolation estimated that 20% capacity fade would occur after nine hundred daily drive cycles. This is a high degradation rate, but experimental data and model were close for the 3 months test period. Zhang et al. (2023) have compared LFP and NMC batteries. Degradation characteristics during charging of LiFePO4 (LFP)/Graphite batteries at voltages of 3.65–4.8V and Li(Ni0.5Co0.2Mn0.3)O2 (NCM)/Graphite batteries at 4.2–4.8V at -10 °C with currents of 0.2–1A were determined. The loss of active material (LAM) causes the height of the highest IC peak ( $dQ/dV-V$ ) to decrease for a given voltage, while the loss of Lithium inventory (LLI) shifts the DV curve ( $dV/dQ-Q$ ) toward lower capacities.

It can also be interesting to see what measurements on real vehicles indicate with respect to capacity losses. Salazar and Bengoechea (2021) have summarized information reported by different Tesla Model 3 owners. One had a decrease of capacity by 4.8% during 136,000km operation, another 2.3% loss during 22,000km, when the vehicle was charged to 10% five days a week. A third had a 2.2% decrease during 32,000km of SoH for 12 months operation. In this case all the cars had LFP batteries.

Shen et al. (2019) tried to make RUL predictions. They were working with NASA data set and the CALCE data set. They compared their own model to another approach. Still, the value is to use common data sets for comparing different modelling approaches. Uddin et al. (2016) used an approach with identification and tracking of electrochemical battery model Parameters. This combines physical and statistical methods. The method was demonstrated on a 3.03Ah LiC6/NCA battery stored at 45°C with 50% SoC for 476 days.

Rahbari et al. (2018) used another approach with an Adaptive Neuro-Fuzzy Inference System for SoH of real-life plug-in hybrid electric vehicles (PHEVs). The model was representing the experimental data in a good way. Dai et al. (2018) showed a SoH estimation method by using prior knowledge-based neural network (PKNN) and Markov chain for a single lithium-ion battery. Shi et al. (2019) used another method with estimation of the state of health (SoH) for a lithium-ion battery based on the ohm internal resistance  $R_0$ . They were considering the variation of  $R_0$  with the state of charge (SoC), which was new.

This overview covers a broad spectrum of methods, although many other techniques are also utilized. From all this, we developed a simplified battery degradation model that can be adapted to different types of batteries and with reasonable values for impact of different variables like temperature, C-rate, DOD/DOC and calendric time.

The paper is organized as follows. Section 2 presents the experiment setup while Section 3 presents the experimental data, both our data and other data from literature. In Sections 4 and 5, we develop a battery degradation model and power demand model from vehicle, respectively. The paper ends with discussion and conclusions in Section 6.

## 2. EXPERIMENTS SETUP

### 2.1 Testing of battery cells

Single cells can be tested with Electrochemical Impedance Spectroscopy (EIS). The spectra are made by measuring the current and the capacity as a function of voltage as the frequency of the supplied current is going from 1000Hz to 0.001Hz. At high frequencies, we see the resistance in the electrolyte, at mid frequencies capacitance over the electrode surface, and at low frequencies the inner resistance of the cells as such. By following the EIS at the start and after a number of cycles, you can get both a quantitative measure of the overall degradation over time as well as an indication of what mechanisms in the battery cells are causing this.

Another measure is to follow  $dQ/dV$  or  $dV/dQ$ , as a function of  $V$ , where  $Q$  is cell capacity (Ah) and  $V$  voltage. By measuring and plotting this over cell cycles a battery performance pattern is achieved. This can be measured continuously during the use of the cell, which is not possible with the EIS, and thus can be a good complement. Figure 1 depicts a system for prediction of RUL, SoH and SoC.

There is a correlation between DoD (depth of discharge)/DoC (depth of charge) and degradation rate. Figure 2 shows this correlation presented by Qadrdan et al. (2018). Real operation with Tesla cars still indicates that this curve is not relevant for NAC batteries in “real life” operations.

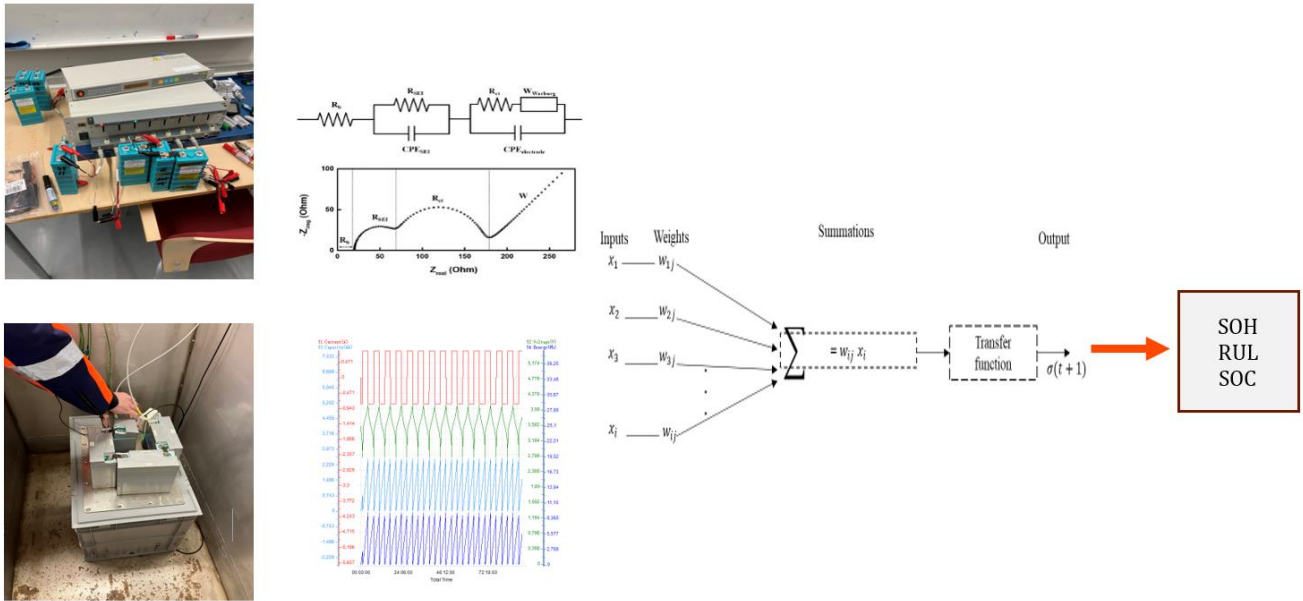


Fig. 1. Experimental setup to collect data that will be used to develop algorithms for prediction of SoH, RUL, and SoC.

$DoD = \frac{\text{removed amount of charge}}{\text{maximum available amount of charge}} = Qd/C * 100$  [%]

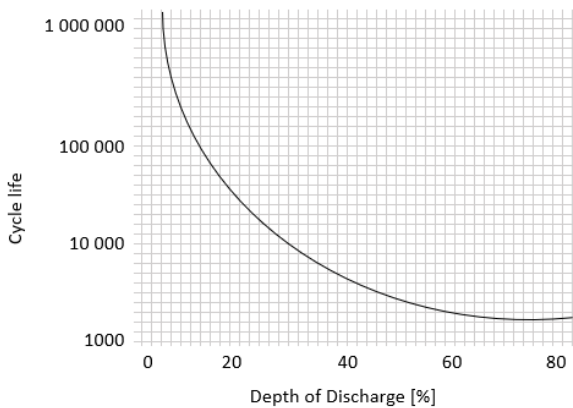


Fig. 2. Battery cycle life as a function to depth of discharge. Adapted from Qadrdan et al. (2018).

The SoH can be shown in an EIS, Electrochemical impedance spectrum. In Fig. 3, we see how the pattern in a  $Z_{real}$  vs  $Z_{im}$  is looking like when a frequency scan is made from 1000Hz to 0.001Hz. Higher frequency is to the right. Closest to y-axis we have ohmic resistance ( $R_{ohm}$ ):

$$\begin{aligned} Z(\omega) &= \tilde{V}(\omega) \tilde{i}(\omega) \\ &= |\tilde{V}(\omega) \tilde{i}(\omega)| (\cos\phi(\omega) + j\sin\phi(\omega)) \\ &= Zr + jZj \end{aligned}$$

The impedance spectrum can also be represented as an equivalent electric circuit model as shown in Fig. 4.

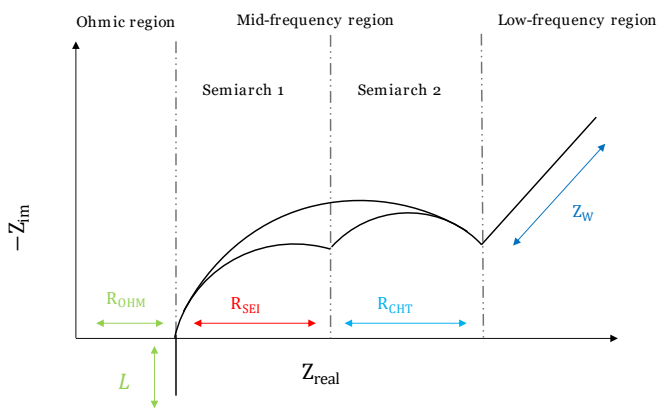


Fig. 3. Analysis of EIS in a Nyquist plot inspired by Li et al. (2020), and Iurilli et al. (2021).

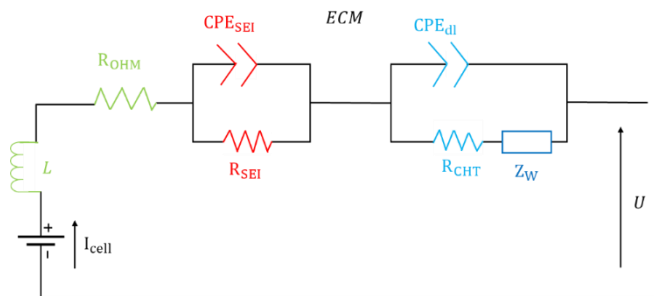


Fig. 4. Common equivalent circuit model inspired by Xiong et al. (2020).

According to Xiong et al. (2020), the model in Fig. 4 is the most common ECM, which is composed of three parts:

- Part 1: a series of  $R_{OHM}$  and  $L$ .
- Part 2: a parallel of  $CPE_{SEI}$  and  $R_{SEI}$ .
- Part 3: a parallel of  $CPE_{dl}$  and series of  $R_{CHT}$  and  $Z_W$

Part 1 indicates the ohmic resistance increase, where  $R_{OHM}$  incorporates the ohmic resistance of electrolytes, electrode, binder and current collector. It can be acquired by resolving the intersection among the impedance spectrum and high frequency region of the Nyquist plot. The inductance incorporated the high frequency phenomena occurring in the collector, can be acquired by the impedance positive imaginary part. Part 2 describes the formation, decomposition and growth of SEI film, where the  $R_{SEI}$  is calculated from the first semi arch span at mid-frequency. Part 3,  $R_{CHT}$  the charge transfer resistance attained by a second semiarch at low frequency,  $CPE_{dl}$  simulates the double-layer affect that occurs during battery discharge for the shape of electrode according to Xiong et al. (2020).

In our cycling tests we have collected spectra with  $dQ/dV$  as a function of voltage and number of cycles. Here we can see how the spectrum for the same battery is changing pattern. In this case we cycled NMC batteries model 18650 up to four hundred cycles from 3.3 to 4.2 Volt. The results are shown in Fig. 5.

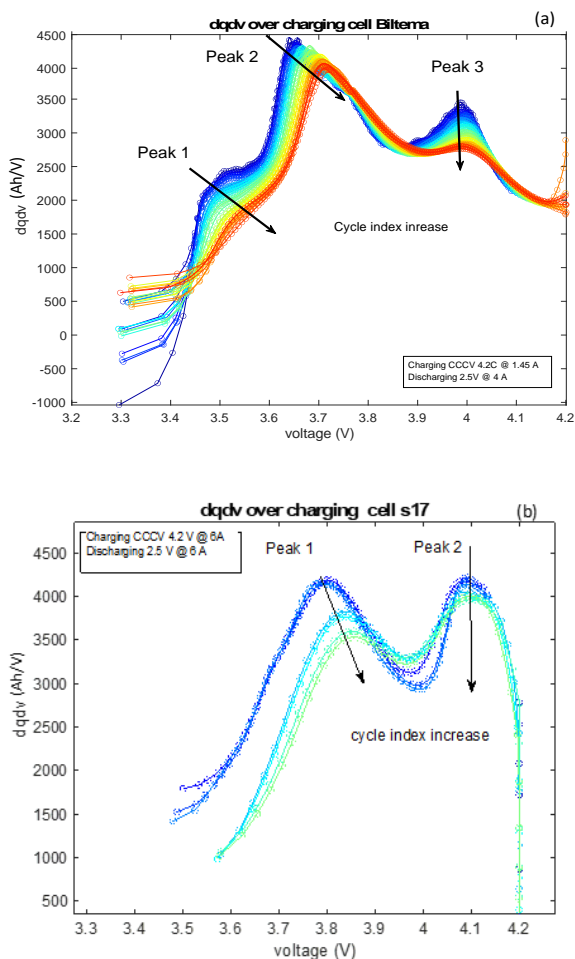


Fig. 5. Results of the incremental capacity analysis over the cell's cycling (a) NMC cell from Biltema and (b) NMC cell from Samsung.

The shift in pattern during cycling is shown as a few arrows. This information can be used to predict the performance of the batteries. It can also be used for prediction of remaining useful

life, RUL, if cycling proceeds until the capacity has faded to below 70-80% of the original SoH.

In Fig. 6, we see how the EIS changes during cycling. The diagram to the left is for fresh batteries while the others are after several cycles. The higher the cycle number, the further to the right.

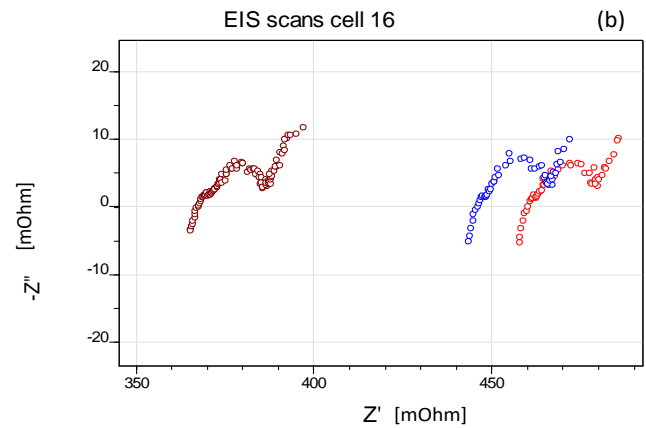


Fig. 6. Electrochemical impedance spectra (EIS) for a NMC battery from Samsung after Nyquist plot for battery cell (marked 16): cycle 200, 207 and 300. Experiments performed at MDU.

From experimental data, regressions and prediction models have been developed using different AI-algorithms. Results from these are exemplified in Fig. 7.

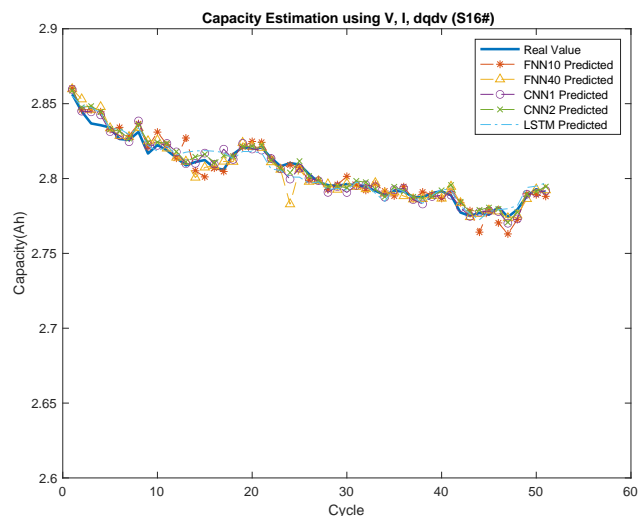


Fig. 7. Battery capacity estimation using different deep-learning algorithms as a function of cycle number (FNN10,FNN40,CNN1,CNN2 and LSTM). Rojas Vazquez (2023).

Some other approaches using different type of models is e.g. capacity degradation estimation using sigma point Kalman filter (Gaddipati and Kuthuri, 2023) and Lin et al (2023) using a data driven approach.

## 2.2 Testing of battery packs

For many cells in series and parallel, it is difficult to perform EIS, while  $dQ/dV$  or  $dV/dQ$  is possible to measure. By comparing this for packs as well as single cells, and



performing EIS as well for the single cells, we can create data that can be used to make AI models for different types of performance numbers. These can be state of charge (SoC), state of health (SoH), remaining useful life (RUL), and similar. By measuring full cycles for batteries and packs when fresh and after different numbers of cycles, prediction models can be developed. Also the depth of discharge (DoD), that is how close to 0% charge you go, is of interest to monitor. C-rate for both charge and discharge as well as temperature in the cell packs are other factors.

### 3. EXPERIMENTAL DATA

#### 3.1 Our experimental data

Previously we presented data from NMC battery cycling performed at our lab. We also have done testing with LFP batteries, or more correct Lithium Iron Manganese Phosphate batteries with 20Ah capacity. The latter are prismatic, while the first ones are cylindrical with a 2.9Ah capacity.

As seen above, the degradation during fifty cycles of a NMC battery was 1.8% or 0.035% per cycle, when cycling between 2.5V and 4.2V. Another cell was cycled 325 cycles between 2.5V and 4.2V with 1.5A charging current (C-rate= 0.52) and 3A discharge current (C-rate =1.04). The decrease in capacity went from 2.91Ah to 2.73Ah, or 0.18Ah. This means 6.2% during 325 cycles or a decrease of 0.019% per cycle. If we just look at the first fifty cycles of the same battery (S20), the decrease was from 2.91Ah to 2.81Ah, or 3.4%, or 0.069% per cycle. There is a faster degradation in the beginning, but it is reduced with time.

For the LFP batteries, we had a degradation of 0.12% per cycle at C-rate 0.15 during thirty-nine cycles, when cycling between 2.5V and 3.7V at 3A for the 20Ah batteries. It was the same for charging and discharging.

We did the same cycle for LFP batteries but with four cells in series with 10A and 40A respectively. This corresponds to a C-rate of 0.5 and 2 respectively. For the reference case with C-rate 0.5 and cycling only between 12.6V to 13.7V for four batteries, corresponding to 3.15-3.4V per cell. We could not see any degradation at all after one hundred cycles with this mild cycling. For C-rate 2, 40A, we saw a small degradation by 0.0079% per cycle during 102 cycles. When running at C-rate 0.5, the temperature increase was around 2-3°C but for C-rate two it was 55-65°C measuring between the cells at various positions. The lower temperature was at the entrance and outlet from the series, while the higher temperature was between the cells.

A problem with these measurements is that we get slightly non-linear degradation. The degradation is higher in the first fifty cycles compared to later. This also will depend on temperature, C-rate, and other factors. To get more insights, we have collected data from the literature as explained in the next subsection.

#### 3.2 Other experimental data

Tests with different Tesla cars with NAC batteries were made with fast charging 90% of the time (Tesla 3 and Tesla Y) and

compared to Tesla models with only 10% fast charging. This was followed 1000- 2000 days. What they found was that the degradation was very similar in both cases (SOH from 99% to 91% for 1000 days and 89% for 2000 days. The % is the percent of the SOH measured as original distance with fully charged batteries after use compared to fresh batteries). Still, here the temperature control has been good, and a charging pattern with low power close to full charge (around 80-90%). Still, it is not known how degradation affects the long-term capacity like 10-20 years.

Later the Tesla user’s organization compared LFP batteries. Salazar and Bengoechea (2021) have summarized information reported by different Tesla Model 3 owners, who have LFP batteries. One had a decrease of capacity by 4.8% during 136,000 km operation; another 2.3% loss during 22,000 km, when the vehicle was charged to 100% five days a week. A third had 2.2% decrease in SOH during 32,000 km and 12 months of operation. In this case, all the cars had LFP batteries. The second Tesla owner says he was charging to 100% five days a week, which is higher DOC than recommended. Also, the others said they often charged to 100%. A Tesla model 3 consumes 1.4-1.7kWh/10km which means 136,000 km corresponds to some 19,000 – 23,100kWh total charging. The battery capacity is some 55-77kWh depending on the model, so it corresponds to 250- 420 full battery cycles. 4.8% total degradation then corresponds to 0.011-0.019% per cycle. The second had a loss of 2.3% during 2,000 km or 3080–3740kWh. This means forty -68 full cycles, or 0.058- 0.034% per cycle.

Zhang et al. (2023) studied the degradation of both LFP and NMC batteries as a function of temperature and C-rate. The LFP batteries were charged to 3.65V while the NMC batteries were charged to 4.2V. The degradation is seen in Table 1.

**Table 1. Degradation of LFP and NMC batteries as a function of C-rate and Temperature according to Zhang et al. (2023)**

C-rate	% degradation per cycle					
	0.2		0.5		1	
Temp	25°C	-10°C	25°C	-10°C	25°C	-10°C
NMC	0.02		0.07		0.05	
NMC			0.0375	1.16	0.4125	3.6
LFP	0.03		0.25		0.36	
LFP			0.0233	0.26		

Lin et al. (2023) have studied SoH in relation to internal resistance. They found a degradation of SoH by 8% during three hundred cycles while the inner resistance increased from 0.18 to 0.213 Ohm. The SoH decrease per cycle was 0.027%.

Shabani et al. (2023) have shown that DOD/DOC has an impact on degradation, but also where in the span charge and discharge occur. With the same total cycle depth but with different spans you see different degradation. With DOC =50%, we can see that the degradation rate goes from 12 to



14.5% for 10 years when having a cycle of 40-90% SoC compared to 10-60% SoC. When we increase the DOC above this span of 50%, we also see an increased fade. How much this depends on temperature, battery type, and C-rate?

### 3.3 Summary of degradation data

If we try to summarize the data, both our own and others' data, we get a high variation as % per cycle, but still, we can see some trends. An increased C-rate above 0.5-1 usually increases the fade of SoH. When the temperature is below 0-10°C and above 30°C, we also see an increased degradation rate. The difference between LFP and NMC batteries is not clear from these data. What can be seen is that the value for the same condition varies significantly at the same temperature and C-rate. We can see that the C-rate above 0.2 is increasing the degradation rate as well as -10°C compared to +25°C. This is for both LFP and NMC batteries. It is usually said that LFP should not get that hot as energy per kg is lower, but we saw a very high increase to 65°C at 40A with four cells in series, with each 20Ah, or C-rate 2. This led to the swelling of the batteries significantly. Concerning DOC/DOD many authors report that this is important, like Shabani et al. (2023), but in absolute numbers, it is not that easy to get a reliable figure.

What we have done with our simulation model is to set some average values on degradation rate and from these estimate parameter values. Adjustment is made for large changes in DOC, temperature, and C-rate. We have made these adjustments for each cycle assuming a full cycle. When the cycle is not full, we assume degradation in SOH is a share of the full cycle.

## 4. BATTERY DEGRADATION MODEL

The battery degradation will depend on several factors like time, temperature, Depth of Discharge (DOD), Depth of Charge (DoC), number of cycles, and C-rate as well as the calendric time as such. It is of interest to define some key numbers to follow that integrate these different factors.

The algorithm we use for the battery simulation is shown in Table 2 (input data), Table 3 (calculations) and Table 4 (calculations for first 21 timesteps during charging) below.

Table 2. Input data to battery simulator. In this case a single LFP cell.

Input	LFP Cell	Sort
$E_{\max,cell}$	64	Wh
$E_{0,cell}$	3.2	Wh
$E_{100,cell}$	60.8	Wh
$U_{\max,cell}$	3.7	V
$U_{\text{low,cutoff}}$	2.5	V
$U_{0,cell}$	2.5	V
$U_{100,cell}$	3.515	V
$I_{0,cell}$	15	A
$I_{100,cell}$	2	A
$U_{\text{normal}}$	3.2	V
$P_{\text{cell}}$	48	W

The calculations in the simulator are chosen as constant voltage, constant current or constant power. In Table 3 below we see calculations for the constant power case during charging. For discharge the calculation of SoC is slightly different compared to during charging.

Table 3. Calculation for constant power (kW) during charge and discharge

Constant power		
dt	0.016667	h (minute)
$E_{t,cell}$	60.8	$E_{100,cell}$
$P_{\text{cell}}$	48	
$U_{t,cell}$	3.515	$U_{100,cell}$
$I_{t,cell}$	13.65576	$P_{\text{cell}}/U_{100,cell}$
$E_{t+1,cell}$	60	$E_{t,cell} - U_{t,cell} * I_{t,cell} * dt$
SoC	0.986111	$(E_{t+1,cell} - E_{0,cell}) / (E_{100,cell} - E_{0,cell})$
$U_{t+1,cell}$	3.500903	$U_{t,cell} - (1 - \text{SoC}) * (U_{100,cell} - U_{0,cell})$
$I_{t+1,cell}$	13.71075	$P_{\text{cell}} / U_{t+1,cell}$
Loop	$E_{t,cell} = E_{t+1,cell}$	
	$U_t = U_{t+1}$	
	$(I_t = I_{t+1})$	

Principally we calculate an update of SoC for each time step depending on the kW discharge or charge. Calculations are done for a single cell, but several cells in parallel and series are configured to get the correct current, voltage, and energy content (kWh).

Table 4. Calculation of first 10 time-steps during charge of 15 cells in series

Constant power 15 cells										
t	1	2	3	4	5	6	7	8	9	10
dt	0.01667	0.01667	0.01667	0.01667	0.01667	0.01667	0.01667	0.01667	0.01667	0.01667
E <sub>0</sub>	48	48	48	48	48	48	48	48	48	48
E <sub>100</sub>	912	912	912	912	912	912	912	912	912	912
U <sub>0</sub>	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
U <sub>100</sub>	52.275	52.275	52.275	52.275	52.275	52.275	52.275	52.275	52.275	52.275
P	720	720	720	720	720	720	720	720	720	720
I <sub>0</sub>	15	15	15	15	15	15	15	15	15	15
I <sub>t</sub>	19.2	19.09548	18.9921	18.88983	18.78865	18.68856	18.58952	18.49153	18.39457	18.29861
E <sub>t</sub>	48	60.0024	72.0048	84.0072	96.0096	108.012	120.0144	132.0168	144.0192	156.0216
U <sub>t</sub>	37.5	37.70525	37.9105	38.11575	38.321	38.52625	38.7315	38.93675	39.142	39.34724
E <sub>t+1</sub>	60.0024	72.0048	84.0072	96.0096	108.012	120.0144	132.0168	144.0192	156.0216	168.024
SoC	0.013892	0.027783	0.041675	0.055567	0.069458	0.08335	0.097242	0.111133	0.125025	0.138917
U <sub>t+1</sub>	37.70525	37.9105	38.11575	38.321	38.52625	38.7315	38.93675	39.142	39.34724	39.55249
I <sub>t+1</sub>	19.09548	18.9921	18.88983	18.78865	18.68856	18.58952	18.49153	18.39457	18.29861	18.20366

If we look at the degradation due to different factors, we can see that cycling conditions can be accounted for with a number of adjustment factors or KPIs:

1. DoD/DoC is calculated as  $SoC_{in}$  when a cycle starts minus  $SoC_{out}$  when we switch from charging to discharging or the opposite. When DoD/DoC is larger than 60% the amount between actual value and 60% is calculated and multiplied by  $KPI_{doc}$ .
2. The temperature is assumed normal between 10°C and 30°C, but increased degradation in proportion to temperature difference higher or lower than this.  
 $KPI_{temp} = T_{operational} - (> 30^{\circ}C \text{ or } < 10^{\circ}C) * C_{temp}$ .
3. Adjustment for C-rate is  $KPI_{c-rate} = C\text{-rate}^{\wedge C\text{-rate}}$ .
4. Calendric time influence  $KPI_{cal} = \text{number of hours since manufacture of battery} * C_{cal}$ .

Degradation now becomes average degradation when DOD/DOC is <60%, temperature 10-30°C and C-rate < 0.5. We then add degradation rate as add-ons to this average value.

Degradation of SoH equals to:

$$SoH_{deg} = SoH_{average} + (SOC-60)*KPI_{doc} + (Temp - >30 \text{ or } <10)*KPI_{temp} + C\text{-rate}*KPI_{c-rate}$$

The  $SoH_{average}$  is calculated from the measured values when conditions are stated as above. For fifty cycles we have eight test sets giving an average of 0.045%/cycle. For 325 and 435 cycles we have used two data sets, giving 0.015%/cycle. These are the base values under “normal conditions”. So  $SoH_{average,50} = 0.045\%$  and  $SoH_{average,400} = 0.015\%$ . We assume the same for both NMC and LFP batteries.

When the temperature went down to -10°C, the degradation rate for NMC batteries was 1.16%/cycle at C-rate 0.5 and 3.6%/cycle at C-rate one. For C-rate going from 0.5 to 2 the degradation rate went from 0.001 to 0.0079 for LFP battery

and from 0.038 to 0.41 for a NMC battery in one set but from 0.07 to 0.05 in another! The tests unfortunately give quite diverse measures! At extreme temperatures we normally see significant degradation of SOH, but sometimes less than expected. We thus have chosen to use conservative values. The plan is to use future measurements to make these factors better by time, including both measurements done in lab as well as including module and pack data from different vehicles. The following values have been set as our initial estimates:  $KPI_{doc} = 0.002$ ,  $KPI_{temp} = 0.005$ ,  $KPI_{c-rate} = 0.01$ . This would give for DOD=90, temperature 0°C and C-rate 2 a  $SoH_{deg,400} = 0.015 + (90-60)*0.002 + 10*0.005 + 2*0.01 = 0.145\%$ .

Battery degradation could be modelled as  $U=I*R$  where R is increased as a function of degradation of SoH. The correlation between inner resistance and SoH is that a decrease of R by 0.18 to 0.213Ohm correspond to a decrease in SoH by 8% during three hundred cycles according to Wang et al. (2023). This means 0.027% per cycle. This is a reasonable value if we assume 25°C and C-rate 0.2-0.5. We also assume DOC/DOD to be 60% (SoC 20-80%).

The actual power then could be calculated as  $P_{actual} = P*(R_{original}/R_{present})$ . When we demand  $P = 980W$ , we only get  $P*(R_{original}/R_{present})$  which is lower than demanded, assuming that total resistance is increasing.

In our simulation model though we are using the SoH degradation depending on temperature, DOC/DOD and C-rate as stated previously. From this we can calculate battery degradation giving RUL and SoH from running with different scenarios with respect to the different conditions.

## 5. POWER DEMAND FROM VEHICLE

The power demand for each time step is given from a model over e.g. a train line going from one station to the next. There is a time schedule that must be followed given the limitations with respect to acceleration, deceleration, and average

velocity. When passing sensitive areas there are speed limitations e.g. 80 km/h to reduce the impact of noise. The power demand for a train line is seen in Fig. 8 below. The power demand is calculated from the weight of the train, distance, inclination, velocity limitations, acceleration, friction due to bearings, and wind. The results are shown in Fig. 9. The power demand is in kW. Positive values are discharging batteries and negative values charging the batteries due to “motor breaking”.

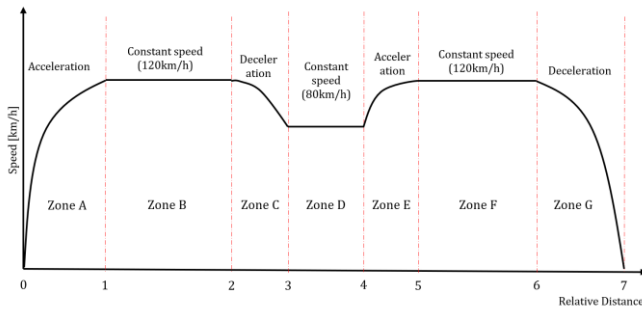


Fig. 8. A train line with seven zones with different conditions with respect to velocity, inclination and acceleration/deceleration.

From the data in Fig. 8 power demand for each zone is calculated and sent to the battery simulator. In Fig. 9 we see the speed of the train in each zone, or more precisely the velocity of the train when it enters and leaves the zone. Thereafter we see the power demand as kW in each zone, the energy output or input to the battery as kWh and finally the state of charge, SoC, calculated for each zone. In this case there were just seven zones, but where there are major accelerations/decelerations each zone may be divided into several zones.

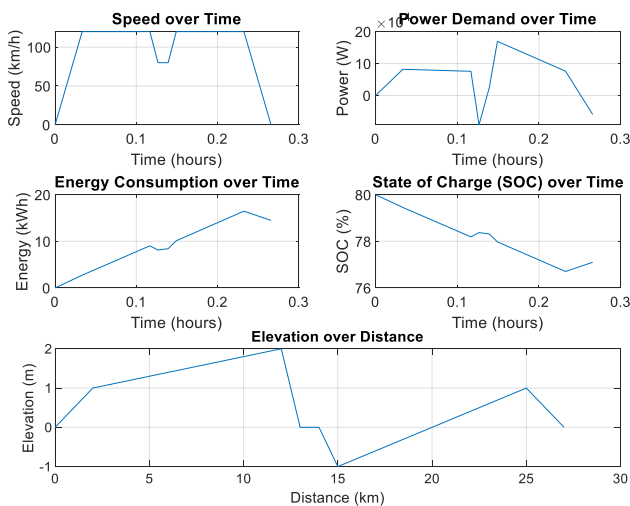


Fig. 9. Speed in km/h in each zone (top), power demand (kW), energy transfer (kWh) and finally state of charge (SoC) in/out of each zone.

For each full cycle, we do a calculation on degradation of SoH due to calendric time, temperature, DOD/DOC and C-rate. For part of a full cycle, that is how much of a full cycle before changing from charging to discharge or vice versa, we do this

calculation as percent of a full cycle, so as a function of total Ah stored and used.

## 6. DISCUSSION AND CONCLUSIONS

From the experimental results, we have made prediction models for RUL and SoH. Concerning the battery simulator, we have formulated the equations for constant power during each section of the distance of a vehicle, or during a certain time period. By running scenarios like a train line as above or some other driving cycle for another type of vehicle, we can simulate future degradation and from this calculate RUL and SoH at certain times, or for “end of life” (EOL). Also calculations can be made on when SoH has reached e.g. 80%, where a second life use would be recommended. From the intensity of the drive cycle, we also can recommend suitable use of the batteries for this second life use. In case of harsh cycles, it may be better to use the batteries for only energy storage like in photovoltaic (PV) cell applications. If low DOD/DOC has been applied generally, a power application, like for example frequency control, can be possible, where high power may be demanded.

## ACKNOWLEDGEMENTS

We thank our partner companies, especially Alstom, in the project SITE1, financed by the Swedish Energy Agency aside from Alstom in the grant 2020-021322 (project number 49896-2). We also thank the Knowledge Foundation (KKS) for the support through the SYNERGY project Circul8 (Smart Circular Battery System) with number 20230024.

## REFERENCES

- Ahmadian, A., Sedghi, M., Elkamel, A., Fowler, M., and Golkar, M.A. (2018). Plug-in electric vehicle batteries degradation modeling for smart grid studies: Review, assessment and conceptual framework. *Renewable and Sustainable Energy Reviews*, 81, 2609–2624.
- Chirumalla, K., Kulkov, I., Parida, V., Dahlquist, E., Johansson, G., Stefan, I. (2024). Enabling battery circularity: Unlocking circular business model archetypes and collaboration forms in the electric vehicle battery ecosystem. *Technological Forecasting and Social Change*, 199, 123044.
- Chirumalla, K., Kulkov, I., Vu, Felix., Rahic, M. (2023). Second life use of Li-ion batteries in the heavy-duty vehicle industry: Feasibilities of remanufacturing, repurposing, and reusing approaches. *Sustainable Production and consumption*, 42, 351-366.
- Dahlquist, E., Wallin, F., Chirumalla, K., Toorajipour, R., Johansson, G. (2023). Balancing power in Sweden using different renewable resources, varying prices and storages like batteries in a resilient energy system. *Energies*, 16(12), 4734
- Dai, H., Zhao, G., Lin, M., Wu, J., and Zheng, G. (2018). A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and markov chain. *IEEE transactions on industrial electronics*, 66(10), 7706–7716.
- de la Vega, J., Riba, J.R., and Ortega-Redondo, J.A. (2023). Mathematical modeling of battery degradation based on direct measurements and signal processing methods. *Applied Sciences*, 13(8), 4938.

- Gaddipati, G.R. and Kuthuri, N.R. (2023). Enhanced lithium-ion battery model for estimation of degraded capacity and soc using sigma point Kalman filter. *International Journal of Renewable Energy Research (IJRER)*, 13(2), 857–870.
- Iurilli, P., Brivio, C., and Wood, V. (2021). On the use of electrochemical impedance spectroscopy to characterize and model the aging phenomena of lithium-ion batteries: a critical review. *Journal of Power Sources*, 505, 229860.
- Lam, L. and Bauer, P. (2012). Practical capacity fading model for li-ion battery cells in electric vehicles. *IEEE transactions on power electronics*, 28(12), 5910–5918.
- Li, Z., Liu, D., Xiong, J., He, L., Zhao, Z., and Wang, D. (2020). Selective recovery of lithium and iron phosphate/carbon from spent lithium iron phosphate cathode material by anionic membrane slurry electrolysis. *Waste management*, 107, 1–8.
- Lin, M., Yan, C., Wang, W., Dong, G., Meng, J., and Wu, J. (2023). A data-driven approach for estimating state-of-health of lithium-ion batteries considering internal resistance. *Energy*, 277, 127675.
- Maheshwari, A., Paterakis, N.G., Santarelli, M., and Gibescu, M. (2020). Optimizing the operation of energy storage using a non-linear lithium-ion battery degradation model. *Applied Energy*, 261, 114360.
- O’Kane, S.E., Ai, W., Madabattula, G., Alonso-Alvarez, D., Timms, R., Sulzer, V., Edge, J.S., Wu, B., Offer, G.J., and Marinescu, M. (2022). Lithium-ion battery degradation: how to model it. *Physical Chemistry Chemical Physics*, 24(13), 7909–7922.
- Panchal, S., Mcgrory, J., Kong, J., Fraser, R., Fowler, M., Dincer, I., and Agelin-Chaab, M. (2017). Cycling degradation testing and analysis of a lifepo4 battery at actual conditions. *International Journal of Energy Research*, 41(15), 2565–2575.
- Pelletier, S., Jabali, O., Laporte, G., and Veneroni, M. (2017). Battery degradation and behaviour for electric vehicles: Review and numerical analyses of several models. *Transportation Research Part B: Methodological*, 103, 158–187.
- Qadrdan, M., Jenkins, N., and Wu, J. (2018). Smart grid and energy storage. In *McEvoy’s Handbook of Photovoltaics*, 915–928. Elsevier.
- Rahbari, O., Mayet, C., Omar, N., and Van Mierlo, J. (2018). Battery aging prediction using input-time- delayed based on an adaptive neuro-fuzzy inference system and a group method of data handling techniques. *Applied Sciences*, 8(8), 1301.
- G. Geetha Ravali\*, K. Narasimha Raju: Enhanced Lithium-Ion Battery Model for estimation of Degraded Capacity and SoC Using Sigma Point Kalman Filter. *International journal of renewable energy research*, Vol.13, No.2, June, 2023
- Rojas Vazquez Josefin: Battery capacity prediction using deep learning. Master thesis work, Malardalen University press, 2023.
- Salazar, E. and Bengoechea, J. (2021). Research paper lithium and its role in the new energy transition. *Bachelor Thesis, Geneva Business School*.
- Shabani, M., Wallin, F., Dahlquist, E., and Yan, J. (2023). The impact of battery operating management strategies on life cycle cost assessment in real power market for a grid-connected residential battery application. *Energy*, 270, 126829.
- Shamarova, N., Suslov, K., Ilyushin, P., and Shushpanov, I. (2022). Review of battery energy storage systems modeling in microgrids with renewables considering battery degradation. *Energies*, 15(19), 6967.
- Shen, D., Xu, T., Wu, L., and Guan, Y. (2019). Research on degradation modeling and life prediction method of lithium-ion battery in dynamic environment. *IEEE Access*, 7, 130638–130649.
- Shi, E., Xia, F., Peng, D., Li, L., Wang, X., and Yu, B. (2019). State-of-health estimation for lithium battery in electric vehicles based on improved unscented particle filter. *Journal of Renewable and Sustainable Energy*, 11(2).
- Uddin, K., Perera, S., Widanage, W.D., Somerville, L., and Marco, J. (2016). Characterising lithium-ion battery degradation through the identification and tracking of electrochemical battery model parameters. *Batteries*, 2(2), 13.
- Wang, Z., Zhao, X., Fu, L., Zhen, D., Gu, F., and Ball, A.D. (2023). A review on rapid state of health estimation of lithium-ion batteries in electric vehicles. *Sustainable Energy Technologies and Assessments*, 60, 103457.
- Wildfeuer, L., Karger, A., Aygül, D., Wassiliadis, N., ossen, A., and Lienkamp, M. (2023). Experimental degradation study of a commercial lithium-ion battery. *Journal of Power Sources*, 560, 232498.
- Xiong, R., Huang, J., Duan, Y., and Shen, W. (2022). Enhanced lithium-ion battery model considering critical surface charge behavior. *Applied Energy*, 314, 118915.
- Zhang, X., Sun, P., Wang, S., and Zhu, Y. (2023). Experimental study of the degradation characteristics of lifepo4 and lini0. 5co0. 2mn0. 3o2 batteries during over-charging at low temperatures. *Energies*, 16(6), 2786.

# Numerical simulation of thermal runaway kinetic mechanisms and battery thermal model for safety assessment of different lithium-ion battery chemistries

S. Mehranfar\*, A. Mahmoudzadeh Andwari\*, J. Könnö \*, A. Garcia Martinez \*\*, C. Mico Reche\*\*

\* *Machine and Vehicle Design (MVD), Materials and Mechanical Engineering, Faculty of Technology, University of Oulu, FI-90014 Oulu, Finland (Sadegh Mehranfar Tel: +358 50 4321302; e-mail: [Sadegh.Mehranfar@oulu.fi](mailto:Sadegh.Mehranfar@oulu.fi), Amin Mahmoudzadeh Andwari [Amin.M.Andwari@oulu.fi](mailto:Amin.M.Andwari@oulu.fi))*

\*\* *CMT–Clean Mobility and Thermo fluids. Universitat Politècnica de València. Camino de Vera s/n, 46022 Valencia, Spain*

**Abstract:** The importance of EVs and li-ion batteries are pinpointed in the automotive industry during the last decade by increased growth of electrified powertrain. Li-ion batteries offer significant improvements in terms of energy and power density; however, safety challenges continue to exist. Different thermal, mechanical, or electrical abuse conditions in li-ion batteries can trigger a series of exothermic chain reactions in the battery cells and finally result in thermal runaway (TR) and combustion of battery cells and EVs. Furthermore, different battery technologies exploit various cell chemistries, leading to the distinct thermal behavior of battery cells during normal and abuse conditions. This work aims at investigating the TR kinetic mechanisms to evaluate thermal behavior of the battery cells under thermal abuse conditions. Furthermore, this work investigates the different li-ion battery cathode, anode and electrolyte materials to assess the safety of battery systems in EV application. The results revealed that unlike batteries with LiCoO<sub>2</sub> cathodes with temperature threshold of 150 °C, Li<sub>1.1</sub>(Ni<sub>1/3</sub>Co<sub>1/3</sub>Mn<sub>1/3</sub>)<sub>0.9</sub>O<sub>2</sub> batteries do not undergo TR process at temperatures below 170 °C. Moreover, the temperature peak is more hazardous in LiCoO<sub>2</sub> batteries with LiPF<sub>6</sub>/PC: DMC electrolyte compared to the same battery with standard electrolyte. In addition, batteries with Lithiated Li<sub>4</sub>Ti<sub>5</sub>O<sub>12</sub> anode showed safer TR process compared to all the previous battery types.

**Keywords:** Numerical simulation, Electric Vehicles, Li-ion batteries, Safety, Thermal runaway kinetic mechanisms

## 1. INTRODUCTION

The importance of EVs and li-ion batteries are pinpointed in the automotive section by the zero-tailpipe emission requirement of EU fleet and increased share of electrified powertrain in the market (IEA). Li-ion batteries offer significant improvements from the first generations of EVs in terms of energy density and power density, however, safety challenges in the way of li-ion EVs continue to exist (Wang et al., 2023). Extensive research has been done in thermal management of EV batteries by proposing hybrid cooling methods or even battery operating under cold climate to enhance the performance of battery system (Gharehghani et al., 2022, 2023). However, battery performance under abuse conditions remains a challenge for battery developers. Different thermal, mechanical, or electrical abuse conditions in li-ion batteries can trigger a series of exothermic chain reactions in the battery cells and finally result in thermal runaway (TR) and combustion of battery cell, battery system and the EV. Therefore, investigations of thermal behavior of battery cells under critical conditions are of utmost importance for EVs security and driver's safety. Moreover, different battery technologies exploit various cell chemistries (Cathode, anode, electrolyte and separator materials), leading to their distinct thermal behavior. Some researchers have

conducted investigations on the thermal behavior of Li-ion batteries under TR. A lumped model is proposed by Hatchard et al. (Hatchard et al., 2001) to model the oven test as a standard procedure of battery TR under thermal abuse conditions. Kim et al. extended the previous models to 3D models for oven tests of cylindrical cells and showed that smaller cylindrical cells can reject heat faster than larger cells and undergo a more moderate TR. Different studies have focused on the thermal stability of cathode materials. MacNiel et al. (MacNiel et al., 2002) studied the thermal stability of seven different cathode materials by differential scanning materials (DSC) and ranked them from safest to the least safe. Jiang et al. studied the three different cathode materials, LiCoO<sub>2</sub>, Li(Ni<sub>0.1</sub>Co<sub>0.8</sub>Mn<sub>0.1</sub>)O<sub>2</sub> and LiFePO<sub>4</sub> using accelerated rate calorimetry (ARC) and showed that LiFePO<sub>4</sub> offers highest thermal stability. Peng et al. (Peng et al., 2016) numerically investigated thermal safety of batteries for five different cathode materials. Wang et al. (Wang et al., 2006) investigated the thermal stability of li-ion battery electrolytes and fitted the chemical reaction kinetics by Arrhenius law. They concluded that the stability of electrolyte plays a substantial role in li-ion safety. The effect of anode material was also investigated by Haung et al. (Haung et al., 2016). The thermal and combustion characteristics of TR over the battery module with Li<sub>4</sub>Ti<sub>5</sub>O<sub>12</sub> anode battery cells were investigated



through heating. Numerical investigation of thermal runaway behavior of lithium-ion batteries with different battery materials is also done by other researchers in the field (Kong et al., 2021). In spite of numerous research on the effect of battery material on the thermal stability of li-ion cells, a comprehensive study to investigate the effect of different cell components TR process of the battery cells is rarely done.

This work aims at the investigation of the most well-proven kinetic mechanism reactions of TR phenomena to evaluate thermal behavior of the battery cells with different materials under thermal abuse conditions. A thermal model with TR kinetic mechanism sub-model is adopted to replicate the behavior of the battery cells under thermal abuse conditions. Furthermore, to investigate the effect of different battery materials of the cell thermal behavior during TR, two different cathode, anode and electrolyte is selected from the literature and also their temperature evolution with heat rates are compared in the simulation of thermal abuse test. The results of this work will facilitate the integration of kinetic mechanisms into battery modelling under critical operation and will improve the safety design of li-ion batteries in EV application.

## 2. METHODOLOGY

This work employs a 2D thermal model integrated with the TR kinetic mechanism to evaluate the thermal stability of different li-ion battery cells under thermal abuse condition. The model was built in the commercial COMSOL Multiphysics software and by setting PDEs describing the undergoing physics of the problem. The simplified schematic of battery cell is presented in Fig.1 based on the 18,650-cell geometry and the model was developed by 2D definition.

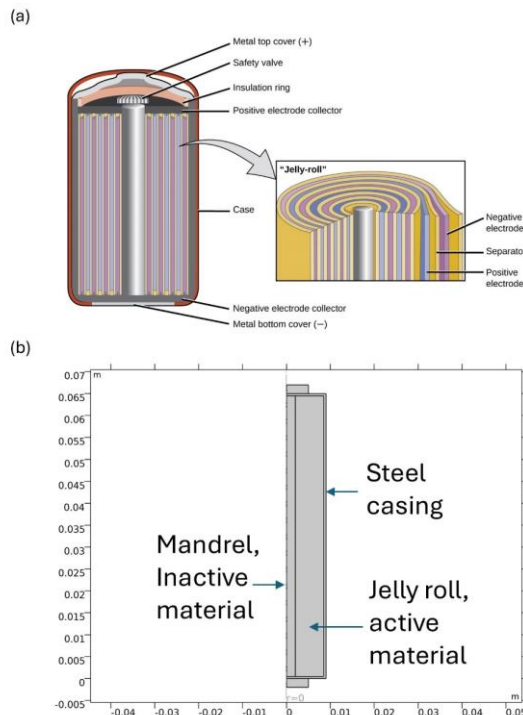


Fig. 1. A) Schematic of battery cell and inner structure B) Model Geometry of present work

The thermal model was adopted by defining the conservation of energy law and introducing the heat source term of TR reactions. Heat source term considers heat of reaction in TR event and by including a variety of different exothermic reactions in each component of battery. The objective of simulation is to replicate ARC test and trigger li-ion battery cells with thermal abuse and by setting oven temperature.

The interplay of heat transfer between cell and environment, heat of exothermic reactions and cell thermal balance determines the temperature dynamics of the cell. The temperature of the cell increases by the enthalpy of each reaction, which further increases the cell temperature and decreases the concentration of that component. The framework of numerical thermal model in the present study is illustrated in Fig. 2. Furthermore, li-ion cell properties in this simulation are listed in Table 1.

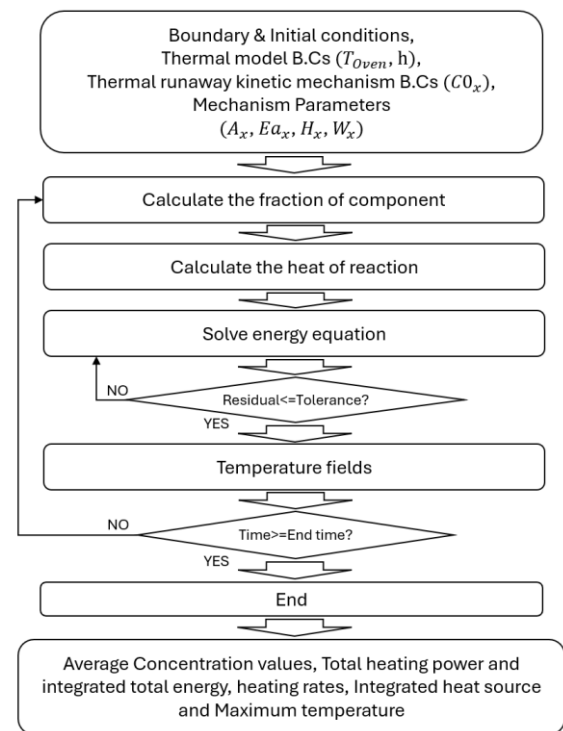


Fig. 2. Thermal modeling and TR framework

### 2.1 Thermal model

The thermal behavior inside the battery is modeled by the conduction heat transfer and conservation of energy as Eq.1 and 2. The generated heat  $Q_{gen}$  by decomposition of each component and dissipated heat  $Q_{diss}$  to the environment is then introduced into the energy equation.

$$\frac{dT_{Bat}}{dt} = \frac{Q_{gen} - Q_{diss}}{MC_p} \quad (1)$$

$$T_{Bat}(t) = T_{Bat,0} + \int \frac{dT_{Bat}}{dt} dt \quad (2)$$

The generated heat of each reaction is calculated by the TR model. Then, total generated heat is expressed as the

summation of multiple heat components following the following equation:

$$Q_{gen} = \sum Q_x \quad (3)$$

The interaction between battery and environment is determined by considering the convection and radiation heat as in Eq.4 and allows battery to reach to the environment temperature and calculate the released heat of reaction in that temperature. This iterative process is illustrated in Fig.2

$$Q_{diss} = Q_{conv} + Q_{rad} \\ = h \cdot A \cdot (T_{ARC} - T_{Bat}) \\ + \varepsilon \sigma (T_{ARC}^4 - T_{Bat}^4) \quad (4)$$

## 2.2 Thermal runaway model

The present model utilizes TR kinetic mechanisms introduced by Kim et al. (Kim et al., 2007). The model follows the basic kinetic mechanism of chemical reactions by the following Arrhenius form:

$$\kappa_x = \frac{dc_x}{dt} = A_x (c_x)^{n_1} (1 - c_x)^{n_2} e^{\frac{E_{a,x}}{R_0 T}} \quad (5)$$

Where  $\kappa_x$  is the reaction rate and  $c_x$  is the normalized concentration. Furthermore,  $A_x$ ,  $E_a$  and  $g_x$  are the pre-exponential factor, activation energy and mechanism function respectively. The concentration of each species is then updated in the TR process as follows and by calculation of the reaction rate. Model parameters for Kim et al. mechanism is presented in Table 2 and.

$$c_x = 1 - \int \kappa_x dt \quad (6)$$

The heat of the reaction is then calculated by multiplication of reaction rate, heating value ( $H_x$ ) and total mass of that component ( $m_x$ ) as the following:

$$Q_x = m_x \cdot H_x \cdot \kappa_x \quad (7)$$

Finally, the generated heat of each reaction calculated by the model is superimposed to determine the total generated heat in the following equation:

$$Q_{gen} = \sum Q_x = Q_{SEI} + Q_a + Q_c + Q_e \quad (8)$$

Where  $Q_{SEI}$  is the heat from the SEI decomposition reaction,  $Q_a$  is the heat from the anode active material and electrolyte,  $Q_c$  is the heat from the cathode active material and electrolyte and  $Q_e$  is the heat from the electrolyte decomposition. The Model parameters of different battery materials are listed in Table 1.

The results of temperature evolution simulation are compared with the experimental data in Kim et al. (Kim et al., 2007) study. The comparison in Fig. 3 shows that temperatures in the simulation are in good agreement with the experimental data.

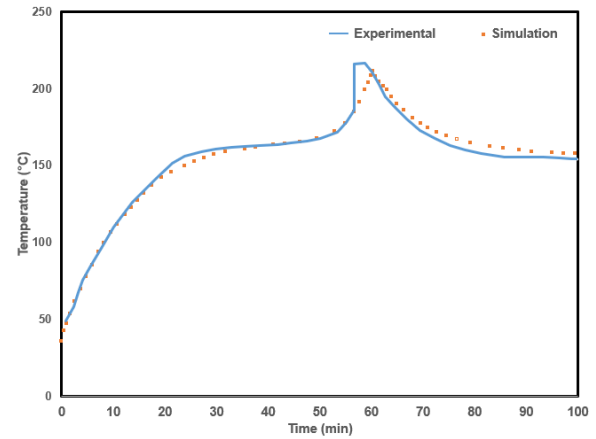


Fig. 3. Comparison of the simulated temperature and reference temperature for Kim et al. mechanism and Oven Temperature of 155 °C

In this study, a  $\text{LiCoO}_2$  battery with standard  $\text{LiPF}_6$  electrolyte and graphite anode is selected for oven test. Furthermore, a  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  cathode (Kong et al., 2021),  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte (Wang et al., 2006) and lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode (Haug et al., 2016) is selected from the literature to assess the effect of different cathode, electrolyte and anode material on the thermal stability of the cell during thermal abuse conditions respectively. The Model parameters of different battery materials are listed in Table 1.

Table 1. Model parameters of different anode, cathode and electrolyte materials.

	Lithiated $\text{Li}_4\text{Ti}_5\text{O}_{12}$ anode (Haug et al., 2016)	$\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$ cathode (Kong et al., 2021)	$\text{LiPF}_6/\text{PC}$ : DMC electrolyte (Wang et al., 2006)
$H_x$	$2.568 \times 10^5$	$7.9 \times 10^5$	$3.209 \times 10^5$
$A_x$	$5.21 \times 10^{19}$	$2.25 \times 10^{14}$	$7.53 \times 10^{19}$
$E_x$	$1.88 \times 10^5$	$1.54 \times 10^5$	$1.882 \times 10^5$
$W_x$	$1.274 \times 10^3$	$1.293 \times 10^3$	$0.96 \times 10^3$

Table 2. Model parameters and Li-ion cell properties for Kim et al. (Kim et al., 2007) mechanism.

Symbol	Description	Value
Cell format	18,650	-
Battery radius, m	$r_{batt}$	0.009
Battery height, m	$h_{batt}$	0.065
Thickness of battery can, m	$d_{can}$	$5E-4$
Mandrel radius, m	$r_{mandrel}$	0.002
Volumetric heat capacity of jellyroll, ( $\text{J m}^{-3} \text{K}^{-1}$ )	$\rho_{ho} \cdot C_{p,batt}$	$2.789E6$
Average jelly roll radial thermal conductivity, $\text{W/cm K}$	$k_{T,batt}$	0.034
Heat transfer coefficient, $\text{W}/(\text{m}^2 \cdot \text{K})$	$h_{conv}$	7.17
Reaction heat, $\text{J} \cdot \text{kg}^{-1}$	$H_{sei}$	$2.57 \times 10^5$
	$H_a$	$1.714 \times 10^6$
	$H_c$	$3.14 \times 10^5$



	$H_e$	$1.55 \times 10^5$
Reaction frequency factor, $s^{-1}$	$A_{sei}$	$1.667 \times 10^{15}$
	$A_a$	$2.5 \times 10^{13}$
	$A_c$	$6.667 \times 10^{13}$
	$A_e$	$5.14 \times 10^{25}$
Reaction activation energy, $J \cdot mol^{-1}$	$E_{a_{sei}}$	$1.3508 \times 10^5$
	$E_{a_a}$	$1.3508 \times 10^5$
	$E_{a_c}$	$1.396 \times 10^5$
	$E_{a_e}$	$2.74 \times 10^5$
Initial value, dimensionless	$C_{0,sei}$	0.15
	$C_{0,a}$	0.75
	$\alpha_0$	0.04
	$C_{0,e}$	1
Reaction order	$m_{sei}$	1
	$m_{a,n}$	1
	$m_{c,p1}$	1
	$m_{c,p2}$	1
	$m_e$	1
Volume-specific content of reacting material, $kg \cdot m^{-3}$	$t_{0,sei}$	0.033
	$W_a$	610
	$W_c$	1300
	$W_e$	406

Table 3. Initial concentrations and reaction rates for each component in Kim et al. (Kim et al., 2007) mechanism.

Component	Initial concentration	$-dc/dt$
Anode	0.75	$-R_{an}$
Cathode	0.04	$-R_{cat}$
Electrolyte	1	$-R_e$
SEI	0.15	$-R_{sei}$
$t_{sei}$	0.033	$-R_{an}$
Binder	-	-

### 3. RESULTS AND DISCUSSION

The TR simulation of the li-ion battery cells with different cathode, anode and electrolyte materials is performed to investigate the thermal stability and safety of li-ion batteries with different materials. The thermal stability of Li-ion batteries is characterized by the onset temperature and time of TR events. In addition, the thermal safety of the TR process can be characterized by the heat rate and peak temperature. The thermal safety of li-ion battery cell with LiCoO<sub>2</sub> cathode, graphite anode and standard LiPF<sub>6</sub> electrolyte is assessed based on Kim et al. kinetic mechanism. The results illustrated in Fig.4 indicate that battery cells are not prone to TR event at the temperatures under 150 °C. However, higher temperatures cause the start of exothermic reactions and further increased temperature. It can be found that higher temperatures can cause more serious hazard TR events in terms of released heat, temperature peak and onset time of TR. The results of temperature diagram, heat rate and average values for components are presented in Fig. 4.

The thermal safety of Li<sub>1.1</sub>(Ni<sub>1/3</sub>Co<sub>1/3</sub>Mn<sub>1/3</sub>)<sub>0.9</sub>O<sub>2</sub> battery cathode with standard LiPF<sub>6</sub> electrolyte and graphite anode is evaluated in Fig. 5. Unlike batteries with LiCoO<sub>2</sub> cathodes that presented temperature threshold of 150 °C, Li<sub>1.1</sub>(Ni<sub>1/3</sub>Co<sub>1/3</sub>Mn<sub>1/3</sub>)<sub>0.9</sub>O<sub>2</sub> batteries do not undergo TR process at temperatures below 170 °C. However, the thermal runaway events are much more intensive and oven temperatures of 180 can result in peak temperatures of 443 °C while LiCoO<sub>2</sub> batteries peak at ≈300 °C during TR at the same oven

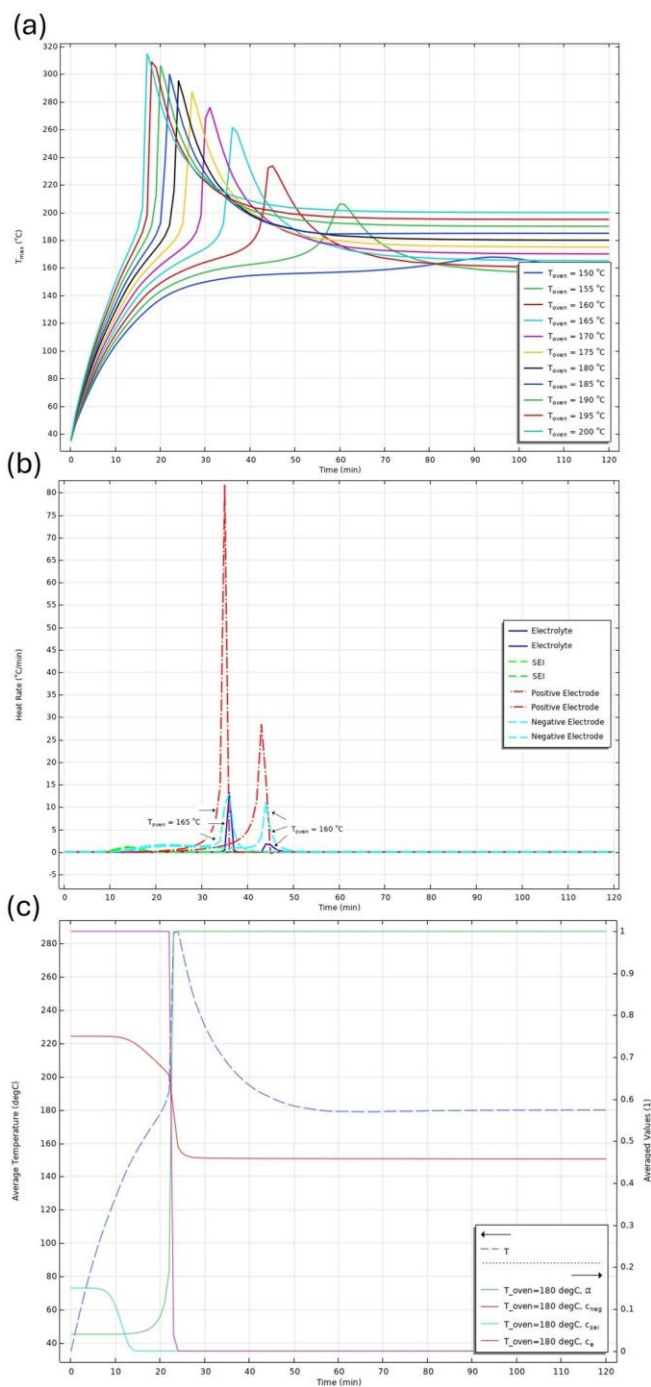


Fig. 4. TR kinetic mechanisms for LiCoO<sub>2</sub> battery with standard LiPF<sub>6</sub> electrolyte and graphite anode. A) Maximum temperature for Oven Temperature 150-200 °C B) Heating rate for Oven Temperature 160 and 165 °C C) Average values and temperature for Oven Temperature 180 °C.

The thermal safety of Li<sub>1.1</sub>(Ni<sub>1/3</sub>Co<sub>1/3</sub>Mn<sub>1/3</sub>)<sub>0.9</sub>O<sub>2</sub> battery cathode with standard LiPF<sub>6</sub> electrolyte and graphite anode is evaluated in Fig. 5. Unlike batteries with LiCoO<sub>2</sub> cathodes that presented temperature threshold of 150 °C, Li<sub>1.1</sub>(Ni<sub>1/3</sub>Co<sub>1/3</sub>Mn<sub>1/3</sub>)<sub>0.9</sub>O<sub>2</sub> batteries do not undergo TR process at temperatures below 170 °C. However, the thermal runaway events are much more intensive and oven temperatures of 180 can result in peak temperatures of 443 °C while LiCoO<sub>2</sub> batteries peak at ≈300 °C during TR at the same oven

temperature. The results of temperature diagram and heat rates for  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  batteries are presented in Fig. 5.

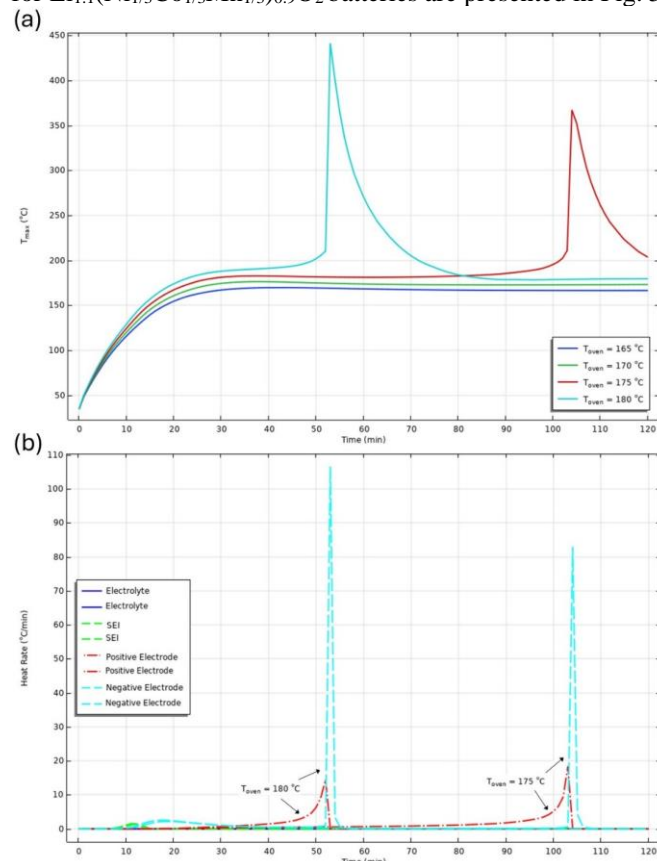


Fig. 5. TR kinetic mechanisms for  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  battery with standard  $\text{LiPF}_6$  electrolyte and graphite anode A) Maximum temperature for Oven Temperature 165-180 °C B) Heating rate for Oven Temperature 175 and 180 °C.

The thermal safety of  $\text{LiCoO}_2$  batteries with  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte and standard graphite anode was also assessed in this study to compare the influence of different electrolyte materials. The results indicate that batteries undergo TR process at 150 °C while this is a safe temperature for  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  batteries or even  $\text{LiCoO}_2$  batteries. The temperature peak is less substantial in  $\text{LiCoO}_2$  batteries with  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte compared to  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  batteries, but more intensive compared to  $\text{LiCoO}_2$  batteries, especially at higher oven temperatures. The results of temperature diagram and heat rates for  $\text{LiCoO}_2$  batteries with  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte is presented in Fig. 6.

Lastly, thermal safety of  $\text{LiCoO}_2$  batteries with standard  $\text{LiPF}_6$  electrolyte and Lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode is assessed in this study to compare the influence of different anode materials. The results indicate that batteries undergo TR process at 160 °C.  $\text{LiCoO}_2$  batteries with Lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode show less intensive TR process compared to all the previous battery types. This is evident from comparison of temperature peaks for different battery materials. The results of temperature diagram and heat rates for  $\text{LiCoO}_2$  batteries with Lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode are presented in Fig. 7.

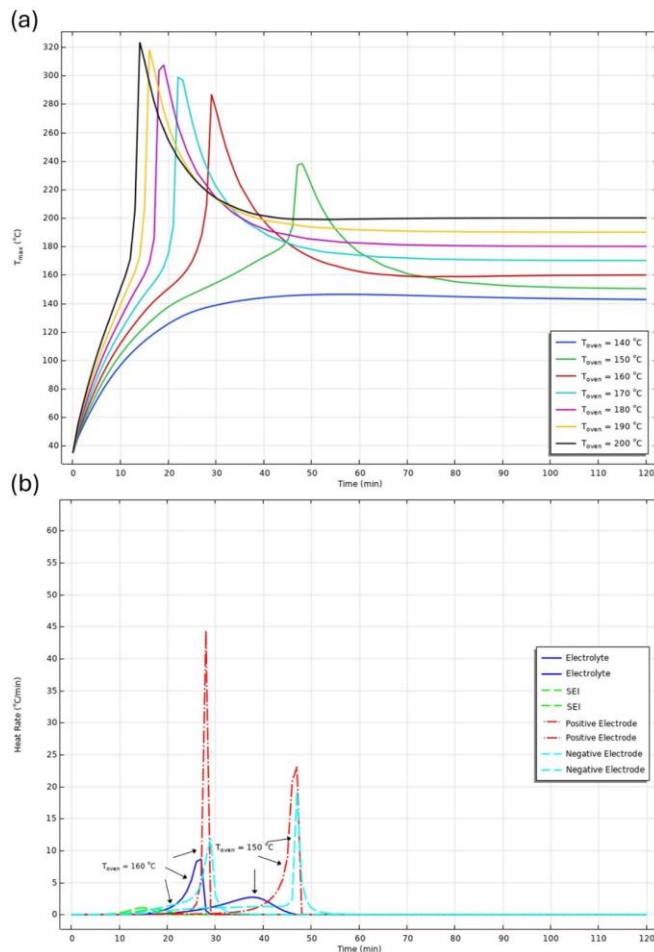


Fig. 6. TR kinetic mechanisms for  $\text{LiCoO}_2$  battery with  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte and graphite anode. A) Maximum temperature for Oven Temperature 140-200 °C B) Heating rate for Oven Temperature 150 and 160 °C.

#### 4. CONCLUSIONS

In this study thermal stability of different battery materials was evaluated for li-ion batteries under thermal abuse conditions. A  $\text{LiCoO}_2$  battery with standard  $\text{LiPF}_6$  electrolyte and graphite anode is selected for oven test as the basic battery. Furthermore, a  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  cathode,  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte and lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode is selected to assess the effect of different cathode, electrolyte and anode material on the thermal stability of the cell during thermal abuse conditions respectively. The results of temperature evolution and heat rate diagram are reported and comparison between different battery materials has been drawn.

It is shown that unlike batteries with  $\text{LiCoO}_2$  cathodes with temperature threshold of 150 °C,  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  batteries do not undergo TR process at oven temperatures below 170 °C. However, the temperature peaks are more substantial in batteries with this type of cathode. Moreover, the temperature peak is more intensive in  $\text{LiCoO}_2$  batteries with  $\text{LiPF}_6/\text{PC}$ : DMC electrolyte compared to the same battery with standard electrolyte but less intensive compared to  $\text{Li}_{1.1}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})_{0.9}\text{O}_2$  batteries. In addition, batteries with Lithiated  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  anode show less intensive TR process compared to all the previous battery types.

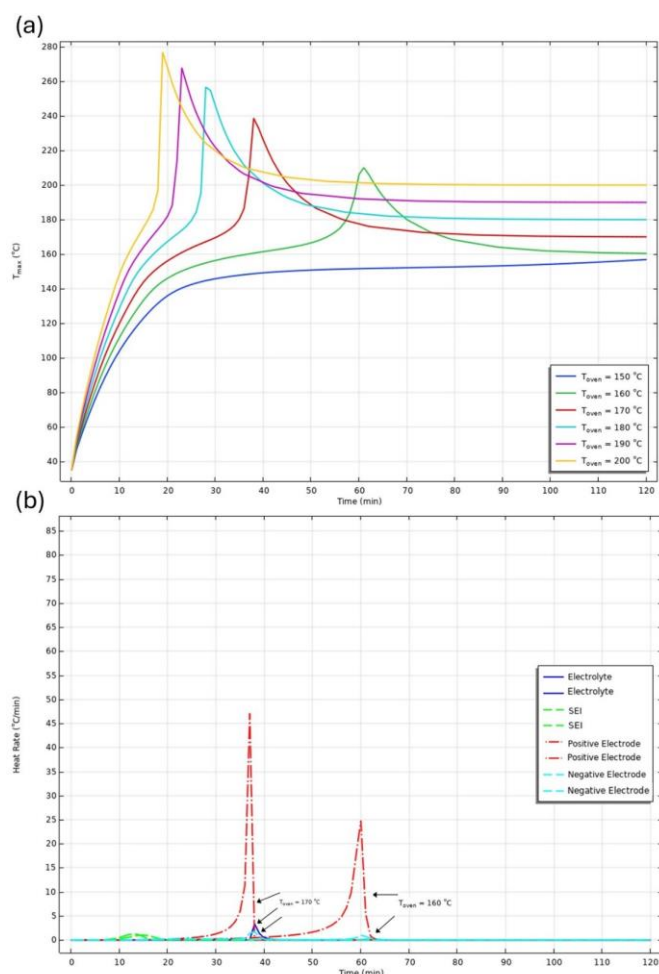


Fig. 7. TR kinetic mechanisms for LiCoO<sub>2</sub> battery with standard LiPF<sub>6</sub> electrolyte and Lithiated Li<sub>4</sub>Ti<sub>5</sub>O<sub>12</sub> anode A) Maximum temperature for Oven Temperature 140-200 °C B) Heating rate for Oven Temperature 160 and 170 °C.

The result of this study provides battery safety researchers with new insights into the thermal stability of different battery types. Further investigation into the assessment of battery materials thermal stability will foster the EV battery safety.

#### REFERNCIES

IEA (2024), Global EV Outlook 2024, IEA, Paris <https://www.iea.org/reports/global-ev-outlook-2024>, Licence: CC BY 4.0.

Wang, Y., Feng, X., Huang, W., He, X., Wang, L., and Ouyang, M. (2023). Challenges and opportunities to mitigate the catastrophic thermal runaway of high-energy batteries, *Advanced Energy Materials*, 13(15), 2203841. [doi:10.1002/aenm.202203841](https://doi.org/10.1002/aenm.202203841)

Gharehghani, A., Gholami, J., Shamsizadeh, P., and Mehranfar, S. (2022). Effect analysis on performance improvement of battery thermal management in cold weather, *Journal of Energy Storage*, 45, 103728. [doi:10.1016/j.est.2021.103728](https://doi.org/10.1016/j.est.2021.103728)

Rabiei, M., Gharehghani, A., and Andwari, A.M. (2023). Enhancement of battery thermal management system using a

novel structure of hybrid liquid cold plate, *Applied Thermal Engineering*, 232, 121051. [doi:10.1016/j.applthermaleng.2023.121051](https://doi.org/10.1016/j.applthermaleng.2023.121051)

Hatchard, T. D., MacNeil, D. D., Basu, A., and Dahn, J. R. (2001). Thermal model of cylindrical and prismatic lithium-ion cells, *Journal of The Electrochemical Society*, 148(7), A755. [doi:10.1149/1.1377592](https://doi.org/10.1149/1.1377592)

MacNeil, D. D., Lu, Z., Chen, Z., and Dahn, J. R. (2002). A comparison of the electrode/electrolyte reaction at elevated temperatures for various Li-ion battery cathodes, *Journal of power sources*, 108(1-2), 8-14. [doi:10.1016/S0378-7753\(01\)01013-8](https://doi.org/10.1016/S0378-7753(01)01013-8)

Peng, P., and Jiang, F. (2016). Thermal safety of lithium-ion batteries with various cathode materials: A numerical study, *International Journal of Heat and Mass Transfer*, 103, 1008-1016. [doi:10.1016/j.ijheatmasstransfer.2016.07.088](https://doi.org/10.1016/j.ijheatmasstransfer.2016.07.088)

Wang, Q., Sun, J., Yao, X., and Chen, C. (2006). Micro calorimeter study on the thermal stability of lithium-ion battery electrolytes, *Journal of Loss Prevention in the Process Industries*, 19(6), 561-569. [doi:10.1016/j.jlp.2006.02.002](https://doi.org/10.1016/j.jlp.2006.02.002)

Huang, P., Ping, P., Li, K., Chen, H., Wang, Q., Wen, J., and Sun, J. (2016). Experimental and modeling analysis of thermal runaway propagation over the large format energy storage battery module with Li<sub>4</sub>Ti<sub>5</sub>O<sub>12</sub> anode, *Applied energy*, 183, 659-673. [doi:10.1016/j.apenergy.2016.08.160](https://doi.org/10.1016/j.apenergy.2016.08.160)

Kong, D., Wang, G., Ping, P., and Wen, J. (2021). Numerical investigation of thermal runaway behavior of lithium-ion batteries with different battery materials and heating conditions, *Applied Thermal Engineering*, 189, 116661. [doi:10.1016/j.applthermaleng.2021.116661](https://doi.org/10.1016/j.applthermaleng.2021.116661)

Kim, G. H., Pesaran, A., and Spotnitz, R. (2007). A three-dimensional thermal abuse model for lithium-ion cells, *Journal of power sources*, 170(2), 476-489. [doi:10.1016/j.jpowsour.2007.04.018](https://doi.org/10.1016/j.jpowsour.2007.04.018)

## Assessing the impact of rising wind power with energy storage on grid resilience in Sweden to mitigate volatility and enhance grid flexibility

J.M. Jorritsma\* S. Vouros\*\* K. Kyprianidis\*\* K. Hubacek\*

\*University of Groningen, \*\*Mälardalen University

\*\*stavros.vouros@mdu.se

---

**Abstract:** This paper assesses the impact of increasing wind power production and energy storage systems on grid resilience in Sweden. Wind power currently makes up 17% of Sweden's electricity mix, and this share is expected to rise significantly in the coming decades as Sweden aims for 100% renewable energy generation by 2040. However, the variable and intermitted output can negatively impact grid stability. A microgrid model is developed, incorporating a wind turbine, battery storage, power grid, and a representative demand profile. Wind speed data is analysed to select profiles representing high and low variability, with variance used as a metric for resilience. Planned production is constructed in 12-hour intervals based on wind speed forecasts. The model compares grid dependency and electricity delivery with and without battery storage of varying capacities. The results show that battery storage reduces grid interactions and grid dependency. Furthermore, it aligns actual wind power production with the planned production profile. Optimisation analyses find that minimising operational costs and battery usage increases grid reliance while minimising costs and grid supplies provides a more stable supply but overuses batteries. Sensitivity analysis demonstrates higher grid dependency in high-variance wind conditions. The paper contributes to understanding how to enhance wind power resilience through improved production planning and battery integration. It proposes using variance analysis in wind profile selection and identifies trade-offs between system stability, costs and battery lifespan under different optimisation strategies.

---

**Keywords:** wind power, planned production, battery storage, resilience, Sweden

### 1. INTRODUCTION

Sweden primarily relies on hydropower and nuclear energy for domestic electricity production (The Swedish Energy Agency, 2023). In 2020, renewable energy sources contributed to 92% of Sweden's electricity production, with hydropower accounting for 45%, nuclear power for 29%, wind power for 17% and solar power for 1%. The remaining 8% was generated through combined heat, power, and industrial processes. Globally, there has been a rapid increase in the adoption of wind power (Benitez, Benitez and van Kooten, 2008), a trend mirrored in Sweden. The installed capacity of wind power in Sweden increased from 3,487 GWh in 2010 to 27,536 GWh in 2020, a growth attributed partly to supportive renewable electricity policies (IEA, 2019). Sweden aims to achieve 100% renewable energy production by 2040, while still retaining nuclear energy as an option (The Swedish Energy Agency, 2023). However, this goal is complicated by the predicted rise in energy demand over the coming decades, driven by various factors including emissions reduction, industry growth, hydrogen production, and the electrification of transportation and the steel industry (Holmberg and Tangerås, 2023). To address this growing demand, it is expected that wind energy production will need to increase over the coming decades (Ministry of the Environment and Energy, 2018). Current projections indicate that the installed capacity of wind power, which stood at 12,100 MW, is expected to rise to 18,500 MW by 2030 and 33,300 MW by 2040 (Swedish Wind Energy

Association, 2021). Integrating wind power into the electricity grid presents several challenges due to its inherent weather-dependent nature, which results in variable and unstable power output (Zhao *et al.*, 2015; Reddy, 2017). This variability may adversely affect the stability and performance of the electric power system, causing frequency and voltage disturbances that may lead to system shutdowns (Li *et al.*, 2021). Additionally, the intermittency of wind power affects market mechanisms for electricity trading, as these mechanisms rely on accurate production planning forecasts. Inaccuracies can lead to price fluctuations in electricity prices, particularly as wind power penetration rises (Peizheng Xuan *et al.*, 2019). The primary goal of grid operation is to meet electricity demand, however variable wind power output complicates this objective. Accurate forecasting of production and demand is crucial for determining the required amount of dispatchable electricity. Despite advancements in forecasting techniques, errors are inevitable, necessitating power reserves for grid operators (Michiorri *et al.*, 2018), ultimately hindering the integration of wind power (Zhao *et al.*, 2015). One proposed solution to mitigate these grid issues is combining wind power with energy storage systems (ESS). ESS can provide the necessary flexibility to smooth out the variability in wind power output (Zhao *et al.*, 2015; Michiorri *et al.*, 2018; Barra *et al.*, 2021). Previous research has explored various aspects of ESS integration: Li *et al.* examined short-term "power-smoothing" applications utilising high-power ESS that rapidly respond to



high power outputs (Barra *et al.*, 2021), Sperstad and Korpås *et al.* investigated the optimal scheduling of ESS in grids with large renewable energy shares, developing a framework to avoid suboptimal operations (Sperstad and Korpås, 2019). Additionally, M. Ghazipour and M. Abardeh *et al.* developed a stochastic optimisation approach for optimising the location and size of ESS in energy systems (M. Ghazipour and M. Abardeh, 2019). These studies collectively aim to enhance the understanding of ESS from various technological perspectives, addressing the volatile energy output of renewable energy sources (RES) and mitigating their adverse effects on the energy system. However, none of these studies specifically address ensuring a guaranteed electricity supply, a critical factor as the increase in wind power reduces the amount of controllable electricity supply. This aspect is vital in the broader context of energy system resilience. Acknowledging resilience is increasingly crucial within the ongoing energy transition, despite its varied definitions across multiple disciplines. Fundamentally, resilience revolves around the capacity to cope with disruptive events (Gasser *et al.*, 2021; Jasiūnas, Lund and Mikkola, 2021). One definition of resilience involves minimising service disruptions by anticipating, resisting, absorbing, adapting to and recovering from disruptive events (Ahmadi, Saboohi and Vakili, 2021). Gasser *et al.* define resilience as the capacity of systems to withstand stress, pressure or disturbance without loss of function (Gasser *et al.*, 2021). This research aims to develop a microgrid model that integrates wind power and battery energy storage, assess the role of battery storage in mitigating wind power variability, and analyse the system's resilience. By evaluating performance during disruptive wind events, this study aims to enhance the broader understanding of how ESS can enhance the resilience of renewable energy systems, ensuring a more stable and reliable electricity supply. The central question addressed is: *How can battery energy storage mitigate volatility and increase the integration of wind turbines?*

## 2. METHODOLOGY

This research employs a case study representative of recent developments in Eskilstuna, Sweden. The primary components of the microgrid model include a wind turbine, battery storage, a power grid and a representative demand profile. Two configurations will be modelled, to assess the value added by battery storage. These configurations are modelled using Modelon Impact, a systems modelling and simulation program. Modelon Impact utilises Modelica's core modelling and simulation capabilities. Modelica is an object-oriented programming language. Modelica allows for a detailed description of the behaviour of physical components and their interactions within the system.

### 2.1 Components and controls

Figure 1 illustrates the microgrid model. The Wind and Temperature blocks contain wind speed and air temperature data, respectively. The Temperature block determines the air density, directly affecting the wind power produced in the Turbine block. By incorporating these data, the model accounts for the impact of temperature-induced density variations on wind power generation. The power generated by

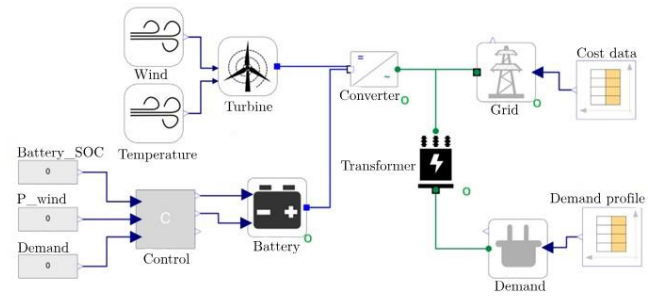


Fig. 1. The microgrid model developed with Modelon Impact.

the wind turbine is then directed to the Converter block which converts the alternative current (AC) to direct current (DC). This DC power flows through the transformer- which converts the high voltage to a lower voltage suitable for distribution within the grid. The electricity is delivered to the Demand block, representing a representative demand profile. The Demand and Demand profile blocks represent the forecasted power demand. The Grid block can provide and receive unlimited electricity to balance the grid. This setup facilitates analysis by comparing the actual power output against the planned output. An important component of the microgrid model is the Control unit. The operation of the battery is based on the net amount of power denoted by  $P_{net}$ , described in (1).

$$P_{net} = P_{wind} - P_{demand} \quad (1)$$

The Control unit measures  $P_{net}$  at each time point and operates according to the following control scheme.

- If  $P_{net} < 0$ , the required power is generated by discharging the battery, or bought from the grid
- If  $P_{net} > 0$ , the surplus is either used to charge the battery sold to the grid or both.

### 2.2 Wind power

The theoretical power that can be extracted from the wind by a wind turbine is proportional to the wind speed to the power of three (Kim, 2013). This relation is described in (2), where  $P$  represents the total wind power production by the turbine, measured in watts. The total area the turbine blades cover in one rotation is described by  $A$ , the swept area of the wind turbine in  $m^2$ ,  $\rho$  is the density of the air in  $kg/m^3$  and  $v$  represents the velocity of the wind in  $m/s$ ,  $C_p$  is the power coefficient, defined as the ratio of power extracted by the wind turbine from the energy available in the wind.

$$P = \frac{1}{2} A \rho v^3 C_p \quad (2)$$

In addition to calculating the theoretical wind power, a suitable wind speed profile is required. Moreover, a thorough wind data analysis is needed to capture resilience in a wind speed profile. Rapid and large changes in wind speeds are identified as disruptive events. One example of such events is sudden drops in high wind speeds. These abrupt changes can be quantified through statistical measures such as variance. Variance assesses the spread of data points relative to their average in the data set. Specifically, in wind speed analysis, variance indicates the degree of variability in wind speeds over time. Greater variability, as indicated by a higher variance signifies

an increased need for system resilience. A timeframe of one week is selected for the modelling phase to calculate the variance of the wind speed of the dataset. Hence the variance will be calculated for each week of the dataset. The formula for the variance is shown in (3). Where  $x_i$  is each value in the data set,  $\bar{x}$  is the mean of all values in the data set and  $N$  is the number of all values in the data set (Hui, 2018).

$$\sigma^2 = \frac{(x_i - \bar{x})^2}{N} \quad (3)$$

The selection of wind speed profiles in this research is based on several criteria. First data was sourced from a location of importance to the research region. Additionally, the profiles were chosen to represent a range of scenarios including average, low and high-wind conditions, to assess the system's resilience under diverse operational conditions.

### 2.3 Planned production

Forecasting wind speed will become increasingly paramount as future wind farms function more like conventional power plants. This transition implies a shift towards more accurate planning of electricity production, leading to the development of guidelines focused on reliability to ensure the safe operation of wind farms. Several factors are driving the shift in the role of wind power. Firstly, wind power's exposure to volatile wholesale electricity prices changes its economic dynamics. Given the relatively low marginal cost of producing wind power, increased wind power tends to decrease electricity prices. Secondly, governmental support schemes, such as feed-in-tariffs (a guaranteed cost-based purchase price for electricity), are being replaced by auctioning systems, incentivising wind farm owners to prioritise profit maximisation over pure electricity production volume. This shift underscores the growing importance of accurate wind speed forecasts in optimising wind farm operations and maximising profitability (Kölle *et al.*, 2022). Wind speed forecasts are constructed for various timeframes depending on the specific application. These include very short-term forecasts (a few seconds to 30 minutes), short-term forecasts (30 minutes to 6 hours ahead), mid-term forecasts (6 hours to a day ahead), and long-term forecasts (1 day to a week or more). Different methods, such as machine learning or statistical approaches are employed for generating these forecasts (Khosla and Aggarwal, 2022). In this research, a mid-term forecast for a half-day ahead is used, with wind speed predictions generated every 12 hours. The variables used in the planned production profile are depicted in (4) and (5), where  $v_{average}$  represents the average wind speed, and  $V_i$  represents the hourly wind speed values  $i = 1, 2, \dots, 12$ . The choice of a 12-hour planning interval aligns with the timeframe of day-ahead wholesale electricity price data, ensuring coherence between the forecasting parameters and pricing data.

$$v_{average} = \frac{1}{12} \sum_{i=1}^{12} v_i \quad (4)$$

$$\rho_{average} = \frac{1}{12} \sum_{i=1}^{12} \rho_i \quad (5)$$

### 2.4 Optimisation

The optimiser minimises the total cost, as presented in (6). Where  $L(x,u,p)$  represents the integral cost depending on the process state  $x$ , the controls  $u$ , and the plant parameters  $p$ . State variables  $x$  denote the dynamic state over time, such as the state of charge (SOC) of the battery, the power output of the wind turbine or the electricity consumption of the grid. Control variables  $u$  are decision variables that can be adjusted to optimise the system performance, such as the charging and discharging rate of the battery or the import and export to and from the grid. Parameters  $p$  are fixed values for a system, including the power efficiency of the turbine, battery capacity, electricity prices, or demand profiles. The cost integrand  $L(x, u, p)$  can be further decomposed into two terms, presented in (7).  $Cost_y$  typically refers to the operational cost per unit time (OPEX<sub>sec</sub>), while  $cost_u$  penalises the controls  $u$  ( $du^2/dt$ ) to promote smoother and more stable operation.

$$\min_{u(t),p} cost = \int_0^{T_{Optimization}} L(x, u, p) dt \quad (6)$$

$Cost_u$  can be defined as minimising the battery's aggressive charging and discharging behaviour, thus extending its technical lifetime. Generally, the penalty of  $cost_u$  is much lower than the operational cost  $cost_y$ . Dynamic optimisation aims to find the optimal trajectory  $u(t)$  while satisfying the system constraints. Modelon Impact utilises the Interior Point OPTimiser (IPOPT) to determine the best next step. IPOPT gradually narrows down search barriers within a feasible region to approach an optimal solution without reaching the boundary until close to finding it.

$$L(x, u, p) = cost_y + cost_u \quad (7)$$

In this research, different objectives are chosen to be minimised. The first scenario combines operational cost and battery controls to minimise total cost while minimising battery operation to extend the technical lifetime. The second scenario considers operational cost and the power output of the grid, aiming to minimise grid dependency. In Table 1 the optimisation scenarios are presented.

Table 1: Optimisation scenarios

Scenario	Cost <sub>y</sub>	Cost <sub>u</sub>
OPEX <sub>sec</sub> , controls	OPEX <sub>sec</sub>	Battery(Power charge * Power discharge)
OPEX <sub>sec</sub> , power grid	OPEX <sub>sec</sub>	Grid power

### 2.5 Data

This section outlines the key properties and sources of the time series data used in the research, which include air temperature, wind speed, electricity prices, and demand profiles. The following tables provide a summary of the data.

Table 2: Demand profile data

Parameter	Resolution	Period	Unit
Demand profile	12 hour interval	Weekly	[MW]

Table 3: Time series data

Parameter	Unit	Range	Period	Source
Air temperature	[K]	01/01/2010 – 01/10/2023	Hourly	(SMHI, 2023)
Wind speed	[m/s]	01/01/2010 – 01/10/2023	Hourly	(SMHI, 2023)
Electricity price	[€/MWh]	01/01/2015 – 01/12/2023	Hourly	(ENTSO-E, 2023)

### 2.6 Modelling assumptions

The battery capacity is assumed to remain constant, meaning that the battery's efficiency does not degrade over the simulation period of one week. This assumption is reasonable given the short simulation period, where the number of charging cycles during this time is insufficient to cause the battery capacity to degrade. Additionally, unlimited import and export from and to the grid is assumed. However, in real-life, the grid may encounter congestion, where the transmission network cannot meet the demand. In such cases, assets like wind farms may receive compensation from Svenska kraftnat (Transmission System Operator) to adjust production or decrease consumption accordingly (Holmberg and Tanagerås, 2022). The total installed cost of batteries decreases as the capacity increases. Most costs are calculated for a battery system with a 2-hour duration, meaning the time it can deliver its full power capacity in MW. For example, a battery with 2 MW and a 2-hour duration has a capacity of 4 MWh. In the case study, the battery system has a 1:1 power ratio (MW:MWh). Data on batteries with a 1-hour duration is limited, therefore it is assumed that the cost of batteries for different capacities is based on 2-hour duration systems. Although the transmission capacity of the power grid is assumed to be unlimited in this research, in reality, exceeding transmission capacity can result in penalties for wind farm owners. For example, if the 220 kV transmission line connecting the wind farm to the grid is exceeded, penalties may be imposed on the wind farm owner for not meeting planned production. To mitigate this, batteries with longer durations and different power ratios could enable more effective operating strategies. A 1:1 power ratio is selected, allowing the battery to discharge completely within one hour.

## 3. RESULTS

### 3.1 Wind profile selection

The wind speed variance at hub height is calculated from 2010 to 2023, plotted in Fig. 2. Each bar in Fig 2. represents a week and its corresponding variance value. A higher variance indicates greater variability in wind speed, while a lower value suggests more stable wind conditions. The highest and lowest variances are 51 and 0, respectively. It is important to note that the variance is rounded up towards the nearest integer. The variance is calculated at hub height, as the wind speed at this height determines the wind turbine's power output.

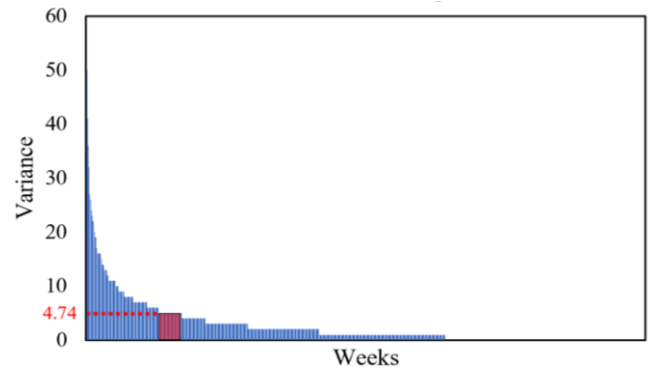


Fig. 2. Variance of the wind speed at hub height, from 2010 until 2023. The chosen wind profile is indicated in red.

### 3.2 Planned production

It is essential to establish a baseline by examining the wind power output generated solely by the wind farm, without any battery storage. This baseline gives insight into how accurate the actual production of the wind turbine is compared to the planned production. Furthermore, the interaction between the wind turbine, power grid, and planned production will be visible. In Fig. 3 the produced power of the wind turbine is presented. The planned power production and the actual power production are not balanced most of the time. During periods of imbalance, the electricity grid functions as a source and sink of electricity. Analysis indicates that 65% of the total exchanged electricity flows into the systems and 35% is delivered to the grid. Integrating battery storage aims to decrease grid interactions, especially the delivered electricity to increase power system autonomy. Battery integration with a wind turbine increases power output. This influences the total amount of electrical energy the grid has to provide. The total electricity delivered is presented in Fig. 4. The blue bar represents the delivered electricity in the scenario when only the wind turbine operates, and the yellow bars indicate the scenario in which both the wind turbine and battery are in operation. With increasing battery capacity, there is a notable decrease in the total energy demand from the grid. For instance, in the wind turbine-only scenario, the grid delivers 430 MWh. However, with 1 MW of installed battery capacity, the grid delivers 37 MWh less. At 30 MW installed capacity, the grid provides a total of 183 MWh.

### 3.3 Battery storage

The straightforward observation of decreased electricity delivered by the grid with increased battery capacity can be further analysed when looking at capacity efficiency. Capacity efficiency is defined as the difference between the delivered electricity by the grid in a scenario with only wind turbines and the electricity delivered by the grid when batteries are installed, divided by the total battery capacity. It measures how effectively the battery is utilised. For example, in Fig. 4, the total electricity delivered by the grid in the wind turbine scenario is 430 MWh, and the delivered electricity for a 2 MW battery is 382 MWh, resulting in a capacity efficiency of  $430 - 382 = 48$  MWh/MW. In Fig. 5 the capacity efficiency



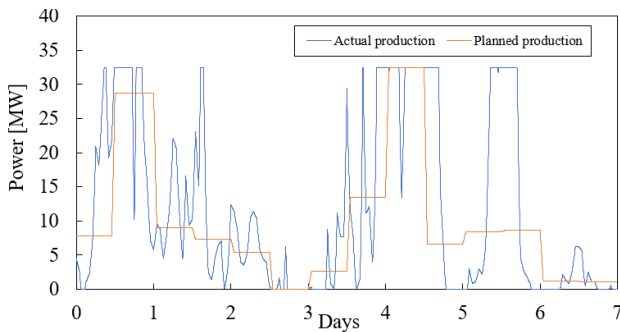


Fig. 3. The actual and planned power production of the wind farm.

for each battery size is plotted against the battery capacity. It can be observed that as battery capacity increases the capacity efficiency goes down. This implies that increasing battery capacity has diminishing returns in terms of its effectiveness in reducing grid dependency.

### 3.4 Optimisation

The results of the two optimisation scenarios in Table 1 are presented in Fig. 6 and compared with the main scenario from Fig. 4, which involves the simulation with the microgrid controller. The control strategy of the microgrid controller, as outlined in (1) focuses on maintaining grid balance by prioritising maximum utilising the battery while minimising reliance on the grid. Unlike optimisation strategies, this method does not involve optimising specific variables but rather adopts a more direct approach to grid management. The optimisation analysis reveals that integrating 1 MW of battery capacity reduces grid-supplied electricity when minimising operational cost and battery controls. Grid-delivered electricity shows a steep increase after installing 10 MW battery capacity. This observation suggests a trade-off, wherein efforts to smooth battery controls to extend battery life elevate the reliance on the grid. In the scenario aimed at minimising the operational cost and grid power output, the dependency on grid-supplied electricity remains relatively stable for each additional battery capacity. However, exceptions are noted with the installation of 2 MW and 30 MW battery capacities, where an increase in grid dependency is observed. Figure 7 presents the operational cost across all scenarios. In each scenario, the operational cost of the wind turbine and the battery are constant as they incur fixed operational expenses. Conversely, the grid's operational costs

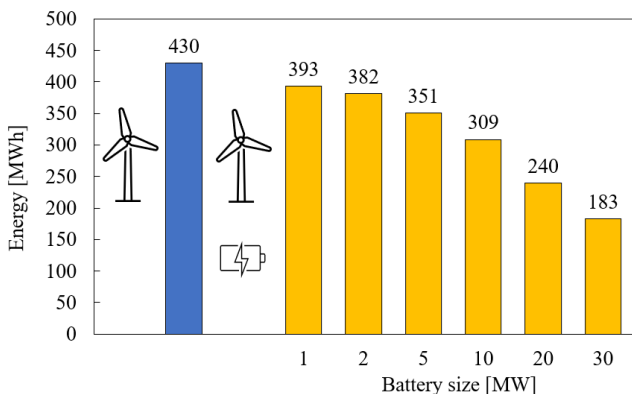


Fig. 4. The total electricity delivered by the grid with and without battery storage.

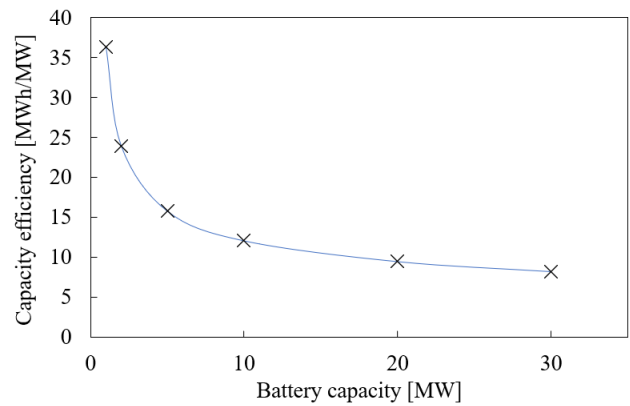


Fig. 5. The impact of battery capacity on the total amount of the grid's delivered electricity.

vary and depend on the power output and electricity price. The optimisation analysis reveals that incorporating up to 2 MW of battery storage leads to a small decrease in operational cost. Compared to the main scenario, incorporating 10 MW, 20 MW and 30 MW of battery storage leads to higher operational costs. The SOC of the battery is presented in Fig. 8. In the optimisation scenario aimed at minimising the operational cost and grid supply, the SOC begins at 0.9 and gradually decreases until 0.1 over the simulation period. Although continuous charge and discharging cycles occur, they constitute only a small fraction of the total battery capacity. In the optimisation scenario of the operational cost and battery controls, the SOC exhibits different patterns of battery utilisation. A more regular pattern is observed in the SOC of the battery, especially in the first two days of the simulation. The small operational cycle during day six indicates a degree of flexibility in deviating from the optimised battery controls to minimise operational costs. Compared to battery controls a higher penalty is associated with optimising operational costs. In contrast, the SOC during the simulation with the microgrid controller indicates a more frequent utilisation of the battery, aligning with the result of Fig. 6, which suggests reduced grid dependency. Specifically, the microgrid controller simulation experiences a total of fourteen operational cycles, whereas the battery controls optimisation scenario only experiences a total of eleven operational cycles. This indicates a trade-off wherein the grid can be stabilised at the expense of potentially overusing the battery, or reducing stress on the battery by increasing dependency on the grid.

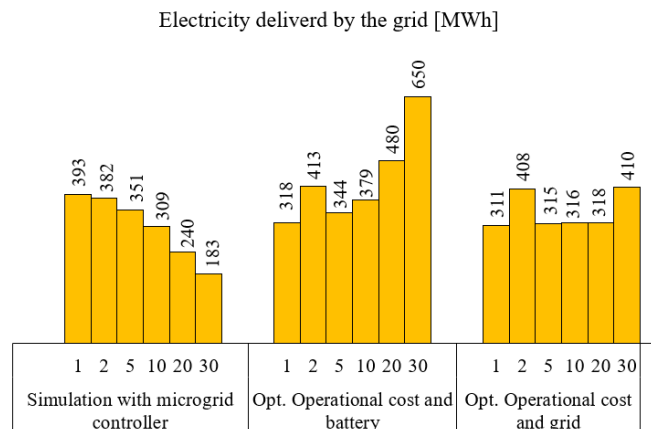


Fig. 6. Delivered power grid electricity for the main and Optimisation scenarios.

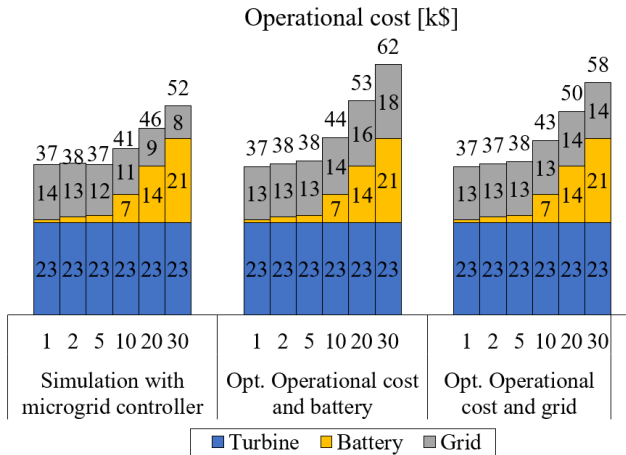


Fig. 7. The operational cost of the simulation and Optimisation scenarios for every battery size.

### 3.5 Sensitivity analysis

To measure the impact of variance on the results, different wind speed profiles and their 12-hour interval average are presented in Fig. 9. What can be observed from this analysis is that the high variance profile exhibits more pronounced peaks and troughs in wind speed throughout the week. High peaks are observed on the first and third days of the simulation. The low-variance profile demonstrates a more stable pattern and gradually increases from 0 m/s on the first day to approximately 5 m/s by the end of the week. Quantitatively, the high variance profile has a variance of 16.01 and the low variance has a value of 2.27. Figure 10 presents the electricity delivered by the grid in the high and low-variance scenarios, showcasing a similar trend to that of the main scenario. In the high variance scenario, more electricity is delivered by the grid in both the wind turbine-only simulation and the incorporation of each additional battery storage capacity. This observation can be attributed to the deviations between planned and actual production caused by the high variance in wind speed. The grid intervenes to align the actual with planned production, resulting in an increased demand for electricity from the grid. Conversely, in the low variance scenario, less electricity is required from the grid as the actual production closely aligns with the planned production due to lower variance in wind speed. Consequently, there is a reduced need for the grid to

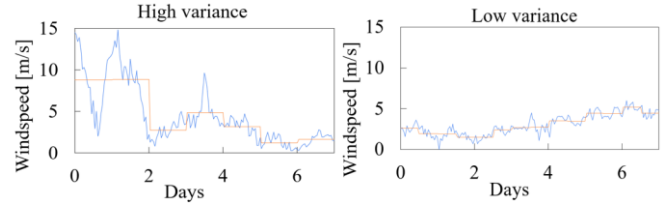


Fig. 9. The selected wind speed profiles and their average wind speeds.

intervene to align the two. For instance, in the wind turbine-only simulation, the electricity delivered by the grid is 116 MWh in the low variance scenario but decreases to 0 MWh when 30 MW of battery capacity is installed, this indicates the significant impact of variance on grid dependency and the effectiveness of battery storage mitigating it. In addition to analysing variations in wind speed, it is essential to consider the potential effects of model prediction uncertainties on the system's performance. Prediction errors, whether in wind speed or demand can lead to imbalances in supply and demand, compromising grid stability. Errors in forecasting can lead to inefficient dispatch resulting in increased operational costs and can reduce system efficiency. Storage sizing depends on accurate forecasts, if errors are not accounted for, resilience strategies may be underutilized.

## 4. DISCUSSION

### 4.1 Battery optimisation

Michiorri et al. researched strategies to minimise power errors in wind turbines and optimise battery storage sizing in a 9 MW wind farm. The wind farm owner provided to the transmission system operator with 30-minute interval power predictions. A 5 MW power-rated battery, resulted in high penalties and periods of disconnection. To address this, a sizing methodology was developed that generated error time series characterised by their autocorrelation. This led to an optimal capacity. A smaller-sized battery performed better because it effectively absorbed prediction errors correlated over timescales of around 6 hours, rather than compensating for all the differences between actual and predicted output over time. Consequently, a smaller battery reduced penalties while still achieving the target level of allowable errors, allowing it to be utilised to its full technical potential (Michiorri et al., 2018). In this research, no error range is employed for the operational strategy, resulting in immediate battery utilisation whenever there is a misalignment between predicted and actual power. This complicates the comparison between the study of Michiorri et al and the current research. However, both studies agree that a smaller battery can better utilise its full potential. This is demonstrated in the present research, where the battery capacity increases and the capacity efficiency decreases.

### 4.2 Production planning interval

This research shows that a 12-hour interval accumulates production planning errors in a high-variance scenario, resulting in increased electricity from the grid. As the variance decreases, the forecast error also decreases, suggesting that a 12-hour forecast interval is more suitable for low-variance wind profiles. In contrast, high-variance wind profiles could

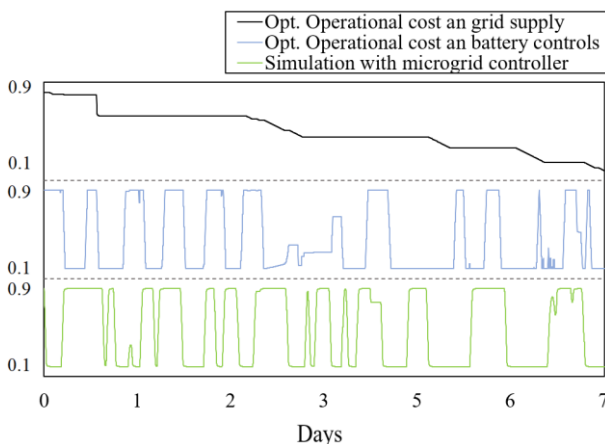


Fig. 8. The battery SOC during the simulation run with the microgrid controller and the optimisation runs.

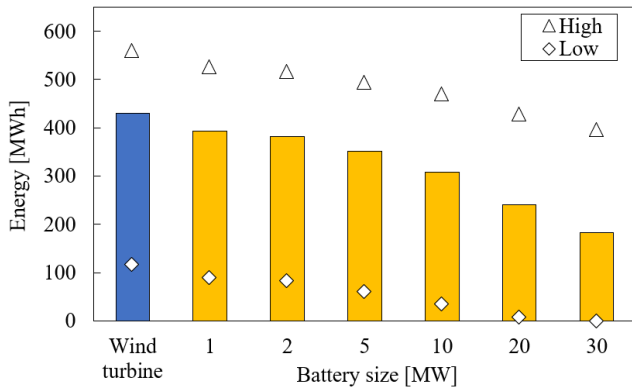


Fig. 10. Electricity delivered by the power grid: high and low wind profile.

benefit from a narrower interval. Y et al. investigated the optimisation of a self-disciplined interval of a wind farm. This interval is calculated assuming an error distribution around the mean of the predicted power output. An estimation technique models the historical error distribution shape between the actual and predicted wind power output. This interval width is optimised using the IPOPT. A case study analyses a 10-minute simulation interval to validate this method. Optimised battery storage technology supplies the necessary power to maintain this interval. It was found that the optimised method can effectively improve the self-disciplined level. Showing shorter intervals is an effective way of constructing robust production planning. A limitation of this research is the exclusion of battery degradation. Furthermore, only one case study is used and the method does not consider any error in actual wind power production (Yu *et al.*, 2020).

#### 4.3 Policy decisions

Policy decisions should focus on resilient energy infrastructure, with investments made in battery optimisation to achieve cost-effective grid independence. Furthermore, policies should address grid congestion and provide compensation mechanisms for energy producers. Increase penalties for overproduction to encourage efficient energy management. Furthermore, high upfront costs and varying electricity prices are barriers to large-scale deployment of battery storage. Governments can provide subsidies for stable pricing mechanisms and long-term contracts to ensure financial security.

### 5. CONCLUSIONS

This research aims to investigate how battery storage can mitigate the volatility of wind power and its implications for the resilience of the Swedish energy system upon integration into the power grid. Given the growing trend of wind power with battery storage in Sweden, this study contributes to our understanding of improving wind turbine resilience through better production planning. Presented below are the main findings stemming from this research:

- Incorporating battery storage significantly reduces dependency on the power grid, especially in the lower-variance wind profiles.

- Enhanced utilisation of batteries is observed as battery capacity decreases.
- The research introduces a method for selecting wind speed profiles based on variance analysis, which captures the dynamic nature of wind behaviour. This approach identifies disruptive events through variance, providing a nuanced understanding of wind variability and system resilience.
- Enhancing system resilience by reducing grid dependency can increase capital and operational costs. Consequently, this leads to a higher variability in the SOC of the battery while smoothing the power grid supply. This creates a trade-off between stabilising the grid by heavily using the battery and protecting battery life by relying more on the grid.
- High-variance wind speed profiles lead to greater discrepancies between planned and actual production, requiring more grid intervention. In contrast, low-variance profiles aligned better with forecasts, reducing grid dependency.

The research identifies a clear trade-off between battery usage and grid dependency. While battery integration reduces grid reliance, it also necessitates careful consideration of battery control strategies to prevent increased operational costs and ensure battery longevity. This insight is crucial for optimising microgrid performance and achieving a more autonomous, cost-effective, resilient power system.

#### ACKNOWLEDGEMENTS

This work has been partially supported by project ProPlan under RESILIENT Competence Center (project number 2021-90273) and project GEFWIN (project number 20 22-02407) funded by Vinnova under Eurostars/Eureka. Authors extend their acknowledgements to Modelon for providing the software, tools, and support in the modelling and simulation part of this work.

#### REFERENCES

- Ahmadi, S., Saboohi, Y. and Vakili, A. (2021) Frameworks, quantitative indicators, characters, and modeling approaches to analysis of energy system resilience: A review, *Renewable and Sustainable Energy Reviews*, 144. doi:10.1016/j.rser.2021.110988.
- Barra, P.H.A., and de Carvalho, W.C. (2021) A review on wind power smoothing using high-power energy storage systems, *Renewable and Sustainable Energy Reviews*, 137. doi:10.1016/j.rser.2020.110455.
- Benitez, L.E., Benitez, P.C. and van Kooten, G.C. (2008) The economics of wind power with energy storage, *Energy Economics*, 30(4), 1973–1989. doi:10.1016/j.eneco.2007.01.017.
- ENTSO-E (2023) *entso-e map of day ahead prices*. <https://www.entsoe.eu/> (Accessed: 12 February 2024).

- Gasser, P., and Lustenberger P. (2021) A review on resilience assessment of energy systems, *Sustainable and Resilient Infrastructure*, 6(5), 273–299. doi:10.1080/23789689.2019.1610600
- Holmberg, P. and Tangerås, T.P. (2023) *The Swedish electricity market – today and in the future*. [https://www.riksbank.se/globalassets/media/rapporte\\_r/pov/artiklar/engelska/2023/230512/2023\\_1-the-swedish-electricity-market--today-and-in-the-future.pdf](https://www.riksbank.se/globalassets/media/rapporte_r/pov/artiklar/engelska/2023/230512/2023_1-the-swedish-electricity-market--today-and-in-the-future.pdf) (Accessed: 31 May 2024).
- Hui, E.G.M. (2018) *Learn R for applied statistics: With data visualizations, regressions, and statistics*, *Learn R for Applied Statistics: With Data Visualizations, Regressions, and Statistics*. Apress Media LLC. doi:10.1007/978-1-4842-4200-1.
- IEA (2019) *Energy Policies of IEA Countries - Sweden 2019 Review*. [https://iea.blob.core.windows.net/assets/abf9ceee-2f8f-46a0-8e3b-78fb93f602b0/Energy\\_Policies\\_of\\_IEA\\_Countries\\_Sweden\\_2019\\_Review.pdf](https://iea.blob.core.windows.net/assets/abf9ceee-2f8f-46a0-8e3b-78fb93f602b0/Energy_Policies_of_IEA_Countries_Sweden_2019_Review.pdf) (Accessed: 31 May 2024).
- Jasiūnas, J., Lund, P.D. and Mikkola, J. (2021) Energy system resilience – A review, *Renewable and Sustainable Energy Reviews*, 150. doi:10.1016/j.rser.2021.111476.
- Khosla, A. and Aggarwal, M. (2022) *Renewable Energy Optimization, Planning and Control Studies in Infrastructure and Control*. <http://www.springer.com/series/16625>.
- Kim, C. (2013) *Design and Simulation of Micro-Power System of Renewables*. <https://www.mwfr.com/homer/Lecture2.pdf> (Accessed: 31 May 2024).
- Kölle, K., Göçmen, T., Eguinoa, I., Alcayaga Román, L. A., Aparicio-Sanchez, M., Feng, J., Meyers, J., Pettas, V., and Sood, (2022) FarmConnors market showcase results: wind farm flow control considering electricity prices, *Wind Energy Science*, 7(6), 2181–2200. doi:10.5194/wes-7-2181-2022.
- Li, X., Song, J., Yang, S. and Li, Z. (2021) Primary Frequency Regulation Technology of Power Grid and Frequency Regulation Potential Analysis of Hydrogen Fuel Cell, in *Proceedings of the 16th IEEE Conference on Industrial Electronics and Applications, ICIEA 2021*. Institute of Electrical and Electronics Engineers Inc., 357–362. doi:10.1109/ICIEA51954.2021.9516368.
- Ghazipour, M. and Abardeh, M. (2019) Optimal Operation of Energy Storage System in a Hybrid Microgrid Considering Load Uncertainty, *IEEE Transactions on Sustainable Energy*, 1–6.
- Michiorri, A., Lugara, J., Siebert, N., Girard, R. and Kariniotakis G. (2018) Storage sizing for grid connected hybrid wind and storage power plants taking into account forecast errors autocorrelation, *Renewable Energy*, 117, 380–392. doi:10.1016/j.renene.2017.10.070.
- Ministry of the Environment and Energy (2018) *Sweden's draft integrated national energy and climate plan*. [https://energy.ec.europa.eu/system/files/2020-03/se\\_final\\_necp\\_main\\_en\\_0.pdf](https://energy.ec.europa.eu/system/files/2020-03/se_final_necp_main_en_0.pdf) (Accessed: 31 May 2024).
- Xuan, P., Zhu, J., Zou, J. and Xie, P. (2019) Research on the Electricity Market Clearing Model and Price Fluctuation with Wind Power, *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2936–2940.
- Reddy, S.S. (2017) Optimal scheduling of thermal-wind-solar power system with storage, *Renewable Energy*, 101, 1357–1368. doi:10.1016/j.renene.2016.10.022.
- SMHI (2023) *Download meteorological observations*. <https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer/#param=wind,stations=core> (Accessed: 12 February 2024).
- Sperstad, I.B. and Korpås, M. (2019) Energy storage scheduling in distribution systems considering wind and photovoltaic generation uncertainties, *Energies*, 12(7). doi:10.3390/en12071231.
- Swedish Wind Energy Association (2021) *Roadmap 2040 Wind power: combating climate change and improving competitiveness*. <https://swedishwindenergy.com/wp-content/uploads/2021/01/Roadmap-2040-ENG-rev-2020.pdf> (Accessed: 6 November 2023).
- Yu, X., Zhang, W., Dong, X., Liu, S., Pang, S. and Zang, H. (2020) Optimization of wind farm self-discipline interval and energy storage system configuration, *IEEE Access*, 8, 79114–79123. doi:10.1109/ACCESS.2020.2989306.

# CFD validation of optimized compact heat exchanger designs

Johan R. Espelund, Geir Skaugen

*SINTEF Energy Research, Sem Sælands vei 11, Trondheim, Norway*

*Corresponding author: geir.skaugen@sintef.no*

**Abstract:** In offshore oil and gas production gas turbines are used for both power production and to provide process heat. CO<sub>2</sub> emissions from the gas turbines accounts for about 25 % of the total Norwegian emissions and installing a bottoming cycle to produce power by recovering heat from the gas turbine exhaust is one way to reduce these missions. When installing a steam bottoming cycle offshore, the total weight and size will be important, and there is a need for a compact heat recovery steam generator (HRSG). A compact HRSG will often need to be designed with smaller tube diameters than conventional on-shore steam generators. To increase confidence in the compact design, the heat transfer and pressure loss models need to be accurate for the relevant geometry ranges. In this work, a compact Once Through Steam Generator (OTSG) is designed using optimisation procedures where the total weight of the steam generator has been minimised for a desired duty with restrictions for pressure losses. A range of correlations from the literature were used for the calculation of the performance. The results from the optimisation show that the 'heaviest' results were about three times the minimum weight than the 'lightest'. To increase confidence in the results, and to provide a recommendation for design models, a validated CFD model was used to perform a numerical analysis of the optimised geometry and compare this with the correlations.

*Keywords:* heat exchanger optimization, finned tube bundle, heat transfer, pressure drop, CFD

## 1. INTRODUCTION

The oil and gas industry is a significant contributor to the Norwegian overall emissions, with offshore installations accounting for a quarter of total greenhouse gas emissions. Most of these emissions come from gas turbines used on the platforms. To reduce CO<sub>2</sub> emissions towards 2030, the installation of steam bottoming cycles for power production has been proposed as illustrated in Fig. 1. Another technology to reduce emissions from gas turbines is to use carbon-free fuels like for instance mixtures including ammonia and hydrogen. Then the fuels must be imported off-shore and installing a bottoming cycle will have considerable fuel-saving potential. Weight and size of these cycles are

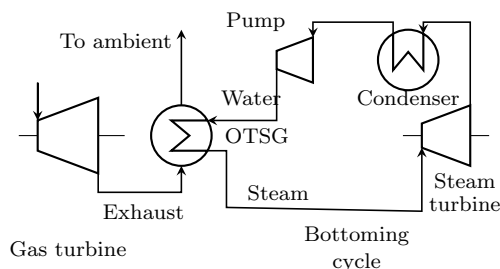


Fig. 1. Gas turbine with a steam bottoming cycle.

currently limiting factors for their widespread implementation, and efforts have already been made to develop designs

<sup>1</sup> This publication has been funded by HighEFF - Centre for an Energy Efficient and Competitive Industry for the Future, an 8-year Research Centre under the FME-scheme (Centre for Environment-friendly Energy Research, 257632/E20). The authors gratefully acknowledge the financial support from the Research Council of Norway and user partners of HighEFF.

that meet these criteria. Previous studies have shown that using small diameter tubes in the heat exchanger is one of the key factors in achieving compact design. However, the design of the heat exchanger currently relies on empirical correlations that are not necessarily validated for offshore geometries. As a result, designs can vary greatly depending on which correlation is chosen. Mazzetti et al. (2021) and Deng et al. (2021) developed an optimisation procedure for steam bottoming cycle design, which demonstrated that the heat recovery steam generator (HRSG) was the main contributor to the total weight of the cycle. They showed that optimising the HRSG with minimum weight as the objective, the lowest possible tube diameter was always selected. In Montañés et al. (2023), a similar optimisation study for a combined heat and power bottoming steam cycle was done. Here, typical available tube diameters and wall thicknesses were chosen for the HRSG while the remaining geometry parameters were optimized. The results showed a clear trend on how the obtained minimum weight increased with the selected tube diameter. The fin height and the fin- and tube spacing were different for each tube diameter. To increase the confidence in these results and alternative to experimental work, numerical studies with CFD can be an option. Lindqvist and Næss Lindqvist and Næss (2018) developed a steady-state CFD model for plain and serrated fin tube bundles, which was validated against available experimental data. The layout angle was limited to 30°, and both solid and serrated fins were investigated. A periodic domain was used, where a single tube row is modeled. The results were compared to a simulation where 8 tube rows were modeled, and it was found that the periodic model provided near identical results at a fraction of the computational cost. The numerical results were also compared with some widely used correlations,

showing that the CFD model generally was more accurate than the correlations, being within 20 % of experimental values. It was found that none of the investigated empirical correlations yielded results similar to those of the CFD model for all geometries.

The current author, Espelund et al. (2022), expanded on this work by including tube bundles also with larger layout angles. It was found that the steady-state approach did not converge for larger angles, but transient simulations gave accurate results. Comparisons with experiments showed that the transient CFD model agreed within 20 %.

In this work, an optimised set of design parameters using the different exhaust side heat transfer and pressure loss models from Table 1 are found and the different minimum weights are compared. In earlier studies of compact Once Through Steam Generators (OTSG's) by Mazzetti et al. (2021), Deng et al. (2021), and Montañés et al. (2023), the ESCOA correlations have been used as the basis for the thermal design. Therefore, the thermal performance (duty and exhaust side pressure loss) when using the other models with the optimized "ESCOA geometry" has been evaluated as a pure simulation. Simulated Nu- and Eu-number are compared to the results from the CFD analysis.

## 2. GEOMETRIES AND CORRELATIONS

The heat recovery heat exchanger in a steam bottoming cycle can be a Once Through Steam Generator (OTSG). The OTSG geometry consists of finned tube bundles, either as a single core or divided into different bundles, each representing the economizer, the evaporator, and the superheater. In this case study, a single tube bundle is used as the base case. Relevant OTSG geometry parameters are illustrated in Fig. 2. A staggered tube pattern is used with a fixed layout angle of 30° as shown in Fig. 2(a). In this study, only models for solid round fins have been evaluated. Figure 2(c) shows the notation for the number of passes and rows per pass. The number of tubes per row  $N_t$ , is the number of tubes along a header.

The performance of the OTSG is calculated as a cross-counter flow heat exchanger with exhaust flowing upward. The exhaust flow is flowing across 60 tube rows (30 · 2). Each parallel circuit, "tubes per row", is defined to have equal performance so the problem is 2-dimensional. Each tube pass is divided into 20 sub-elements and the heat balance between gas and water/steam is solved for each of the the 1200 heat exchanger elements. For each sub-element the transferred heat,  $Q$ , is calculated from

$$Q = U \cdot A \cdot \Delta T. \quad (1)$$

The overall heat transfer coefficient  $U$  (W/(m<sup>2</sup>K)) is calculated from (2) based on internal and external heat transfer coefficient,  $\alpha_i$  and  $\alpha_o$ , the tube side and fin side surface areas,  $A_i$  and  $A_o$ , and the conductive resistance through the tube wall,  $R_w$ .

$$U = \left[ \frac{1}{\alpha_i \cdot \frac{A_o}{A_i} + R_w + \frac{1}{\alpha_o}} \right] \quad (2)$$

In the performance calculations, the correlation from Bennett and Chen (1980) is used for the two-phase heat

transfer when evaluating  $\alpha_i$  in (2). Two-phase frictional pressure loss is calculated with the Friedel (1979) correlation. The extended surface "apparent" heat transfer coefficient  $\alpha_o$  in Eq. (2) is calculated from a correlation for the Colburn  $j$ -factor or Nu-number that uses the detailed fin and tube geometry into account. The derived heat transfer coefficient is the  $\alpha_c$  in Eq. (3) where  $\eta_f$  is the fin efficiency,  $A_f$  is the fin surface area and  $A_o$  is the total outside surface area.

$$\alpha_o = \alpha_c \cdot \left[ 1 - (1 - \eta_f) \frac{A_f}{A_o} \right] \quad (3)$$

The relations between the heat transfer coefficient  $\alpha_c$ , the  $j$ -factor, Stanton- (St), Prandl- (PR) and the Nusselt (Nu)-number are shown in Eqs. (4) and (6).

$$j = \text{St Pr}^{2/3} = \alpha_c \frac{\text{Pr}^{2/3}}{\rho \cdot u_{\max} \cdot c_p} \quad (4)$$

$$\text{Nu} = \frac{\alpha_c d_h}{\lambda} \quad (5)$$

$$\text{Pr} = \frac{\eta \cdot C_p}{\lambda} \quad (6)$$

where  $\eta$  (Pa.s) is the dynamic viscosity,  $C_p$  (J/(kg K)) is the specific heat capacity,  $\lambda$  (W/(m K)) is the thermal conductivity and  $d_h$  is the hydraulic diameter. The Re-number is based on the maximum velocity,  $u_{\max}$  inside the tube bundle and with the hydraulic diameter as the diameter at the fin base, namely

$$\text{Re} = \frac{u_{\max} d_h}{\nu}. \quad (7)$$

We have assumed L-fin where the fin base diameter, so the hydraulic diameter,  $d_h$ , is the outside tube diameter  $d_o$  plus  $2 \cdot F_t$  where  $F_t$  is the fin thickness. The fin efficiency  $\eta_f$  in Eq. (3) is calculated as recommended by the correlation while the temperature difference  $\Delta T$  in Eq. (1) is calculated for the arithmetic mean temperature difference between inlet and outlet fluid temperatures on inside and outside. Since the performance calculation model is called from an optimisation routine, we want to avoid unnecessary iterations so the OTSG is solved from the "warm" end, following the exhaust flow from the warm inlet. From a desired steam outlet temperature and pressure, the calculation is done backward to the feed water flow direction inside the tubes. When solved, the unknown inlet state for the steam and outlet state for the exhaust can be found. To solve the heat balance in each integration step, the 2'nd order Heun's method is used to accurately predict the temperature difference and the transferred heat,  $Q$ , from the warm to the cold end. After the integration, the total heat duty and the pressure losses for the exhaust and waterside are known and used by the optimisation routine in the constraint evaluations. The various correlations used for heat transfer and pressure loss on the exhaust side are listed along with the in Table 1

The pressure drop is calculated using the Euler number Eu, which is defined as the pressure drop across a tube row normalized by the dynamic pressure,

$$\text{Eu} = \frac{\Delta p}{\frac{1}{2} \rho u_{F_{\min}}^2} \quad (8)$$



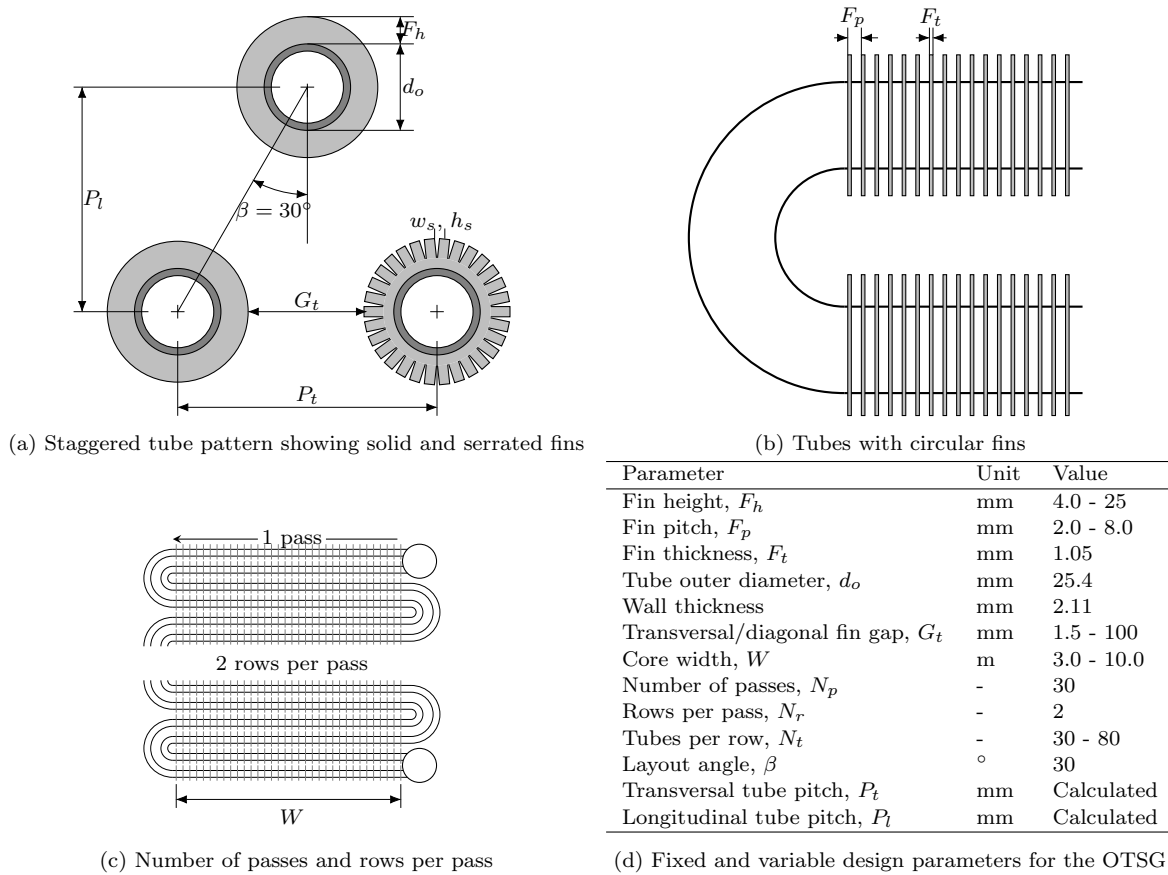


Fig. 2. Geometry definitions for the OTSG used by the design optimisation model.

where  $\Delta p$  is the pressure drop across one tube row and  $U_{F_{min}}$  is the velocity through the minimum flow area.

Table 1. List of evaluated correlations for outside heat transfer and pressure loss for round solid fins

Correlation	Heat transfer Pressure loss
	Briggs and Young (1963)
Briggs & Robinson	Robinson and Briggs (1966)
Stasiulevicius	Stasiulevicius et al. (1988)
PFR	Rosenman (1976)
Mon	Mon (2003)
ESDU	Hewitt (1998) (ch 2.5.3)
ESCOA	Sec. Ganapathy (2003)
Holfeld	Holfeld (2016)
Lindqvist	Lindqvist (2019)
Rabas	Rabas et al. (1981)

### 3. OPTIMIZATION PROCEDURE

The optimisation procedure is set up as a flowsheet optimisation model with only the OTSG as a single unit model. It is implemented in an in-house heat exchanger modelling software by Skaugen et al. (2013). The optimisation was done with the NLPQL model from Schittkowski (1986). NLPQL requires a function for the calculation of an objective and functions for the calculation of all equality and inequality constraints.

The optimisation model in this work is configured to minimise the weight of the OTSG in Fig. 1 for a given steam

production and duty. The steam turbine and other components in the steam cycle are not included. The gas turbine is a natural gas-fired gas turbine with around 30 MW power output. The exhaust flow rate is 86.1 kg/s with a temperature of 510°. The corresponding water/steam boundary conditions used are: Feed water flow rate and temperature of 10.2 kg/s and 17.0 °C. The inlet feed water pressure is 29.5 bar.

The geometry design parameters, and their range, are listed in Table 2d. These are the free optimisation variables for the problem.

The underlying heat exchanger model consists of the *thermal simulation model* described in Sec. 2 combined with *weight calculation model* for the fin and tube weights plus an estimation for the total weight of the casing with plates, beams, and insulation. So, for a set of geometry input parameters, the free optimisation variables, the thermal model finds the transferred heat and pressure losses while the weight model estimates the total weight. The calculated results are converted into constraints and objective for the optimisation routine. By minimising the total weight, and not only the bundle weight, the model ensures that the size and shape of the tube bundle will require unnecessary large inlet/exit transition ducts for instance. The tube bundle will typically account for about 50% or less of the total weight for the OTSG. The optimisation problem is defined in Table 2. The condition for the required duty is the only equality constraint. NLPQL evaluates the constraints and the objective function and



proposes a new set of design variables (from Table until the minimum is found and none of the constraints are violated.

Table 2. Defined constraints and objective function for the optimisation problem

Description	Condition
OTSG Duty	$Q = 34$
Exhaust pressure loss	$\Delta p_{ex}$
Water/steam pressure loss	$\Delta p_{ws}$
Maximum exhaust velocity	$U_{max} \quad /s$
Diagonal tube pitch	$P_d < 3$
Objective	min (Total weight)

The optimisation routine does not guarantee that the global minimum is found, so to improve confidence in the result a multi-start where each case has been run with 10 random sets of initial values for the free variables within the specified range.

#### 4. CFD MODEL

A CFD model for plain and serrated fin tube banks has previously been developed by the current author (Espelund et al., 2022). A detailed description can be found in (Espelund, 2022), but a summary of the model is given here. The simulations were run on the IDUN HPC cluster (Själänder et al. 2021). The incompressible Navier-Stokes equations are solved for the external flue gas in a periodic domain, while the heat equation is solved in the fins. The domains are coupled with thermal boundary conditions, and the equations are solved using the `chtMultiRegionFoam` solver in OpenFOAM v2206.

##### 4.1 Governing equations

The Navier-Stokes equations are solved in the gas region. They are constituted by the continuity, momentum, and energy equations. In this section, Einstein notation is used, with  $i \in \{1, 2, 3\}$  corresponding to the three spatial coordinates. The continuity equation reads,

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (9)$$

where  $u_i$  is the velocity component in the  $i$  direction. The gas is modelled as incompressible and with constant thermal and transport properties, yielding the following momentum balance equations,

$$\frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu_{eff} \frac{\partial^2 u_i}{\partial x_j \partial x_j} + S_i, \quad (10)$$

where  $p$  is the pressure field,  $\rho$  is the density and  $\nu_{eff}$  is the effective kinematic viscosity (accounting for both molecular and turbulent viscosity). The turbulence model by Spalart and Allmaras (1992) is used.  $S_i$  is an external driving force (corresponds to the pressure loss  $\frac{\partial p}{\partial x_i}$ ) to drive the flow through the cyclic domain, as described by Patankar et al. (1977).

The energy conservation equation is formulated using the specific enthalpy  $h$ , viz.

$$\frac{\partial(\rho h + e_K)}{\partial t} + \frac{\partial}{\partial x_i} (u_i(\rho h + e_K)) - \frac{Dp}{Dt} = -\frac{\partial q_i}{\partial x_i}, \quad (11)$$

where the heat flux is given by Fourier's law  $q_i = \rho \alpha_{eff} \partial h / \partial x_i$ , where the effective thermal diffusivity is

defined as  $\alpha_{eff} = \kappa_{eff} / (\rho c_p)$  and  $e_K = \frac{1}{2} \rho u_i u_i$  is the specific kinetic energy. Here  $\kappa_{eff}$  is the effective thermal conductivity. In the solid region, the special case of  $u_i = Dp/Dt = 0$  in Equation (11) is solved. The solid region is also assumed to have constant thermal properties.

To reduce the time needed to reach pseudo-steady-state conditions, steady-state simulations are run initially, using the `chtMultiRegionSimpleFoam` solver. Once the average fin temperature  $T_f$  stabilises, the steady-state simulation is terminated, and the transient simulation is initiated. To ensure converged results and a sufficiently long sampling interval, the transient simulations were run for 30 fluid exchange times  $\tau = P_l / U_{F,max}$ . Data were sampled for  $\tau > 15$ .

##### 4.2 Geometry and discretization

A periodic domain is used to model the heat exchanger, and the discretization follows the procedure in Lindqvist and Næss (2018). The geometry is specified according to the optimised solid fin OTSG. The mesh consists of hexahedral cells, with wall refinement at the fins sides and at the tube surface. The mesh at the edges of the fins are not refined towards the wall boundary, which means that wall functions are needed to model the turbulent profiles here. The numerical mesh is illustrated in Fig. 4.

##### 4.3 Boundary conditions

By adding an external pressure force  $S_i$  in the momentum equation, Eq. (10), cyclic boundary conditions can be used in all directions. As the pressure gradient is not known a priori, the `meanVelocityForce` option is used in OpenFOAM. This will adjust the source term to reach a specified mean velocity, which was used to fix Re to relevant values from the optimized OTSG design.

The temperature field needs additional treatment since the total heat transfer to the domain is not known a priori. To keep a fixed average inlet temperature, a `cyclic jump` boundary condition is used between the inlet and outlet for temperature.

$$T_{in}(x, y) = T_w + \frac{T_{out}(x, y) - T_w}{\bar{T}_{out} - T_w} \cdot (T_{in,target} - T_w), \quad (12)$$

where  $T_{in}(x, y)$  and  $T_{out}(x, y)$  is the inlet and outlet temperature fields, respectively,  $T_w$  is the constant wall temperature and  $T_{in,target}$  is the constant target inlet temperature. The average temperature is defined as

$$\bar{T} = \frac{1}{\dot{m}} \int_A T(x, y) \rho u_i dn_i, \quad \text{where} \quad \dot{m} = \int_A \rho u_i dn_i, \quad (13)$$

and  $n_i$  is the patch normal vector and  $A$  is the domain of the patch. Thus, this is a mass flux weighted average which ensures a fixed inlet temperature, but with a cyclic profile that also ensures that the temperature is constant at the walls. In this work, the conditions of  $T_{in} = 320$  K and  $T_w = 300$  K has been used, which corresponds to experimental conditions for correlation development.

At the interface between the gas and solid regions, no-slip and no-penetration are prescribed the velocity field and

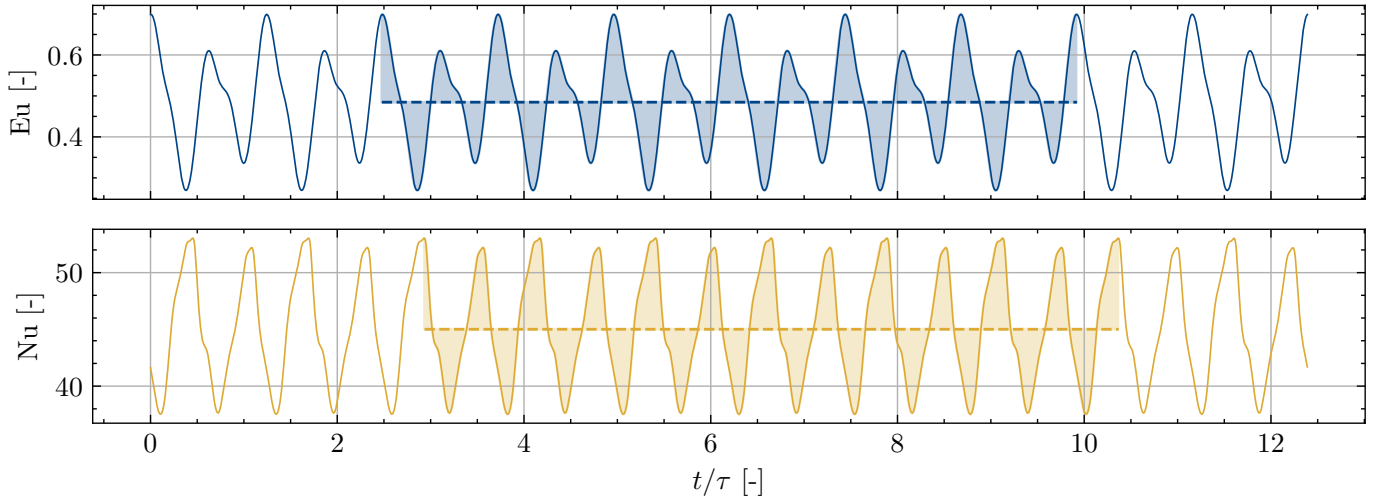


Fig. 3. Illustration of sampling. Time signal for Eu and Nu, with the sampling region shaded around the time average value. The signals are from the coarsest mesh simulation with  $Re = 7500$ . The illustration shows 6 periods used for sampling, while the actual sampling was performed with 15 periods.

von Neumann-boundary conditions for the pressure. The temperature field is coupled in the two domains and obeys continuity in temperature and energy conservation, i.e.

$$T_{w,gas} = T_{w,fin} \quad \text{and} \quad q_i n_i|_{w,gas} = -q_i n_i|_{w,fin}, \quad (14)$$

where  $n_i$  is the wall normal direction.

#### 4.4 Post-processing

The transient temperature, heat flux and pressure drop data were time averaged as

$$X = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t) dt, \quad (15)$$

where  $t_1$  and  $t_2$  is the start and end time of the sampling interval, respectively.  $x(t)$  is the transient time signal and  $X$  is the time average. The sampling of Nu and Eu is illustrated in Fig. 3.

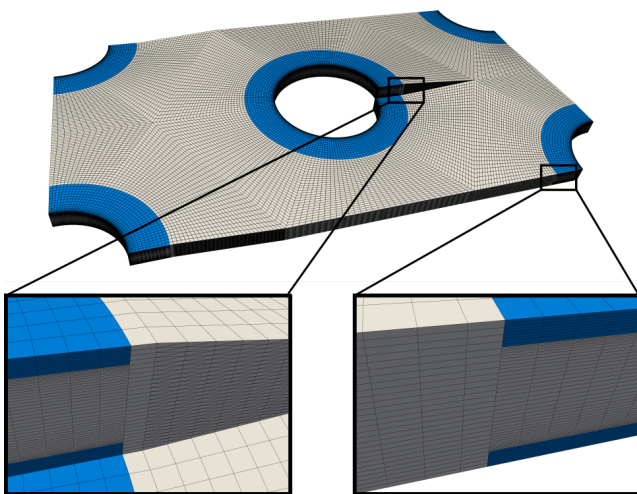


Fig. 4. Computational mesh used in the CFD simulations.

Eu is calculated directly by using Eq. (8), with  $\Delta p = P_l \cdot \partial p / \partial x$ . To calculate Nu, the temperature driving force  $\Delta T$  in Eq. (1) needs to be extracted from the CFD simulations. This is taken to be the difference between the bulk gas temperature and the wall temperature

$$\Delta T = T_b - T_w, \quad (16)$$

where the bulk temperature is defined as

$$T_b = \overline{T_{front}} + \overline{T_{back}}, \quad (17)$$

where  $T_{front}$  and  $T_{back}$  are the temperature fields at the planes located at  $-P_l/2$  and  $+P_l/2$  relative to the center tube, respectively.

The heat flux and fin temperatures are sampled at the center tube, and Nu is then calculated using the definition in Eq. (5).

#### 4.5 Thermophysical properties

The thermal and transport properties are assumed to be constant for both the gas and fins. They are summarised in Table 3. The fin region properties correspond to that of carbon SS-304 steel, while the gas phase is modelled as air at 310 K. These conditions correspond to the experimental conditions in which most correlations are developed.

Table 3. Constant thermophysical properties used in the CFD simulations.

Property	Gas	Fins
Density, $\rho$ (kg/m <sup>3</sup> )	1.1614	7950
Specific Heat Capacity, $C_p$ (J/kg·K)	1007	520
Thermal Conductivity, $\kappa$ (W/m·K)	0.0263	15.5
Dynamic Viscosity, (Pa·s)	$1.8455 \times 10^{-5}$	–

#### 4.6 Grid refinement study

To assess the grid sensitivity of the solution, a grid refinement study was performed. The simulations were run with

$Re = 7\,500$  using four different mesh resolutions, where the wall cell sizes were equal for all meshes, ensuring the same  $y^+$  values for each simulation. An initial steady-state simulation was conducted using the coarsest mesh, and its final time step was used as the initial condition for all simulations in the grid refinement study. The simulations were run until a quasi-steady state was achieved, where the time-averaged values of  $Eu$  and  $Nu$  did not change significantly. The resulting values are presented in Fig. 5. A mesh resolution of 730 000 cells was deemed sufficient, as both  $Eu$  and  $Nu$  are within 1% of the values at the finest grid (1 850 000 cells).

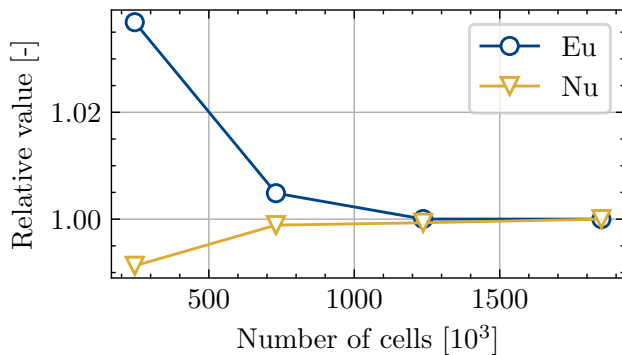


Fig. 5. Results from grid refinement study.  $Eu$  and  $Nu$  are normalised by the value at the finest grid.

## 5. RESULTS AND DISCUSSION

### 5.1 Results from optimisation

The main results from the optimisation are shown in Table 4 and are an indication of the different models ability to extrapolate. The optimisation is run twice. The first time the number of tubes per row is treated as a decimal number, providing a continuous flow area. As a second run, its value is a fixed integer value while the other three parameters, tube length, fin height, and fin pitch are re-optimised with a slight minimum weight increase as the penalty. It was also observed that all the models were constrained by the minimum tube pitch of 3 times the outer tube diameter, so a fixed triangular tube pitch of 76.2 mm was specified in the 2<sup>nd</sup> optimisation run.

As seen from Table 4, the obtained minimum weights range from 84 to 124 tonnes. The tendency is that the models obtaining the lowest weight seem to favour very low fin-height (3-7.5 mm) and pitch below 4 mm while the others are generally in the range of 7-15 mm and 4-8 mm. In a practical situation, the fin height and pitch cannot be varied freely meaning that the spread in the weight could be larger depending on the underlying model. In an OTSG, the heat transfer on the outside is limiting, so therefore fins are used to increase the surface area. However, when optimising for a specific duty and minimising the total weight, the weight contribution from the fins seems to be significant, and thus the optimisation routine finds the alternative solutions as discussed above.

The optimal tube lengths generally are between 5 and 6 m with the number of tubes per row around 50. These two parameters make up the exhaust cross-flow area and

Table 4. The result for the free geometry variables and the objective function for the different models

Model	Tube length [m]	Tubes per row [-]	Fin height [mm]	Fin pitch [mm]	Minimum Weight [Ton]
Briggs	6.29	56	10.1	5.6	124.7
Holfeld	5.91	50	10.5	4.4	114.6
Lindqvist	6.84	52	14.1	9.7	109.6
Rabas	5.71	50	14.5	8.1	102.6
PFR	5.32	42	3.5	1.9	99.3
ESCOA	5.83	48	7.6	4.0	97.7
Mon	5.47	51	3.0	1.5	97.3
ESDU	5.41	49	6.2	3.5	91.5
Stasiulevicius	4.87	50	6.5	3.7	84.1

determine the size of the enclosing casing which accounts for about half the total weight

Each of the correlations from Table 1 has also been used to simulate the performance of the optimised geometry from the ESCOA models. In Fig. 6 the comparison between the predictions is shown.

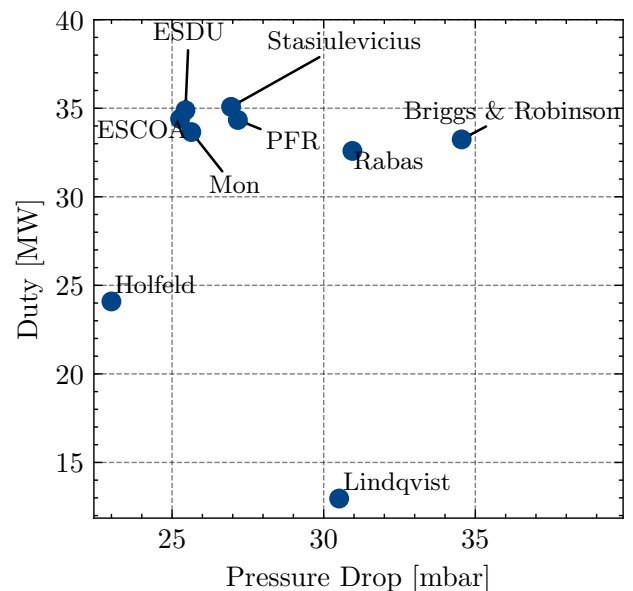


Fig. 6. Resulting duty and pressure losses for the simulation of the fixed geometry.

Some of the models predict the duty and pressure loss close to 34.4 MW and 25 mbar. Outliers seem to either predict the duty with too high pressure-loss. The correlation from Holfeld (2016) seems to underpredict both the duty and the pressure loss although this was developed based on experimental data that also included small-diameter tubes. A similar trend can be observed from Lindqvist and Næss (2018) that also were developed with compact geometries in mind.

### 5.2 Results from CFD simulations

Figure 7 presents the streamlines sampled during transient simulations at  $Re = 7500$ , coloured according to the normalised temperature. The figure also displays the temperature field of the fins. The streamlines distinctly illustrate vortex shedding occurring behind the tubes. This

vortex shedding is likely the primary cause of the significant oscillations observed in  $Eu$  and  $Nu$ , as depicted in Fig. 3. These oscillations were also observed after a certain number of iterations in the preliminary steady-state simulations, indicating that the steady-state assumption is un-physical. These oscillations are indicative of highly transient flow and heat transfer phenomena taking place within the OTSG, suggesting that steady-state simulations are not sufficient to model these flows accurately.

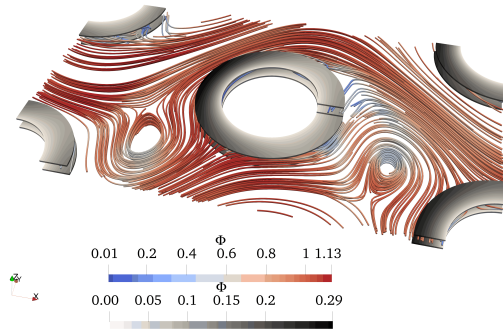


Fig. 7. Velocity streamlines and fins coloured by normalised temperature, for  $Re = 7\,500$ . Here  $\Phi = (T - T_w)/(T_{in,target} - T_w)$  is the normalised temperature.

Figure 8 shows that the correlation from Rabas et al. (1981) is closest to the CFD results for the  $Nu$ , with an average deviation of 6%, while the model from Holfeld (2016) seem show best agreement for  $Eu$ , with an average deviation of 16%. In general, most of the empirical correlations seem to overestimate the  $Eu$  and thus also predict a higher pressure loss, with the correlation by Lindqvist being the most extreme, deviating by 120%. This correlation also has the largest deviation for  $Nu$ , with 53%. In Fig. 6 the Holfeld-correlation shows the lowest pressure loss of all the models for the optimized geometry from using the ESCOA correlation. We note from looking at the streamlines in Fig. 7 that the geometry is quite "open" with regions with considerable turbulence and backflow. One might suspect that some local "pressure recovery" is not captured by any of the correlations. On the other hand, the correlation by Holfeld was designed with a bias toward compact geometries.

An additional optimisation, using the Stasiulevicius et al. (1988) correlation for the  $Nu$  and the Holfeld (2016) correlation for the  $Eu$  was performed.

The result from using these two models found a design with a very low fin pitch (1.07 mm) and with very low fins (2.2 mm). The resulting weight was 76.5 tons, about 20 tons less than the "reference" weight from the use of the ESCOA correlations. This design resembles a tube-bundle with low-fin tubes and this is probably not to be trusted to extrapolate a correlation developed for individually finned tubes to such type of low-fins. The more constrained the optimisation problem is, the more the minimum weight will increase. So, where weight is important, like in offshore installations, more experimental data or a systematic approach with CFD and machine learning would be useful to increase confidence in the design result.

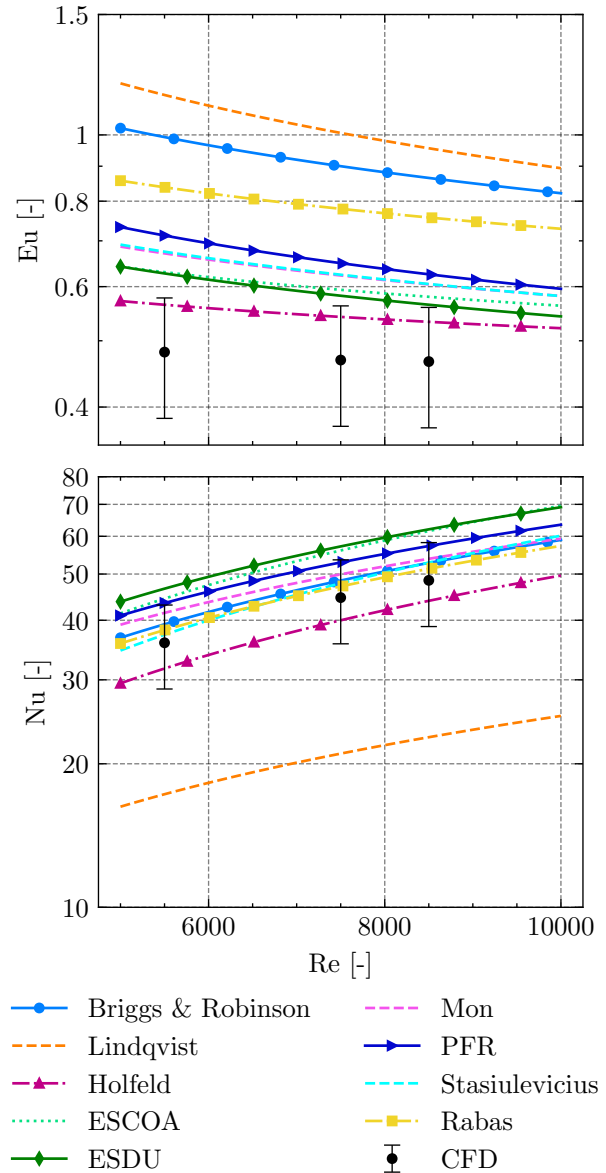


Fig. 8. Results from CFD simulations compared with empirical correlations. The bars on the CFD markers denote  $\pm 20\%$ .

### 6. CONCLUSIONS

In this work, we performed design optimisation of an offshore Once Through Steam Generator (OTSG) using nine different correlations to predict pressure loss and heat transfer. A reference optimised design was selected, and a validated CFD model was used to simulate pressure loss and heat transfer. The simulation results were compared against the correlations. The following conclusions were drawn:

- There is a significant variation in optimised designs, depending on the selected correlations.
- CFD results indicate transient behaviour in the selected configuration.
- A comparison with CFD data reveals significant deviations in pressure loss predictions for several correlations.



- The "best" models appear to extrapolate excessively.
- Additional experimental data and/or a systematic CFD approach to developing new models would be beneficial.

Without a CFD analysis of the optimal geometries found for the typical outliers from Fig. 5, (Holfeld, Lindqvist, Rabas and Briggs and Young) for a cross-check, it is difficult to conclude and recommend which of the models that is suitable for designing a compact OTSG.

## REFERENCES

- Bennett, D.L. and Chen, J.C. (1980). Forced convective boiling in vertical tubes for saturated pure components and binary mixtures. *AIChE Journal*, 26(3), 454–461. doi:10.1002/aic.690260317. URL <http://doi.wiley.com/10.1002/aic.690260317>.
- Briggs, D.E. and Young, E.H. (1963). Convection heat transfer and pressure drop of air flowing across triangular pitch banks of finned tubes. In *Chem. Eng. Prog. Symp. Ser.*, volume 59, 1–10.
- Deng, H., Skaugen, G., Næss, E., Zhang, M., and Øiseth, O.A. (2021). A novel methodology for design optimization of heat recovery steam generators with flow-induced vibration analysis. *Energy*, 226, 120325. doi:10.1016/j.energy.2021.120325.
- Espelund, J.R., Meyer, O.H.H., and Skaugen, G. (2022). Numerical modelling of fin side heat transfer and pressure loss for compact heat recovery steam generators. 263–270. doi:10.3384/ecp192037. URL <https://ecp.ep.liu.se/index.php/sims/article/view/521>.
- Espelund, J.R. (2022). Numerical modelling of fin side heat transfer and pressure loss for compact heat recovery steam generators. URL <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3016714>. Accepted: 2022-09-08T17:19:43Z.
- Friedel, L. (1979). Improved Friction Pressure Drop Correlations for Horizontal and Vertical Two-Phase Pipe Flow. In *European Two-Phase Flow Group Meeting*. Ispra, Italy.
- Ganapathy, V. (2003). Heat Transfer Equipment Design and Performance. In *Industrial Boilers*, 9, 1689–1699. doi:10.1017/CBO9781107415324.004.
- Hewitt, G. (1998). *Heat Exchanger Design Handbook*, 1998. Number pts. 1-2 in Heat Exchanger Design Handbook, 1998. Begell House.
- Holfeld, A. (2016). *Experimental investigation of heat transfer and pressure drop in compact waste heat recovery units*. Doctoral thesis, NTNU.
- Lindqvist, K. (2019). *Modeling of Flow and Heat Transfer in Karl Lindqvist Computational Fluid Dynamics Modeling of Flow and Heat Transfer in Fin-Tube Bundles*. Doctoral thesis, NTNU.
- Lindqvist, K. and Næss, E. (2018). A validated CFD model of plain and serrated fin-tube bundles. *Applied Thermal Engineering*, 143, 72–79. doi:10.1016/j.applthermaleng.2018.07.060. URL <https://www.sciencedirect.com/science/article/pii/S1359431117359598>.
- Mazzetti, M.J., Hagen, B.A.L., Skaugen, G., Lindqvist, K., Lundberg, S., and Kristensen, O.A. (2021). Achieving 50% weight reduction of offshore steam bottoming cycles. *Energy*, 230, 120634. doi:10.1016/j.energy.2021.120634. URL <https://www.sciencedirect.com/science/article/pii/S0360544221008835>.
- Mon, M.S. (2003). Numerical investigation of air-side heat transfer and pressure drop in circular finned-tube heat exchangers. 151.
- Montañés, R.M., Hagen, B., Deng, H., Skaugen, G., Morin, N., Andersen, M., and J. Mazzetti, M. (2023). Design optimization of compact gas turbine and steam combined cycles for combined heat and power production in a FPSO system—A case study. *Energy*, 282, 128401. doi:10.1016/j.energy.2023.128401.
- Patankar, S.V., Liu, C.H., and Sparrow, E.M. (1977). Fully Developed Flow and Heat Transfer in Ducts Having Streamwise-Periodic Variations of Cross-Sectional Area. *Journal of Heat Transfer*, 99(2), 180–186. doi:10.1115/1.3450666. URL <https://asmedigitalcollection.asme.org/heattransfer/article/99/2/180/382469/Fully-Developed-Flow-and-Heat-Transfer-in-Ducts>.
- Rabas, T., Eckels, P., and Sabatino, R. (1981). The effect of fin density on the heat transfer and pressure drop performance of low-finned tube banks. *Chemical Engineering Communications*, 10(1-3), 127–147. doi:10.1080/00986448108910930.
- Robinson, K.K. and Briggs, D.E. (1966). Pressure drop of air flowing across triangular pitch banks of finned tubes. In *Chem. Eng. Prog. Symp. Ser.*, volume 62, 177–184.
- Rosenman, T. (1976). Heat transfer and pressure drop characteristics of dry tower extended surfaces. part i. heat transfer and pressure drop data. doi:10.2172/7334913. URL <https://www.osti.gov/biblio/7334913>.
- Schittkowski, K. (1986). NLPQL: A fortran subroutine solving constrained nonlinear programming problems. *Annals of Operations Research*, 5(2), 485–500. doi:10.1007/BF02022087.
- Själänder, M., Jahre, M., Tufte, G., and Reissmann, N. (2021). EPIC: An Energy-Efficient, High-Performance GPGPU Computing Research Infrastructure. *arXiv:1912.05848 [cs]*. ArXiv: 1912.05848.
- Skaugen, G., Kolsaker, K., Walnum, H.T., and Wilhelmssen, (2013). A flexible and robust modelling framework for multi-stream heat exchangers. 49, 95–104. doi:10.1016/j.compchemeng.2012.10.006. URL <https://linkinghub.elsevier.com/retrieve/pii/S0098135412003067>.
- Spalart, P. and Allmaras, S. (1992). A one-equation turbulence model for aerodynamic flows. In *30th Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, Reno,NV,U.S.A. doi:10.2514/6.1992-439. URL <https://arc.aiaa.org/doi/10.2514/6.1992-439>.
- Stasiulevicius, J., Skrinska, A., Zhukauskas, A., and Hewitt, G.F. (1988). *Heat transfer of finned tube bundles in crossflow*. Hemisphere Pub. Corp.

## Simulation of Ammonia Cracker Process with Aspen HYSYS

Per Morten Hansen<sup>1</sup>, Chinthaka Attanayake, Vahid Farrokhi, Mohsen Gholizadefalah, Ladan Samaei, Zahra Sanidanesh, Lars Erik Øi<sup>1</sup>

University of South-Eastern Norway, N-3901 Porsgrunn, Norway

(Tel: +47 97 68 55 39; e-mail: [perha@usn.no](mailto:perha@usn.no)).

(e-mail: [lars.oi@usn.no](mailto:lars.oi@usn.no))

---

**Abstract:** This paper presents simulations of an ammonia cracker process using Aspen HYSYS. Ammonia is identified as both a promising low-emission maritime fuel and an energy carrier. This study focuses on converting ammonia to hydrogen through an ammonia cracker process. In the literature, there are found simulations of similar processes, but not much about optimization of the ammonia cracker process. A centralized ammonia cracking process was designed using the Peng-Robinson fluid package and Gibbs reactor in Aspen HYSYS. Gibbs reactors were employed to simulate both the cracker and the furnace (ammonia combustion reaction). Simplified assumptions included using a 100 % efficient splitter instead of a pressure swing adsorber. The ammonia feed had a molar flow rate of 500 kmole/h. The simulations included a base case scenario and an improved case for energy optimization. The base case scenario resulted in a total production of 0.13 kg of hydrogen per kg of ammonia feed. The improved case resulted in a production of 0.14 kg hydrogen. This was due to using the energy content present in the hydrogen and nitrogen product streams for warming up the ammonia before entering the cracker. This work demonstrates that Aspen HYSYS is a useful tool for optimizing the energy efficiency of an ammonia cracker process.

**Keywords:** Ammonia cracking, hydrogen, Aspen HYSYS, Aspen Plus, simulation, energy optimization

---

### 1. INTRODUCTION

Ammonia is a zero-emission energy carrier that could play a vital role in the shift towards more sustainable energy systems. A promising option is to transport the energy carrier as ammonia by ship and then transfer it to another ship equipped with an onboard cracker unit. The ammonia could then be converted into hydrogen in the cracker process before it is transported to the onshore (and offshore) end-users. Plans are known for building a maritime pilot/demonstration facility for ammonia cracking within the next couple of years (Wärtsilä, 2023). However, there is a need for additional research and development to obtain the level of knowledge needed to materialize the idea and build a large-scale facility. Simulation of the ammonia cracker process is a key step in the development to establish the mass and energy balance of the system and improve the process parameters with respect to energy usage and cost. This paper presents results from ammonia cracking simulations using Aspen HYSYS V12.

### 2. LITERATURE

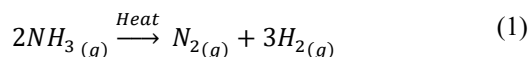
Thermal and catalytic cracking of ammonia are mature technologies that involve the controlled decomposition of ammonia into nitrogen and hydrogen, by applying heat and/or suitable catalysts. The interest in the use of ammonia as an energy carrier has increased in recent years which is reflected in an increasing number of scientific publications. Machaj et al. (2022) published a review paper on the use of ammonia in the maritime sector highlighting that the price of green ammonia is expected to drop significantly by 2050. Mallouppas, Ioannou and Yfantis (2023) examined key

barriers to the use of green ammonia as an alternative fuel in the maritime industry. The barriers included high production costs, availability, the challenge of ramping up current ammonia production and the development of ammonia-specific regulations (Mallouppas et al., 2023). Ashcroft and Goddin (2023) published a technical review of the ammonia cracking process, comparing centralized and localized hydrogen production by ammonia decomposition. Minimizing capital and operations costs are important aspects to consider in designing industrial solutions for ammonia cracking. Optimizing the recovery of waste streams containing ammonia and hydrogen could contribute to higher energy efficiency (Ashcroft and Goddin, 2023). Hansson et al. (2023) published an article on energy systems modeling and multi-criteria decision analysis to examine the potential role of ammonia as a marine fuel. They concluded that while the use of ammonia is promising, there are still unresolved issues that need to be addressed before it can be introduced on a large scale as a maritime fuel.

Ammonia cracking has been given increasing interest in clean energy production and industrial chemistry. It offers a sustainable path to produce high-purity hydrogen, a crucial resource for fuel cells, and as a heat source (Speight, 2023).

Regardless of the scale of hydrogen production, ammonia is cracked by applying heat, typically through the combustion of a fuel or an energy source such as electricity. When ammonia is thermally decomposed, it produces a 1:3 molar ratio of nitrogen and hydrogen. See Equation 1.

The first time ammonia was cracked thermally was by Bruke in 1933. This process was conducted at temperatures between 550 - 600 °C and achieved a 90% conversion rate (Yousefi Rizi and Shin, 2022). Different reactor types for NH<sub>3</sub> cracking are discussed by Mukherjee et al. (2018).



**Table 1. Ammonia conversion at different temperatures in an ideal Gibbs reactor simulation (Chiuta et al., 2013).**

Temperature ( °C)	Ammonia conversion (%)
250	89.2
300	95.7
350	98.1
400	99.1
450	99.5
500	99.7
600	99.9
700	99.95

This section presents a brief review of the simulation studies of hydrogen production using the ammonia cracking process. A literature search showed that a majority of the process simulations were done in Aspen Plus. This preference is likely due to Aspen Plus being more flexible and therefore more suitable for these simulation tasks compared to the Aspen HYSYS software.

Chiuta et al. (2013) simulated the ammonia cracking reaction in HSC Chemistry 7 software through Gibbs free energy minimization. The results, presented in Table 1, demonstrate an increase in equilibrium conversion with rising temperature.

Cha et al. (2021) conducted a study on an efficient process for sustainable green hydrogen production from ammonia decomposition, using Aspen Plus for simulation. The Peng Robinson thermodynamic model was applied in this simulation. Pure NH<sub>3</sub> from liquid storage was released through a feed valve (10–2.5 bar) at the start of the process. The main heat sinks in the process included Ammonia heating, recycled ammonia desorption from the adsorbent material (by raising the temperature from 31 to 310 °C), and endothermic reaction heat from ammonia decomposition. The majority of the hydrogen was separated using a pressure swing adsorption (PSA) section, as the remaining products were primarily a 3:1 mixture of H<sub>2</sub>/N<sub>2</sub>. At fixed conditions with a pressure of 2.5 bar and a temperature of 20 °C, the results showed that 71 % of the hydrogen could be recovered (with a purity of 99 mol%). The process used approximately 200 kg of activated carbon

per kmole of the PSA inlet gas. The simulation results showed good agreement with the experimental results. (Cha et al., 2021).

Devkota et al. (2023) published a study on the process design and simulation of onsite hydrogen production from ammonia decomposition. The designed process was modeled using Aspen Plus, with Peng-Robinson equation of state employed to estimate the thermodynamic properties. The study utilized a feed rate of 4000 kg/h of pure ammonia, maintained at a temperature of 298 K and a pressure of 10 bar, as the input parameters for the calculations. Given the endothermic nature of the decomposition reaction, the necessary heat energy was generated by burning carbon-free ammonia, requiring about 9% of fresh feed. The resulting product stream contained a small amount of unreacted ammonia, which was subsequently separated and sent to the furnace after being mixed with fresh fuel and air.

The waste stream from this process, containing a large amount of hydrogen and nitrogen gas, was recycled to the furnace for thermal energy production. The study employed a steady-state model of a multi-catalytic packed bed reactor for ammonia decomposition, incorporating an intermediate heating system to enhance the reaction rate. Additionally, the ammonia gas was preheated in a fired furnace to achieve a decomposition temperature of 773 K before entering the reactor, utilizing the catalyst Ru/Al<sub>2</sub>O<sub>3</sub> for the decomposition process (Devkota et al., 2023).

Another study was published by Lee et al. (2023) on carbon-free hydrogen production using an induction heating-based reactor for ammonia decomposition, achieving a hydrogen production rate of 150 Nm<sup>3</sup>/h. This study was conducted both experimentally and through simulation. The process design included a reactor for ammonia decomposition, a pre-heater, an adsorption column, and a pressure swing adsorption (PSA) unit. To maintain a carbon-free process and produce green hydrogen, the external heat for ammonia decomposition was supplied using electricity from renewable energy sources. The hydrogen recovery rate of the PSA was fixed at 79 % and incorporated into the process simulation. Experimental results demonstrated that the induction heating reactor achieved an ammonia conversion exceeding 90 % at a temperature of 600 °C and a pressure of 7 barg. This conversion rate was subsequently used as an assumption for the conversion reactor model in the process simulation.

Restelli et al. (2023) conducted a comprehensive techno-economic analysis of green hydrogen production via ammonia decomposition. This research includes various hydrogen production processes, including a centralized ammonia cracking process where all the stored ammonia in the inlet is converted to hydrogen. The process was simulated using Aspen Plus (Restelli et al., 2023). The process flow diagram of this simulation is shown in Fig. 1. In this process, the Ammonia stream is pressurized to 30 bar and preheated through in a series of heat exchangers, before being directed to the cracking reactor, taking advantage of the high enthalpic content of the reaction products. The reactor simulation was conducted using the Gibbs module within Aspen Plus. Consequently, the conversion of ammonia aligns with thermodynamic equilibrium at



the reactor's operating conditions of 30 bar and 900 °C. These conditions are consistent with those typically used for commercially available nickel-based catalysts. To facilitate the cracking reaction, the necessary heat is generated by combusting a portion of the supplied ammonia, combined with waste streams possessing a high H<sub>2</sub> content that originate from the purification section. Air is employed as an oxidizer in slightly excess quantities to ensure complete combustion. Following the reaction stage, the separation of the hydrogen product from any unreacted ammonia and nitrogen is achieved through pressure swing adsorption (PSA).

### 3. METHODS

The central focus of this study is the simulation of the ammonia cracking process, employing Aspen HYSYS V12 with the Peng-Robinson fluid package. The ammonia combustion reaction is modelled as an equilibrium reaction, with equilibrium parameters determined through the minimization of Gibbs free energy. The simulation is based on the following key assumptions:

**Assumption 1.** To facilitate the combustion of ammonia, a Gibbs reactor (furnace) was defined. This was necessary since ammonia alone did not exhibit combustion reactions in the furnace.

**Assumption 2.** In the absence of catalyst data, the process employs the minimization of Gibbs energy for ammonia cracking. This approach assumes thermal cracking, relying on heat to break down ammonia into hydrogen.

**Assumption 3.** Instead of using adsorber and Pressure Swing Adsorber (PSA) units, a component splitter was employed. This approach assumes no temperature or pressure losses and assumes 100% efficiency in separating unreacted ammonia, hydrogen, and nitrogen in both separation units.

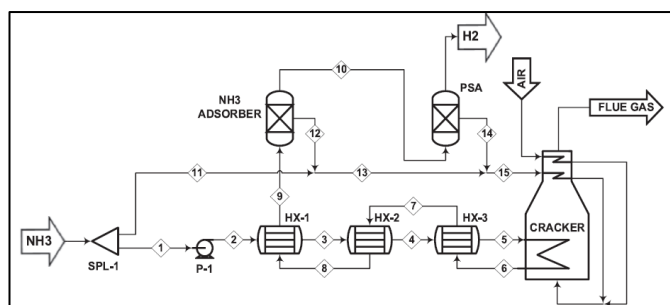


Fig. 1. Process flow diagram of centralized ammonia cracking (Restelli et al., 2023).

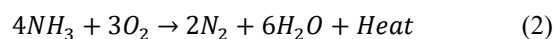
Finally, the study optimizes energy within the process by using waste heat from outlet streams (hydrogen and nitrogen) to preheat the ammonia stream heading to the cracker. To achieve this, two coolers are introduced to lower the temperatures of

the hydrogen and nitrogen streams to ambient levels. The combined heat from these streams is then employed to preheat the ammonia using a heater, facilitated by a recycling manipulator. This optimization not only improves ammonia pre-heating but also reduces ammonia consumption as fuel, thereby increasing the efficiency of the ammonia-cracking process.

Based on the literature review, the process design proposed by Restelli et al. (2023) was selected as the basis for the Aspen HYSYS simulation work presented in this article. Specifications for the streams are given in Table 2. The chemical reaction used in the main combustion reaction in this study is specified in (2).

Table 2. Inlet and outlet stream specifications.

Stream name / Parameter	NH <sub>3</sub>	H <sub>2</sub>	AIR	FLUE GAS
Temperature (°C)	-27.6	25.5	25.0	139
Pressure (bar)	1.3	30	1.01	1.01
Feed TOTAL (kmoles/h)	592.33	698.07	608.13	999.28
Mole Fractions				
H <sub>2</sub>	0	0.999	0	0.0003
N <sub>2</sub>	0	0.001	0.79	0.775
H <sub>2</sub> O	0	0	0	0.191
NO	0	0	0	0.0029
NH <sub>3</sub>	1	0	0	0
O <sub>2</sub>	0	0	0.21	0.0309



## 4. RESULTS AND DISCUSSION

### 4.1 Base case

Figure 2 depicts the Process Flow Diagram (PFD) of the base model for simulation of ammonia cracking. Simplifying assumptions were listed at the start of Section 3 (methods).

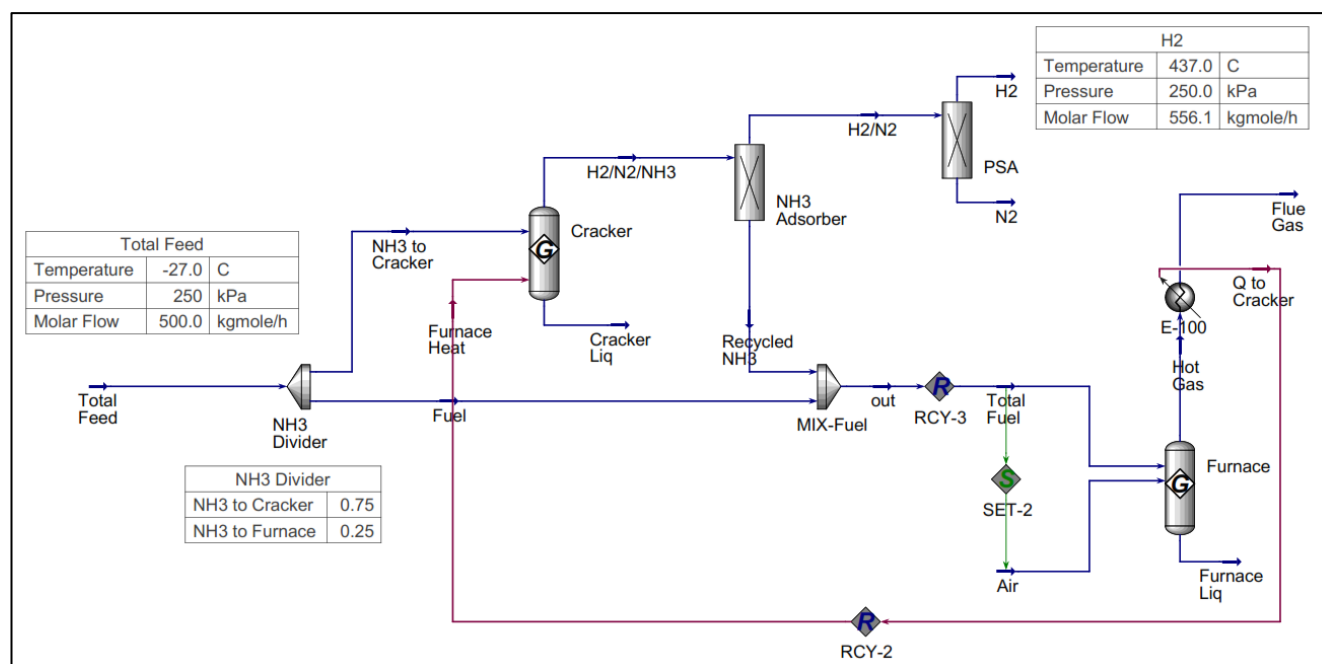


Fig. 2. Aspen HYSYS Process Flow Diagram (PFD) of ammonia cracking (Base case model).

Simulation results from the base case are presented in Table 3.

Table 3: Results from the process simulation - base case.

Species	Parameter	Unit	Value
NH <sub>3</sub>	Total Feed	(kmoles/h)	500
	Cracker Feed	(kmoles/h)	375
	Unreacted	(kmoles/h)	4.3
	Burned in Furnace	(kmoles/h)	129.3
	Cracker conversion	(%)	99.42
	Total conversion	(%)	74.14
	Cracker yield on H <sub>2</sub>	(%)	98.88
	Total yield on H <sub>2</sub>	(%)	74.14
H <sub>2</sub>	Total Production	(kmoles/h)	556.1
	Total H <sub>2</sub> to NH <sub>3</sub>	(kmoles / kmoles)	1.11
	Total H <sub>2</sub> to NH <sub>3</sub>	(kg/kg)	0.1316

#### 4.2 Improved case

In this section, the available heat streams within the process, specifically the heat from the hydrogen and nitrogen product streams, were utilized to increase the hydrogen production, the desired final product. To achieve this objective, a new case was introduced, referred to as the "energy consumption improved case" or simply the "improved case".

To optimize the energy of the process, the energy flow from the outlet streams, namely the produced hydrogen and nitrogen, was used to preheat the ammonia stream fed to the cracker. For this purpose, and simulation simplicity, instead of defining heat exchangers, two coolers were introduced to cool down the hydrogen and nitrogen streams to ambient temperature. These two heat flows were then combined and used to preheat the ammonia stream to the cracker using a heater. Furthermore, optimizing the energy in this manner aids the ammonia cracking process by heating the ammonia stream to the cracker, allowing for an increased fraction of ammonia to be directed to the cracker. Consequently, this reduces the ammonia flow used in the furnace. As a result, the optimization not only involves utilizing waste heat to preheat the ammonia to the cracker but also results in burning less ammonia to supply heat for cracking. The flowsheet of the improved case is illustrated in Fig. 3.

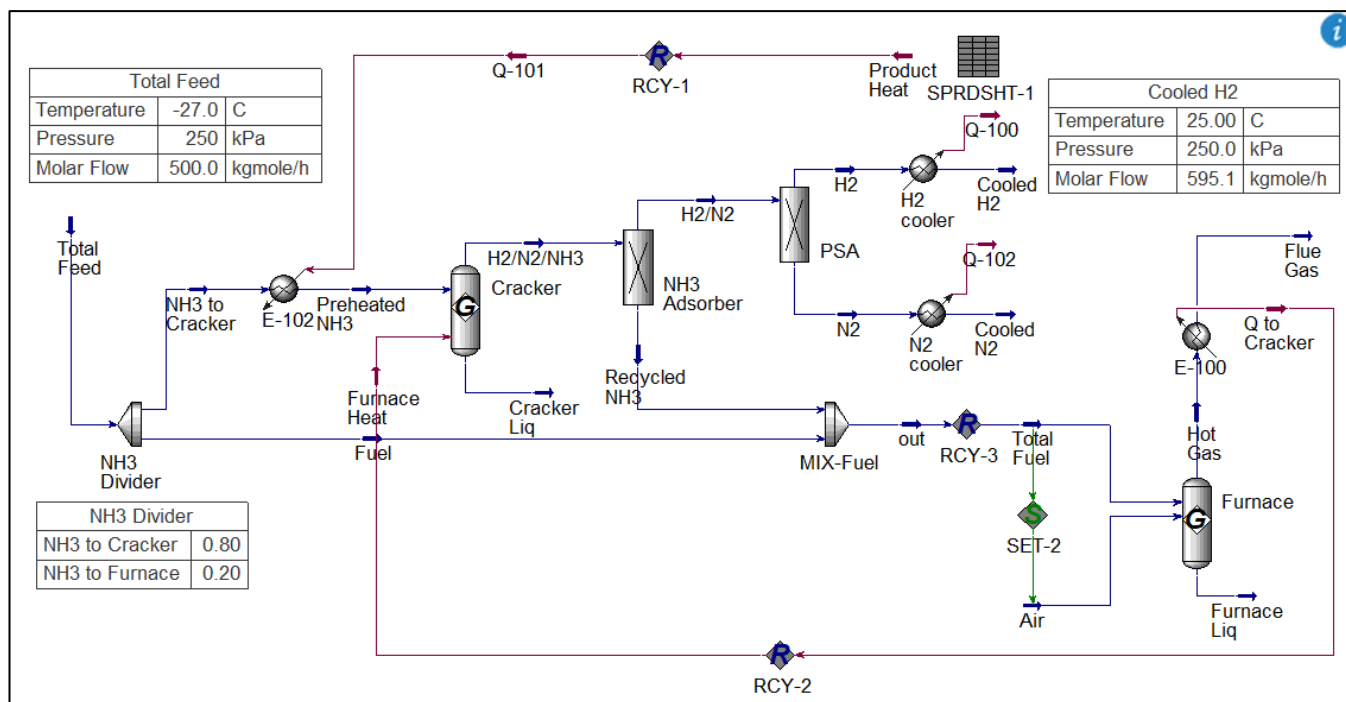


Fig. 3. Aspen HYSYS Flowsheet of the improved case.

The simulation results of the improved case are presented in Table 4.

Table 4. Results from the process simulation - improved case.

Species	Parameter	Unit	Value
NH <sub>3</sub>	Total Feed	(kmoles/h)	500
	Cracker Feed	(kmoles/h)	400
	Unreacted	(kmoles/h)	3.26
	Burned in Furnace	(kmoles/h)	103.3
	Cracker conversion	(%)	99.6
	Total conversion	(%)	74.35
	Cracker yield on H <sub>2</sub>	(%)	99.59
	Total yield on H <sub>2</sub>	(%)	79.35
H <sub>2</sub>	Total Production	(kmoles/h)	595.1
	Total H <sub>2</sub> to NH <sub>3</sub>	(kmoles / kmoles)	1.19
	Total H <sub>2</sub> to NH <sub>3</sub>	(kg/kg)	0.141

#### 4.3 Discussions

The Gibbs reactor, which cracks the ammonia operates without any kinetic model and provides a similar conversion to the studies done by Chiuta et al. (2013) and Ojelade and Zaman (2021). This study has a 99.4% conversion at a cracker temperature of 437 °C in the base case and a 99.6% conversion at a cracker temperature of 466 °C in the energy-optimized case. This reflects on previous studies mentioned in the literature review. In Table 5, the obtained hydrogen production

in the simulations is compared with literature values. It shows that the simulations in this work are in the same order of magnitude compared to earlier simulations. It also shows that 8 % increased hydrogen production from 0.131 to 0.141 is obtained by process improvements.

Simplifying assumptions have been made in the simulations in this work. Assumption 1 and 2 are that Gibbs equilibrium reactors are used to simulate the ammonia combustion and ammonia cracking reactions. This is also done in literature. These assumptions are optimistic, and more realistic simulations would result in lower cracker conversion and lower hydrogen production. Assumption 3 is that an ideal component splitter is used to simulate the PSA unit. This is also optimistic, and this tends to overestimate the hydrogen production.

Table 5. Results - comparison with previous studies.

Sources	hydrogen production (kg)/ ammonia feed (kg)
This study - base case	0.131
This study - improved case	0.141
Restelli et al. (2023)	0.140
Lee et al. (2023)	0.127
Devkota et al. (2023)	0.129

Heat exchange in the process is simulated with ideal heaters and cooling, and it is assumed no heat loss. These assumptions are assumed to be negligible because the heat loss is assumed to be much smaller than the heat of reactions. To make a comparison of these two fluid packages, the base case

simulation was executed using both Peng Robinson and Soave-Redlich-Kwong (SRK). There was not a significant difference in the results of the two equation of state models.

In this work, a portion of ammonia was used as fuel to provide energy for the cracking reactor. This has been done in a Gibbs reactor instead of a furnace. Suggestions for further work include using a furnace, not an ideal Gibbs reactor for burning ammonia. Aspen HYSYS has a limitation of using ammonia as a fuel in the furnace. A potential for energy consumption is identified that can be investigated in future work. Also, energy optimization based on Pinch Technology can be conducted to improve the energy efficiency of the process.

Additionally, doing a techno-economic analysis of the hydrogen production from the ammonia cracking process can be valuable. As an example, this economic investigation can also consider the economic analysis of the energy resource for the cracker.

## 5. CONCLUSIONS

The paper showed results from simulations of an ammonia cracking process to produce hydrogen using Aspen HYSYS. Ammonia represents a promising way of transporting hydrogen over long distances. A base case that replicated previous work from literature was improved with respect to energy consumption. The improved case gave 0.141 kg hydrogen per 1 kg ammonia feed, marking an 8 % improvement. Recommendations for further work include using available experimental kinetic data and replacing the Gibbs reactor in Aspen HYSYS. Additionally, heat exchangers could be used instead of heaters and coolers to enable energy optimization analysis, perform equipment sizing, and conduct techno-economic evaluations.

## REFERENCES

- Ashcroft, J. and Goddin, H. (2022). Centralised and Localised Hydrogen Generation by Ammonia Decomposition: A technical review of the ammonia cracking process, *Johnson Matthey Technology Review*, 66, 375–385. doi: 10.1595/205651322X16554704236047
- Cha, J., Park, Y., Brigljević, B., Lee, B., Lim, D., Lee, T., Jeong, H., Kim, Y., Sohn, H., Mikulčić, H., Lee, K.M., Nam, D.H., Lee, K. B., Lim, H., Yoon, C. W. and Jo, Y.S. (2021). An efficient process for sustainable and scaleable hydrogen production from ammonia. *Renew. Sustain. Energy Rev.* 152, 111562.
- Chiuta, S., Everson, R.C., Neomagus, H.W.J.P., van der Gryp, P. and Bessarabov, D.G., 2013. Reactor technology options for distributed hydrogen generation via ammonia decomposition: A review. *International Journal of Hydrogen Energy*, 38, 14968–14991. doi: 10.1016/j.ijhydene.2013.09.067
- Devkota, S., Shin, B.-J., Mun, J.-H., Kang, T.-H., Yoon, H.C., Mazari, S.A. and Moon, J.-H. (2023). Process design and optimization of onsite hydrogen production from ammonia: Reactor design, energy saving and NOX control. *Fuel*, 342, 127879. doi:10.1016/j.fuel.2023.127879
- Hansson, J., Brynolf, S., Fridell, E. and Lehtveer, M. (2020). The Potential Role of Ammonia as Marine Fuel—Based on Energy Systems Modeling and Multi-Criteria Decision Analysis. *Sustainability* 12, 3265. doi: 10.3390/su12083265
- Lee, J., Ga, S., Lim, D., Lee, S., Cho, H. and Kim, J. (2023). Carbon-free green hydrogen production process with induction heating-based ammonia decomposition reactor. *Chemical Engineering Journal*, 457, 141203. doi: 10.1016/j.cej.2022.141203
- Machaj, K., Kupecki, J., Malecha, Z., Morawski, A.W., Skrzypekiewicz, M., Stanclik, M. and Chorowski, M. (2022). Ammonia as a potential marine fuel: A review. *Energy Strategy Reviews*, 44, 100926. doi: 10.1016/j.esr.2022.100926
- Mallouppas, G., Ioannou, C. and Yfantis, E.A. (2022). A Review of the Latest Trends in the Use of Green Ammonia as an Energy Carrier in Maritime Industry. *Energies*, 15, 1453. doi: 10.3390/en15041453
- Mukherjee, S., Devaguptapu, S.V., Sviripa, A., Lund, C.R.F. and Wu, G. (2018). Low-temperature ammonia decomposition catalysts for hydrogen generation. *Applied Catalysis B: Environmental*. 226, 162–181. doi: 10.1016/j.apcatb.2017.12.039
- Ojelade, O.A. and Zaman, S.F. (2021). Ammonia decomposition for hydrogen production: a thermodynamic study. *Chemical Papers*. Vol 75, pp 57–65. doi: 10.1007/s11696-020-01278-z
- Restelli, F., Spatolisano, E., Pellegrini, L.A., de Angelis, A.R., Cattaneo, S. and Roccaro, E. (2024). Detailed techno-economic assessment of ammonia as green H<sub>2</sub> carrier. *International Journal of Hydrogen Energy*, 52, 532–547. doi: 10.1016/j.ijhydene.2023.06.206
- Speight, J.G. (2024). *The Chemistry and Technology of Petroleum* [WWW Document]. Routledge & CRC Press. URL <https://www.routledge.com/The-Chemistry-and-Technology-of-Petroleum/Speight/p/book/9781439873892> (accessed 5.30.24).
- Yousefi R., H.A. and Shin, D. (2022). Green Hydrogen Production Technologies from Ammonia Cracking. *Energies* 15, 8246. doi: 10.3390/en15218246
- Wärtsilä (2023). Wärtsilä, Höegh LNG and partners receive funding of 5.9 million euros for major clean energy initiative. URL <https://www.wartsila.com/media/news/13-04-2023-wartsila-hoegh-lng-and-partners-receive-funding-of-5-9-million-euros-for-major-clean-energy-initiative-3253477> (accessed 6.1.24).

# Steady State and Transient Modelling of A Three-Core Once-Through Steam Generator

Håvard Falch, Geir Skaugen

*SINTEF Energy Research, Sem Sælands vei 11, Trondheim, Norway  
Corresponding author: haavard.falch@sintef.no*

---

**Abstract:** To reduce emissions and save fuel in offshore power production using gas turbines, one can use the gas turbine exhaust as a heat source for a bottoming cycle for heat and power production. This can replace about one in four gas turbines. In offshore applications weight and size become more important and thus a once-through steam generator (OTSG) is a way to achieve low weight for the bottoming cycle. To reduce the size and weight of the OTSG further, one can reduce the tube diameter in the tube bundles. In this work a three-core OTSG, representing the economizer, evaporator, and superheater, was modelled and the design optimized to achieve minimum weight, while producing a certain amount of power and keeping within constraints of flue gas and steam pressure losses. This was done for varying tube diameters in each of the cores, in steady state. Afterwards transient simulations were performed for each optimized design to find their response times to a step change in the gas turbine load. The evaporator has the biggest impact on both the weight and the response time, while the superheater and economizer had similar and smaller impacts on both the weight and response time.

*Keywords:* Heat exchanger, OTSG, power production, steam production, optimization, transient modelling

---

## 1. INTRODUCTION

In offshore oil and gas production a large source of CO<sub>2</sub> emissions is gas turbine exhaust, and in 2023 it accounted for 80% of CO<sub>2</sub> emissions from Norwegian oil and gas production according to the Norwegian Petroleum Directorate (2023). One opportunity for reducing these emissions is to install a steam bottoming cycle to use the hot exhaust gas as a heat source. By installing a bottoming cycle, the fuel consumption of a gas turbine can be reduced by about 25% (Nord and Bolland, 2012; Mazzetti et al., 2014). For a processing facility with a fleet of several gas turbines, installing a bottoming cycle will mean that one or more of the gas turbines can be removed. In the bottoming cycle, illustrated in Fig. 1, pressurized water is heated, evaporated, and superheated through a heat recovery steam generator (HRSG) and then expanded in a steam turbine to generate power. The low-pressure steam is condensed and cooled by sea water before being pumped back to the HRSG. Other non-conventional fluids have also been studied for use in bottoming cycles, including air by Pierobon and Haglind (2014), CO<sub>2</sub> by Walnum et al. (2013) and Skaugen et al. (2014), and organic Rankine cycles by Pierobon et al. (2014) and Motamed and Nord (2022). However, only steam bottoming cycles have been installed on the Norwegian Continental Shelf, making it

the most mature technology. The HRSG is a large installation with a weight that can reach several hundred tons, so, for offshore installation a compact bottoming cycle will be necessary. This can be achieved by optimizing the HRSG. In offshore installations, the HRSG without a steam drum will normally be suggested and the excess heat is recovered in a once-through steam generator (OTSG). In (Mazzetti et al., 2021) an outline on how a compact steam bottoming cycle can be designed is discussed - where one of the main size/weight drivers was the tube diameter selection in the OTSG. Similar analysis was shown by Montañés et al. (2023) for a compact steam bottoming cycle for heat and power production on a floating production plant. Deng et al. (2021) studied vibrations of the tubes in an OTSG and found that vibrational constraints increased the optimal weight of the OTSG, and that the increase was larger for the single-core case compared to the three-core case.

In this work a three-core OTSG is studied in detail in order to optimize the tube diameters and circuiting in the three cores, the economizer, the evaporator, and the superheater, individually, in order to find the minimum weight for a specified duty and with restriction on pressure losses. In earlier works optimizing OTSG designs, the tube diameter has been used as an optimization parameter in the single-core case and fixed in the three-core case. However, in this work, we investigate four different industry standard tube diameters to get a thorough understanding of how each part of the OTSG is affected by changing the diameter of the tube and if having different diameters

\* This publication has been produced with support from the LowEmission Research Centre ([www.lowemission.no](http://www.lowemission.no)), performed under the Norwegian research program PETROSENTER. The authors acknowledge the industry partners in LowEmission for their contributions and the Research Council of Norway (296207).



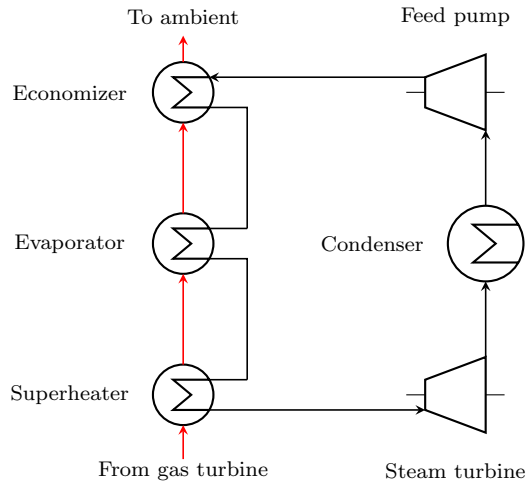


Fig. 1. Heat recovery steam generation from gas turbine exhaust in a bottoming cycle for power production.

in each of the cores will be optimal. This is done by optimizing each core with respect to weight and afterwards looking at the response time for a load change.

## 2. NUMERICAL METHODS

To optimize the geometry of the OTSG and perform dynamic simulations we used an in-house modeling tool developed by SINTEF Energy Research (Skaugen et al., 2013). The OTSG geometry optimization was done in several steps. First, a thermodynamic optimization of the whole bottoming cycle was done to investigate the potential power production and find a suitable mass flow and pressure for the geometric optimization. After this, the geometric optimization of the three cores was performed in three steps. From the thermodynamic optimization, we knew the mass flow and the pressure of the steam at the superheater outlet which coincides with the exhaust inlet, both of which are at the bottom of the superheater. Thus, we began with the superheater optimization, and using the lowest weight design we used the flow conditions from the top of the superheater to optimize the evaporator and again using the flow conditions from the lightest evaporator design we optimized the economizer.

### 2.1 Thermodynamic optimization

The thermodynamic optimization considers the entire bottoming cycle including the OTSG at a flowsheet level, as seen in Fig. 1, however, it does not include the full OTSG geometry. To solve the model it takes both external and process variables. The external variables are the mass flow, pressure, and composition of the exhaust gas coming from the gas turbine and a pinch point temperature difference (PPTD), i.e. the minimum temperature difference between the exhaust gas and the water. The process variables are the outlet temperature of the exhaust gas, the water pressure at the pump outlet, the water temperature at the turbine inlet, the water pressure at the turbine outlet, and the temperature increase of the cooling water in the condenser. The heat transferred from the exhaust gas to the water  $Q$  is calculated as

$$Q = (h(p, T)_{\text{ex, in}} - h(p, T)_{\text{ex, out}}) \dot{m}_{\text{ex}}, \quad (1)$$

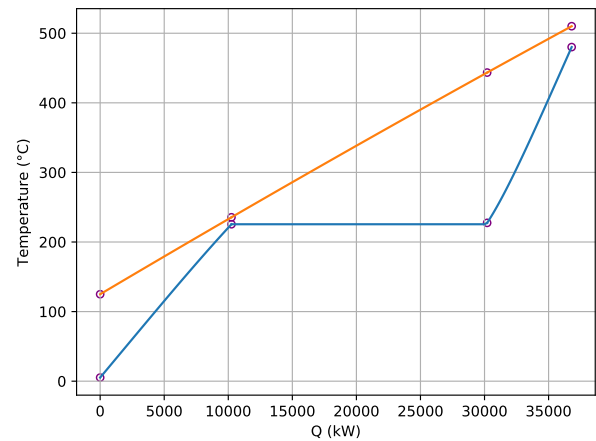


Fig. 2. Typical temperature profile in a counter flow heat exchanger as a function of the heat transferred along the flow direction of the cold fluid. The pinch point, where the temperature of the cold and the hot liquids are the closest, is here marked by the second set of dots and is  $10^{\circ}\text{C}$ .

where  $h$  is the enthalpy,  $p$  the pressure,  $T$  the temperature, and  $\dot{m}$  is the mass flow rate, while the subscript ex denotes the exhaust and the subscripts in and out refer to the inlet and outlet states of the OTSG. This heat transfer is then used to calculate the water flow rate  $\dot{m}_w$  as

$$\dot{m}_w = \frac{Q}{(h(p, T)_{w, \text{out}} - h(p, T)_{w, \text{in}})}. \quad (2)$$

The amount of power produced in the turbine expansion  $W_{\text{exp}}$  is calculated from an isentropic expansion with isentropic efficiency 0.85, while the pump work  $W_{\text{pump}}$  is calculated from an isentropic compression with efficiency 0.7. Together these give the net power produced by the bottoming cycle  $W_{\text{net}}$  as

$$W_{\text{net}} = W_{\text{exp}} - W_{\text{pump}}. \quad (3)$$

The optimization is performed using the gradient-based constrained optimization solver NLPQL by Schittkowski (1986) to optimize the process variables with the objective of maximizing the net power produced. The optimization variables are all given initial values as well as lower and upper bounds when passed to NLPQL. Instead of optimizing all the parameters one or more can also be given a fixed value. In addition, it is subject to several constraints, such as a minimum vapour fraction at the steam turbine outlet and minimum pinch temperature, which are all inequality constraints, and an equality constraint to ensure continuous pressure for the water. In our case, we fix the water temperature at the turbine inlet and then the pinch is at the onset of boiling of the water, and thus the PPTD decides the temperature difference between the exhaust and water at this point. This can be seen as the second set of points in Fig. 2 which shows standard temperature profiles of the exhaust (above) and water (below) as a function of heat transferred in the OTSG.

### 2.2 Geometric optimization

The geometry of the OTSG is described by three cores, the economizer, evaporator, and superheater, each consisting



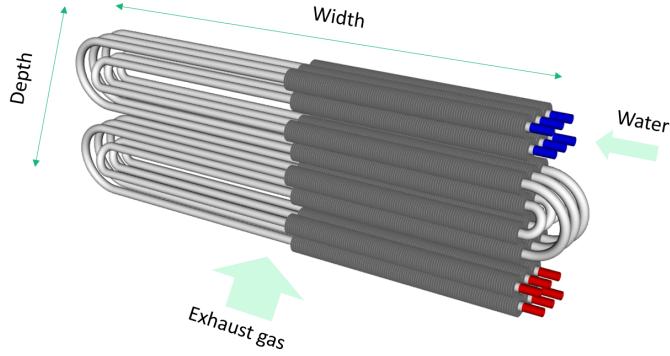


Fig. 3. A simple tube bundle, the fins are only shown for half of the width to get a better view of the tubes.

of a tube bundle along with support beams, as well as a surrounding duct for each tube bundle to contain the exhaust gas and an inlet duct for the superheater and an exit duct for the economizer. The duct size and thus weight is dependent on the size of the tube bundles. The weight of the entire core consists of the tube bundle weight, as well as the duct surrounding the bundle and the inlet or exit duct when applicable. A simple example tube bundle is shown in Fig. 3. We see that each tube passes through 4 times, i.e. 4 passes, there are 2 parallel rows of tubes in each pass, i.e. 2 rows per pass, and there are 3 tubes per row. The width of the tube bundle equals the length of the water flow in one row, the depth is the length of flow for the exhaust gas along the tube bundle. The tubes are laid out with a  $30^\circ$  offset between each row which gives a hexagonal pattern. In addition, the tubes have fins on the outside to improve the heat transfer from the exhaust gas. The fins have a set thickness, but the distance between them and their height can be varied.

To solve each of the cores we start by assuming that all tubes in one row are equal to get a 2D problem and then discretizing along the width and depth such that each tube is divided in several tube segments, or nodes, and the exhaust gas flows in several separate columns. This allows us to start from the outlet of the bottom tube and solve the heat transfer for each node. The heat transfer  $Q$  is given by

$$Q = UA\Delta T, \quad (4)$$

where  $U$  is the overall heat transfer coefficient,  $A$  is the heat transfer area and  $\Delta T = T_{\text{ex}} - T_{\text{w}}$  is the temperature difference between the exhaust gas and water. In the simplest case one can use the initial temperature difference as  $\Delta T$ . However, in the OTSG there is quite a big temperature change in each node and thus we instead use Heun's method to update the temperature difference with intermediate steps and thus get a better heat transfer calculation. The heat transferred is then used to update the fluid enthalpies as

$$h_{\text{out}} = h_{\text{in}} - \frac{Q}{\dot{m}}, \quad (5)$$

where both the exhaust and the water get a minus sign as we are calculating backwards along the water. Each node is solved iteratively until the end of the tube pass and then the tube above is solved in the same fashion and so on until one reaches the inlet. As the pressure and enthalpy of the water might vary for parallel passes, they are here

Table 1. Variables with initial guess and bounds for the geometric optimization.

Variable	Unit	Initial guess	Range
Tubes per row	[ ]	50	5-180
Transversal fin tip gap	[mm]	50	5-125
Fin height	[mm]	6.5	5-20
Fin pitch	[mm]	5.5	2-8

mixed to get a common outlet condition and the same is also done for the exhaust gas. The outlet conditions can then be used as inlet conditions for the next core.

Just as for the thermodynamic optimization we also use NLPQL for gradient-based constrained optimization of the core geometry. The objective of the optimization is to minimize the total weight of the core. The variables with initial values as well as their range are shown in Table 1. The constraints consist of a single equality constraint for the heat recovered as well as inequality constraints for pressure losses for both fluids, maximum exhaust velocity, and minimum fin and tube spacing. There are also a few variables that NLPQL does not optimize, namely the number of passes, the number of tubes per row, the width of the core, and the diameter of the tubes. These are given as inputs and the NLPQL optimization is run a separate time for each combination. The number of passes and tubes per row are both integer values and thus not suited for NLPQL to optimize. The width is also controlled separately to ensure that all three cores have the same width. Finally, the tube diameter is given certain fixed values based on industry-standard tubes.

After the thermodynamic optimization, we have the heat transferred in each core, the mass flow of the water, as well as the temperature and pressure at the steam turbine inlet and thus the superheater outlet. With these values as well as the mass flow, temperature, and pressure of the exhaust gas we can begin with optimizing the superheater. After finding the lightest superheater design, we can use the inlet conditions of the superheater as the outlet conditions of the evaporator and optimize it, and finally we can do the same with the economizer after optimizing the evaporator.

As we calculate the heat transfer and pressure drop at each node, we need correlations for both suitable for our geometry. For the heat transfer and pressure loss of the exhaust we use the ESCOA correlation (Ganapathy, 2002). For the water/steam inside the tubes the heat transfer coefficient is calculated from the Gnielinski (1976) correlation for single phase flow and the Bennett and Chen (1980) correlation during evaporation. The pressure losses are calculated with the Blasius correlation for single phase flow and the Friedel (1979) correlation during evaporation. To calculate the thermodynamic properties of the water the IAPWS formulation is used (Wagner and Prüss, 2002), while the cubic Peng-Robinson equation of state from Thermopack (Wilhelmsen et al., 2017) is used for the exhaust gas.

### 2.3 Transient modelling

The transient model uses many of the same principles as the steady state optimization. We assume equal conditions for all tubes in a row, so we only need to solve for one and we discretize each bundle along its width and depth.

However, in contrast to the steady state model, we use as input the exhaust entering the bottom of the superheater and the water entering the top of the economizer. In addition, we also calculate the wall temperature for each node, such that the heat transfer is calculated using Eq. 4 between the wall and each fluid separately, and conductive heat transfer along the wall is included. This means that the wall interacts with the two fluids and the temperature change  $\Delta T_{\text{wall}}$  in the wall is given by

$$\Delta T_{\text{wall}} = \frac{Q_{\text{ex}} - Q_{\text{w}} + Q_{\text{cond}}}{C_p}, \quad (6)$$

where  $Q_{\text{ex}}$  is the heat transferred from the exhaust to the wall and  $Q_{\text{w}}$  is the heat transferred from the wall to the water, both of which are usually positive,  $Q_{\text{cond}}$  is the net conductive heat transfer along the tube wall, and  $C_p$  is the combined heat capacity of the wall and fins at the temperature of the wall.

To start the transient modeling we first specify the temperature of each wall node and we then solve along each fluid flow. For the water we start from the inlet of the economizer, and for each node we calculate the heat transferred from the wall at that node to the water as well as the pressure loss of the water. This is then done all the way through until the end of the economizer which gives the inlet conditions for the water entering the evaporator. The same procedure is applied to the evaporator and superheater. For the exhaust gas we start from the exhaust gas inlet at the bottom of the superheater. Along each exhaust gas column, we calculate the heat transferred to the wall at that node as well as the pressure loss, then move to the next node above and so on until leaving the top of the economizer. Calculating the conductive heat transfer for each wall node allows us to use Eq. 6 and the time step to update all wall temperatures. This constitutes a single step in the time integration and this procedure is repeated until the chosen finish time of the integration. To ensure each time step is of a sufficient length the fourth-order Runge-Kutta-Fehlberg algorithm from GSL (Galassi et al., 2009) is used. In addition, during the integration, the inlet conditions of the water and exhaust gas, i.e. pressure, temperature, and composition or gas turbine load, can be changed freely at any specified time.

### 3. RESULTS

#### 3.1 Thermodynamic results

We first consider the simplified case of a heat exchanger without geometry, where the heat recovered and net power is calculated for varying values of the PPTD. For all the simulations the exhaust gas is assumed to come from a LM2500+G4 gas turbine running at 90% load, giving an exhaust inlet temperature at 510.1°C, pressure of 1bar and a mass flow of 86.12kg/s unless otherwise specified and we fix the outlet steam temperature to 480.1°C. The optimization was performed with PPTDs ranging from 0 to 30°C to give a good grasp on how the PPTD affects the power produced and the heat recovered. At a PPTD of 30°C we have pinch at both the onset of boiling and the water outlet and going lower than this we would only have pinch in the hot end. Having pinch in only the hot end would lead to a lower mass flow and less power

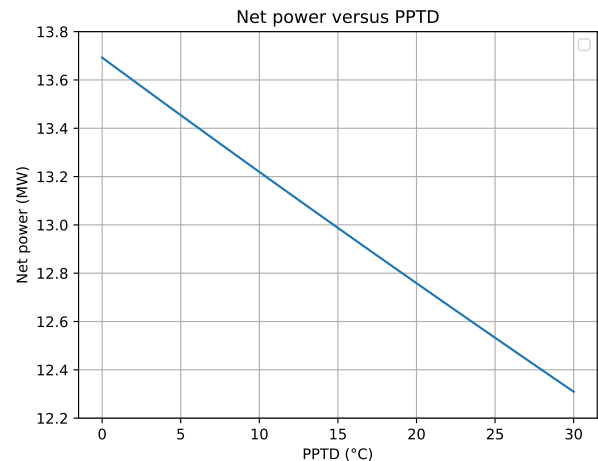


Fig. 4. Net power produced in the combined cycle OTSG without geometry for varying PPTD for fixed outlet temperature 480.1°C.

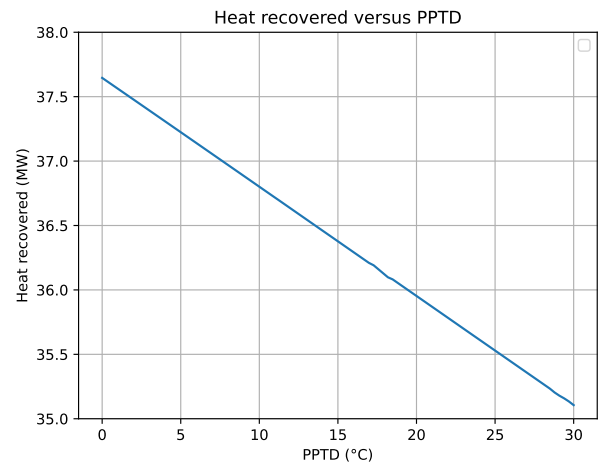


Fig. 5. Heat recovered in the OTSG without geometry for varying PPTD for fixed outlet temperature of 480.1°C.

produced and has not been investigated here. The net power produced is shown in Fig. 4 and the heat recovered from the exhaust gas is shown in Fig. 5. We see that the amount of heat recovered is larger than the net power produced by a factor of almost 3, with a maximum power produced of 13.7MW and a maximum heat recovered of 37.6MW, and that both decrease linearly with an increase in the PPTD. Thus, we want the PPTD to be as small as possible, but this does not take into account the size of the heat exchanger and the pressure loss in the heat exchanger. When the PPTD becomes smaller the necessary area increases and thus at a very small PPTD the OTSG needs to be very large. As obtaining a PPTD close to 0 is very hard we will use the flow conditions calculated for a PPTD of 10°C to optimize the geometries of the separate cores of the heat exchanger. The flow conditions are given in Table 2, and it also contains the heat duties used for the optimization where some extra duty is moved from the economizer to the evaporator.

Table 2. Input values for the water and exhaust gas stream for the geometric optimization.

Variable	Unit	Value
Inlet temperature of exhaust gas	[C]	510.1
Outlet temperature of steam	[C]	480.1
Inlet pressure of exhaust gas	[bar]	1
Outlet pressure of steam	[bar]	25.74
Mass flow of exhaust gas	[kg/s]	86.12
Mass flow of steam	[kg/s]	10.846
Superheater heat duty	[MW]	6.2
Evaporator heat duty	[MW]	23.0
Economizer heat duty	[MW]	6.2

### 3.2 Geometric results

Based on the flow conditions in Table 2 found in the thermodynamic optimization we performed the geometric optimization. The heat loads used are also given in the table, but they are slightly different from the thermodynamic optimization. In Fig 2 we see three distinct phases for the water, between the first and second pairs of dots it is heated, then, between the second and the third it boils and between the third and the fourth it is superheated. This corresponds to the economizer, evaporator, and superheater. However, for the geometric optimization we move some extra heat duty to the evaporator from the economizer to ensure that we avoid boiling in the economizer for different part loads. This could lead to instabilities such as the Ledinegg instability (Ledinegg, 1938) which we want to avoid. The pressure loss restriction was set to 600Pa for the exhaust in each core to get a total pressure drop of less than 3000Pa when including the inlet and exit transition ducts. The water pressure loss restriction was set to 100000Pa or 1bar in the evaporator and superheater and to 50000Pa in the economizer as here there is only liquid water which has a lower pressure loss. As lower tube diameters increase pressure losses it can become favourable to have many rows per pass when optimizing with respect to weight. However, this can become quite complicated to manufacture. Therefore we locked the number of rows per pass to be 2 in the evaporator and superheater where we have boiling and steam and thus higher pressure losses, and the number of rows per pass to 1 in the economizer as it has a lower pressure loss.

The weight of each of the cores, including both the tube bundle and ducting, and inlet duct for the superheater and outlet duct for the economizer, were optimized with widths varying from 3m to 7m with 0.5m intervals. For each width the design with the number of passes which gave the lightest weight were used as the input to the next core, i.e. from superheater to evaporator and evaporator to economizer. The investigated tube diameters are shown in Table 3 and are the tube diameters labeled as sensible range for offshore systems by Montañés et al. (2021) based on the ASME standard (STEELTUBE, 2021).

To investigate the weight savings of having different tube diameters in different cores we also consider cases where not all three cores have the same tube diameter. However, with the four tube diameters from Table 3 there are a total of 60 such combinations. Thus, we need to reduce this further. To begin with we notice that in the economizer there only flows liquid water, in the superheater there is only steam, while the evaporator has a mix. This

Table 3. The different tubes with inner and outer diameter that were investigated in this study.

Inner diameter [mm]	Outer diameter [mm]	Outer diameter
32.56	38.1	1 1/2"
27.53	31.75	1 1/4"
21.18	25.4	1"
15.75	19.05	3/4"

Table 4. The different combinations of inner diameters in the superheater, evaporator and economizer that were optimized for minimal weight.

Superheater [mm]	Evaporator [mm]	Economizer [mm]
32.56	32.56	32.56
32.56	32.56	27.53
32.56	27.53	27.53
27.53	27.53	27.53
32.56	27.53	21.18
27.53	27.53	21.18
27.53	21.18	21.18
21.18	21.18	21.18
27.53	21.18	15.75
21.18	21.18	15.75
21.18	15.75	15.75
15.75	15.75	15.75

means that we expect the pressure loss to be largest in the superheater and smallest in the economizer. In addition, the superheater is closest to the gas turbine so it experiences the harshest exhaust gas conditions while the economizer which is furthest away experiences the least harsh exhaust gas conditions. Thus, we expect that the superheater should be made of the sturdiest and largest tubes. Because of these two reasons we only consider combinations where the tube diameter does not increase when moving from the superheater to evaporator or evaporator to economizer, however, two or three of the cores can have the same diameter. In addition, we expect an actual design to not have huge differences in the tube diameters and thus we only consider designs where the change in tube outer diameter from one core to the next is at most 6.35mm or 1/4". Including the designs where the tube diameter is equal in all three cores leaves us with the designs consisting of the combinations shown in Table 4.

The optimized weights as a function of width for each of the cores in the OTSG for the tube diameter combinations given in Table 4 are shown in Fig. 6, where repeating combinations are excluded for the superheater and evaporator. For both the evaporator and the economizer we note that the weight has a very weak dependence on the previous core(s) and almost only depends on the tube diameter in that core. This is due to the fact that we fix the heat transfer in the previous cores and the pressure loss is relatively small in the evaporator and restricted in the superheater, leading to very similar conditions for the core independent of the geometry of the previous ones. However, for larger superheater diameters the pressure loss does not reach the maximum and thus we see a slightly lower weight of the evaporator as the water has a lower pressure and thus also a lower evaporating temperature. We also observe that reducing the tube diameter always leads to a weight reduction for the evaporator, the economizer and the total OTSG, and sometimes for the superheater. The reason

for this is that a smaller tube diameter gives more heat transfer area per volume which improves the overall heat transfer and thus less weight is needed. However, the smaller tubes also give a larger pressure drop in the tubes. In the superheater there is only steam flowing inside the tubes which leads to a pretty high pressure loss. As this pressure loss is restricted this leads to significantly more tubes per row being needed for smaller tube diameters to reduce the flow in each tube and thus the pressure loss, which counteracts the weight saving of having smaller tubes. Thus, we see that for a tube diameter of 21.18mm there is only a weight saving for widths up to 5m compared to the larger diameter, while for 15.75mm there is only a weight saving up to a width of 3.5m. In addition, this larger pressure drop for smaller diameters leads to the optimal width becoming smaller such that the length of each tube is shorter which also helps reduce the pressure loss. In the evaporator and economizer the pressure loss is not close to the restriction and thus we do not see the same effect in either of them.

If we compare the weight savings in each of the cores we see that in the superheater and economizer we only save about 9 and 8.5 tons respectively changing from the largest to the smallest tube diameter, while we save about 50 tons in the evaporator. As the evaporator transfers a lot more heat than the other two cores it is heavier and there is more to save by reducing the evaporator tube diameter. Finally, the total weight of the OTSG shows the same trend as each of the cores, with a noticeable weight reduction for each tube diameter. There are also four bands where the weights are relatively close, one for each evaporator diameter, which shows that reducing the evaporator diameter has the biggest effect on the weight of the entire OTSG.

### 3.3 Transient results

While the weight of the OTSG is very important for their use offshore, reducing the diameter of the tubes also changes the dynamic response of the OTSG. This is due to the fact that when reducing the tube diameter we get more heat exchanger area per volume of flow and per mass of the tubes. This means that we both reduce the amount of water stored inside the OTSG tubes and the mass of the tubes which reduces the total heat capacity of the system. This means that we expect faster response times for the designs with smaller tube diameters, however, we need to check how much each different core affects the response time. To do this we did a simple transient test. The OTSG was initialized with the flow values used for the steady state optimization and ran for a sufficiently long time such that it was in steady state. Then the gas turbine load was changed instantaneously from 90% to 50%, which amounts to a change in the exhaust gas temperature from 510.1°C to 510°C and a change in exhaust gas mass flow from 86.12kg/s to 66.34kg/s. This was then run until a new steady state and we found the time for the water outlet temperature from the superheater to reach within 1% and 0.1% of the new steady state value. The resulting times are shown in Table 5.

Looking at the response times we see that just like the weights of the OTSG the response time is generally reduced for reduced diameters, with some slight increases for

a few cases, which could be due to the width only being optimized in discrete steps. Changing from the heaviest design with the largest tube diameters to the lightest design with the smallest tube diameters halves the response time. Also, the reduction is not evenly distributed between the cores, instead, the evaporator clearly has the biggest impact. This is again due to the evaporator being significantly heavier and thus also having significantly more thermal mass in the tube bundle which takes longer to cool. The reduction in response time is very similar when reducing the diameter of the economizer and of the superheater, which is because of their weight saving, and thus the thermal mass reduction is approximately the same. Finally, we note that the response time is changed more for larger tube diameters. For example, reducing the evaporator diameter from 32.56mm to 27.53mm reduces the 1% response time by about 1900s while changing from an evaporator diameter of 21.18mm to one of 15.75mm only reduces the response time by about 1200s. However, if we look at the percentage-wise reduction this is about a 20% reduction for both cases. For comparison reducing the superheater diameter gives a reduction of -0.8%-2.2% and the economizer 0.3%-2.4% when ignoring the outlier with 21.18mm diameter in all three cores. Compared with the evaporator the relative change is significantly smaller for the response time contribution of the economizer and the evaporator compared to the weight savings, showing that the evaporator is significantly more important for the response time. This could also explain why the response time of the design with 21.18mm tube diameter in all three cores has a significant increase in response time compared to the designs with similar diameters. In this case, the superheater weight increases quickly with the diameter, leading to an optimal width smaller than the design with a larger superheater diameter. Thus, the evaporator becomes heavier and as it contributes more to the response time than the other cores compared to its weight this then leads to an increase in the response time.

## 4. CONCLUSION

In this study, a numerical framework has been used to minimize the weight of an OTSG consisting of three cores for varying tube diameters in each of the cores. First, the OTSG combined cycle was investigated without a geometric heat exchanger to find suitable operating conditions. The chosen operating conditions were then used to minimize the weight of each of the three cores in the OTSG. It was found that the evaporator was the heaviest and thus had the largest weight savings when reducing the tube diameter, while the superheater had diminishing weight savings due to higher pressure losses for steam in smaller tubes. Finally, the transient response for each of the optimal designs was investigated when reducing the exhaust gas flow rate and keeping the water flow rate constant. Here it was found that reducing the tube diameter and thus the weight generally reduced the response time of the OTSG due to reducing the thermal mass of the system. Just as for the weight the evaporator had clearly the biggest effect due to it having the largest thermal mass reduction when reducing the tube diameter.

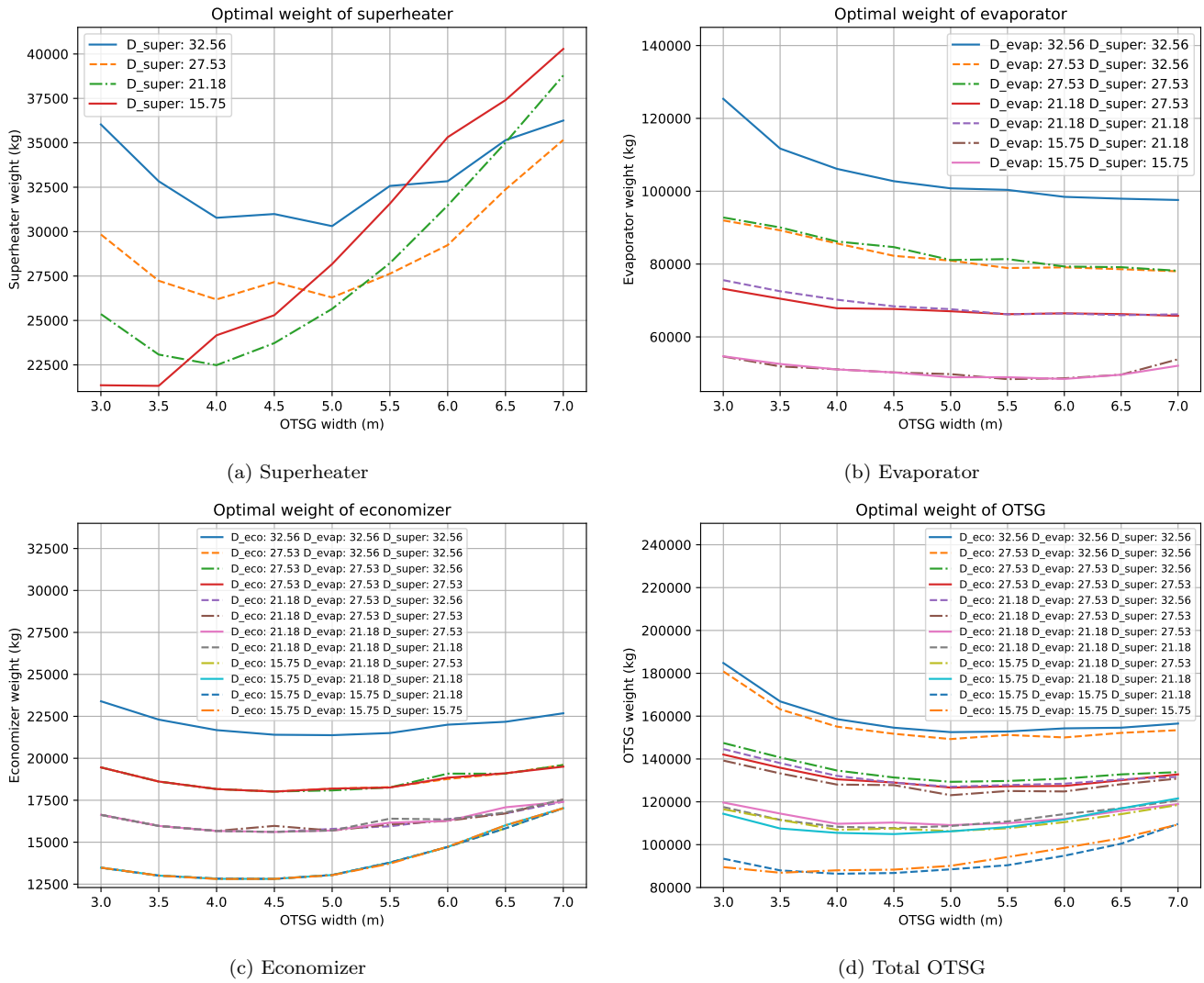


Fig. 6. Weight of the optimal design of the three different cores of the OTSG as well as the total weight of the OTSG for varying widths and diameter combinations given in Table 4. Repeating combinations in the superheater and evaporator are not included to ease readability. The minimum weight is reduced with tube diameter and the evaporator is the heaviest and has the biggest reductions.

Table 5. Time for each three-core design to reach 1% and 0.1% of the final steady state value of the water at superheater outlet when changing the gas turbine load from 90% to 50%. The response time is reduced for smaller tube diameters, and the evaporator has the biggest impact.

Superheater [mm]	Evaporator [mm]	Economizer [mm]	Time for ±1% [s]	Time for ±0.1% [s]
32.56	32.56	32.56	9521	19995
32.56	32.56	27.53	9563	19935
32.56	27.53	27.53	7657	16072
27.53	27.53	27.53	7698	16205
32.56	27.53	21.18	7596	15873
27.53	27.53	21.18	7517	15824
27.53	21.18	21.18	6484	13672
21.18	21.18	21.18	6657	14118
27.53	21.18	15.75	6393	13462
21.18	21.18	15.75	6280	13160
21.18	15.75	15.75	5079	10992
15.75	15.75	15.75	5114	10805

## REFERENCES

- Bennett, D.L. and Chen, J.C. (1980). Forced convective boiling in vertical tubes for saturated pure components and binary mixtures. *AIChE Journal*, 26(3), 454–461. doi:10.1002/aic.690260317.
- Deng, H., Skaugen, G., Næss, E., Zhang, M., and Øiseth, O.A. (2021). A novel methodology for design optimization of heat recovery steam generators with flow-induced vibration analysis. *Energy*, 226, 120325. doi:10.1016/j.energy.2021.120325.
- Friedel, L. (1979). Improved Friction Pressure Drop Correlation for Horizontal and Vertical Two-Phase Pipe Flow. In *European Two-Phase Flow Group Meeting*. Ispra, Italy.
- Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Alken, P., Booth, M., and Rossi, F. (2009). *GNU Scientific Library Reference Manual*. Network Theory Ltd., Bristol, third edition.
- Ganapathy, V. (2002). *Industrial Boilers and Heat Recovery Steam Generators: Design, Applications, and Calculations*. CRC Press, Boca Raton. doi:10.1201/9780203910221.
- Gnielinski, V. (1976). New equations for heat and mass transfer in the turbulent flow in pipes and channels. *International Chemical Engineering*, 16, 359–368.
- Ledinegg, M. (1938). Instability of flow during natural and forced circulation. *Die Waerme*, 61.
- Mazzetti, M.J., Hagen, B.A.L., Skaugen, G., Lindqvist, K., Lundberg, S., and Kristensen, O.A. (2021). Achieving 50% weight reduction of offshore steam bottoming cycles. *Energy*, 230, 120634. doi:10.1016/j.energy.2021.120634.
- Mazzetti, M.J., Ladam, Y., Walnum, H.T., Hagen, B.L., Skaugen, G., and Nekså, P. (2014). Flexible Combined Heat and Power Systems for Offshore Oil and Gas Facilities With CO<sub>2</sub> Bottoming Cycles. In *ASME 2014 Power Conference*. American Society of Mechanical Engineers Digital Collection. doi:10.1115/POWER2014-32169.
- Montañés, R.M., Hagen, B., Deng, H., Skaugen, G., Morin, N., Andersen, M., and J. Mazzetti, M. (2023). Design optimization of compact gas turbine and steam combined cycles for combined heat and power production in a FPSO system—A case study. *Energy*, 282, 128401. doi:10.1016/j.energy.2023.128401.
- Montañés, R.M., Skaugen, G., Hagen, B., and Rohde, D. (2021). Compact Steam Bottoming Cycles: Minimum Weight Design Optimization and Transient Response of Once-Through Steam Generators. *Frontiers in Energy Research*, 9. doi:10.3389/fenrg.2021.687248.
- Motamed, M.A. and Nord, L.O. (2022). Development of a simulation tool for design and off-design performance assessment of offshore combined heat and power cycles. In *63rd International Conference of Scandinavian Simulation Society, SIMS 2022, Trondheim, Norway, September 20-21, 2022*, 1–8. doi:10.3384/ecp192001.
- Nord, L.O. and Bolland, O. (2012). Steam bottoming cycles offshore - Challenges and possibilities. *Journal of Power Technologies*, 92(3), 201–207.
- Norwegian Petroleum Directorate (2023). Emissions to air. <https://www.norskipetroleum.no/en/environment-and-technology/emissions-to-air/>.
- Pierobon, L., Benato, A., Scolari, E., Haglind, F., and Stoppato, A. (2014). Waste heat recovery technologies for offshore platforms. *Applied Energy*, 136, 228–241. doi:10.1016/j.apenergy.2014.08.109.
- Pierobon, L. and Haglind, F. (2014). Design and optimization of air bottoming cycles for waste heat recovery in off-shore platforms. *Applied Energy*, 118, 156–165. doi:10.1016/j.apenergy.2013.12.026.
- Schittkowski, K. (1986). NLPQL: A fortran subroutine solving constrained nonlinear programming problems. *Annals of Operations Research*, 5(2), 485–500. doi:10.1007/BF02022087.
- Skaugen, G., Kolsaker, K., Walnum, H.T., and Wilhelmsen, Ø. (2013). A flexible and robust modelling framework for multi-stream heat exchangers. *Computers & Chemical Engineering*, 49, 95–104. doi:10.1016/j.compchemeng.2012.10.006.
- Skaugen, G., Walnum, H.T., Hagen, B.A.L., Clos, D.P., Mazzetti, M.J., and Nekså, P. (2014). Design and Optimization of Waste Heat Recovery Unit Using Carbon Dioxide as Cooling Fluid. In *ASME 2014 Power Conference*. American Society of Mechanical Engineers, Baltimore, Maryland, USA. doi:10.1115/POWER2014-32165.
- STEELTUBE (2021). Dimensions and Weights of Seamless Tubes According to Standard ANSI/ASME B36.10M.
- Wagner, W. and Pruß, A. (2002). The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *Journal of Physical and Chemical Reference Data*, 31, 387.
- Walnum, H.T., Nekså, P., Nord, L.O., and Andresen, T. (2013). Modelling and simulation of CO<sub>2</sub> (carbon dioxide) bottoming cycles for offshore oil and gas installations at design and off-design conditions. *Energy*, 59, 513–520. doi:10.1016/j.energy.2013.06.071.
- Wilhelmsen, Ø., Aasen, A., Skaugen, G., Aursand, P., Austegard, A., Aursand, E., Gjennestad, M.A., Lund, H., Linga, G., and Hammer, M. (2017). Thermodynamic Modeling with Equations of State: Present Challenges with Established Methods. *Industrial & Engineering Chemistry Research*, 56(13), 3503–3515. doi:10.1021/acs.iecr.7b00317.



# A Comparative Study of Conventional Lime Kilns and Plasma Calcination: Techno-Economic Assessment and Decarbonization Potential

Maximilian Dylong\*, Moksadur Rahman\*\*

\*KTH Royal Institute of Technology, Stockholm, Sweden  
(e-mail: [dylong@kth.se](mailto:dylong@kth.se)).

\*\*ABB Corporate Research Center, Västerås  
(e-mail: [moksadur.rahman@se.abb.com](mailto:moksadur.rahman@se.abb.com))

---

**Abstract:** Lime production is essential in the chemical recovery cycle of chemical pulping mills, typically relying on fuel combustion and thus contributing to greenhouse gas emissions. While Nordic pulp mills mainly use carbon-neutral biofuels, future biomass scarcity underscores the need for sustainable biomass management and alternative lime calcination methods. Electrification presents a promising solution, as CO<sub>2</sub> emissions depend on the carbon intensity of the electricity grid, which increasingly relies on renewable sources. Electrified solutions offer chemical pulp mills the opportunity to function as biorefineries and potentially produce higher-value biofuels in a constrained market. Plasma calcination provides benefits over conventional lime kilns, such as faster reaction times, reduced reactor volume, and lower shell losses. This work develops mathematical models for conventional kilns and plasma calcination to evaluate their techno-economic feasibility and decarbonization potential. A sensitivity analysis identifies influential parameters, and energetic requirements for both technologies under different fuel scenarios are assessed along with CO<sub>2</sub> emissions and economic factors. Results indicate that while plasma calcination's current decarbonization potential depends on the electricity grid's carbon intensity, future projections show its competitiveness over conventional kilns, with significantly lower CO<sub>2</sub> emissions across regions. The economic viability of plasma calcination is further influenced by projected carbon prices and process parameters, which impact its specific electricity consumption.

*Keywords:* Plasma calcination, Decarbonization, Electrification of heat, Lime kiln, Energy modelling, Pulp and paper

---

## 1. INTRODUCTION

The pulp and paper industry belongs to one of the most energy intensive industries in the world. Globally, it accounts for around 5% of total industrial energy consumption and is responsible for around 2% of emissions related to industry (IEA 2020). Despite rapid developments in digitalization, the pulp and paper industry is growing with a rate of 2% annually and is projected to continue rising due to several sectors, such as packaging or tissue production, more increasingly relying on paper-based products (IEA 2020; Summanen 2022). Currently, however, the pulp and paper industry is regarded as not being on track to reach net-zero emissions in 2050 which are obligated by the European Green Deal and legally binding through the European Climate Law (EC 2020; IEA 2020). More efforts need to be made in order to reduce carbon dioxide emissions by moving away from fossil fuels and adopting new technologies in the production process. Electrification is seen as a promising option for supplying the necessary heating demand currently covered by fuels as carbon dioxide emissions are ultimately subject to the carbon intensity of the electricity grid. With an increased development of renewable and low carbon sources, electrification is therefore a crucial strategy to reach ambitious climate goals and

contribute to the decarbonization of energy supply chains (IEA 2024).

One of the main heat intensive process steps in the pulp and paper industry is the production of lime within the chemical recovery cycle for chemical pulp mills. This is conventionally achieved in lime kilns that rely on fuel combustion to meet thermal energy demand. Lime kilns are primarily powered by fossil fuels such as natural gas or coal which is directly linked to carbon dioxide emissions (Falcke et al. 2017). Alternatively, biofuels can be utilized as a renewable energy source in lime kilns which is mostly the case in Nordic pulp mills. In 2020, 90% of the total energy used in Swedish lime kilns was supplied by biofuels, mostly tall oil pitch (63%) and bark powder or sawdust (24%) (Berglin and Schenck 2022). In Finnish lime kilns, roughly 45% of the energy was supplied by biofuels, mostly through gasified bark (18%), tall oil pitch (13%), and lignin powder (8%) (Berglin and Schenck 2022). Even though biofuels are considered a renewable energy source due to their participation in the carbon cycle, carbon dioxide emissions from combustion still occur (Newell 2010). Simultaneously, biomass demand is projected to increase significantly in the future, potentially surpassing the available supply (Material Economics 2021). This supply bottleneck

may be further exacerbated by the European Parliament's environmental committee, which advocates for restricting the use of primary woody biomass for energy purposes, arguing that an ecological limit on biomass harvesting must be considered (Material Economics 2021; Svebio 2022). This would conclusively result in a significantly tighter market for biofuels and an urgent need to consider how to best use biomass in Europe. Pulp mills therefore can play a significant role in becoming bio-refineries by converting available biomass residues into higher-value biofuels while at the same time opening the door for electrified calcination solutions.

Plasma calcination is seen as a promising, electrically driven solution for lime production (Madeddu et al. 2020). It can offer several advantages over conventional lime production through lime kilns such as faster reaction times, reduced reactor volume and decreased shell losses (Andersson and Skogström 2020). It is, however, not yet commercially available for the pulp and paper industry and comparative studies to conventional lime kilns showcasing the potential of plasma calcination have not been conducted. The key question, therefore, is to evaluate the comparative benefits of plasma technology versus lime kilns under both current and projected future conditions.

This work aims to conduct a comparative techno-economic assessment between the plasma calcination technology and conventionally used lime kilns. Mathematical models for lime kilns and plasma calcination are developed to analyze the energetic requirements for both technologies under the investigation of different fuels. Corresponding carbon dioxide emissions are further calculated to assess the decarbonization potential of the plasma calcination technology and an economic assessment including operational and capital expenditures is performed.

## 2. PROCESS DESCRIPTION

Two technologies are being considered and compared for the production of lime, those being conventional lime kilns and plasma calcination. Energy models for both technologies are developed to estimate the energy and fuel consumption as well as production related CO<sub>2</sub>-emissions under consideration of the relevant process steps and corresponding energy and mass balances.

Long rotary kilns are conventionally used in the pulp and paper industry for the production of lime. Fuel is combusted at the discharge of the lime kiln and its flue gases move counter-current to the flow of the lime mud, providing the necessary energy to cover the heating demand for lime production. In its core, a lime kiln can be divided into four heating zones as seen in Fig. 1. First, wet lime mud enters the kiln and is dried to remove all water. After that, the dried lime mud is heated up until it reaches the calcination temperature. The calcination reaction itself is endothermic and requires additional heat to occur. In a last step, the produced lime is sintered by further raising its temperature and agglomerating smaller particles to bigger ones. The hot lime is then discharged and cooled while preheating a secondary air stream to reduce the fuel consumption of the kiln (Bajpai 2018).

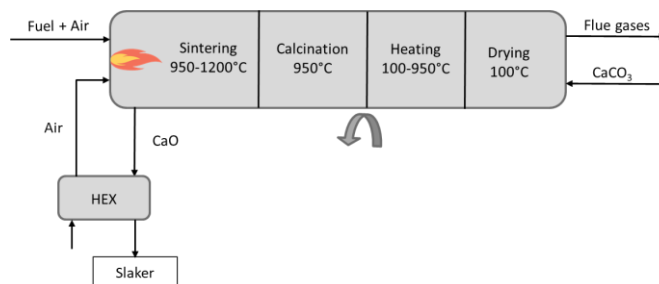


Fig. 1. Illustration of lime kiln with corresponding heating zones.

The plasma calcination technology aims to replace fuel driven lime kilns by providing the necessary heat for lime production with an ionised gas stream, also referred to as plasma, in a calcination reactor. The plasma stream itself is generated from an electric arc through which the gas stream passes, partially ionises and leaves as a gas-plasma mixture reaching temperatures of up to 5000 °C. The hot gas-plasma stream can then drive the calcination reaction (Andersson and Skogström 2020).

Figure 2 shows a simplified process schematic of the electric plasma calcination technology. Lime mud needs to be dried before entering the plasma reactor to reduce its electricity consumption as the evaporation enthalpy of water is comparatively high. For the scope of this work, a heat exchange network design is modelled in which the heat content of the hot discharge gases is used to dry the lime mud partially or fully. In the plasma generator, CO<sub>2</sub> is heated when getting in touch the electrically generated arc and leaves the plasma generator as a gas-plasma mixture to drive the calcination reaction in the plasma reactor together with the dried lime mud. The CO<sub>2</sub> stream after the plasma reactor needs to be cooled before the CO<sub>2</sub> stream is separated into two streams due to the additional CO<sub>2</sub> created during the calcination reaction. The first stream is compressed and recycled into the process while the second CO<sub>2</sub> stream is emitted and thus leaving the system boundaries. CO<sub>2</sub> created during plasma calcination is, however, of high purity meaning that the non-recycled CO<sub>2</sub> stream can be used in other process steps within the pulp production process, such as pulp washing (Bjotveit et al. 2003).

Plasma calcination can offer significant benefits compared to conventional lime kilns. Lime mud reburning with plasma can significantly reduce the process time of lime production from several hours in lime kilns to only a few seconds in a plasma reactor offering improvements in process control and reduced start and stop times. Additionally, equipment size is significantly smaller with a plasma reactor only having 1% of the volume of a lime kiln and no moving parts allowing for reduced operating costs and decreased shell losses (Andersson and Skogström 2020).

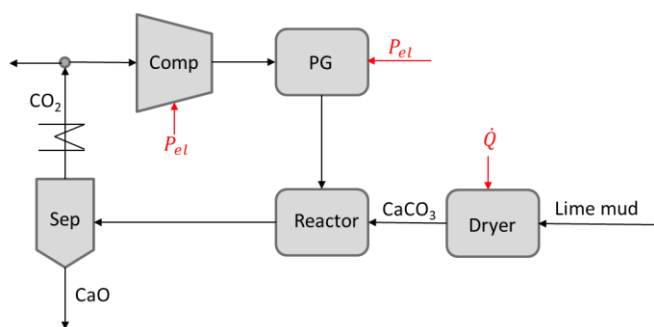


Fig. 2. Simplified process diagram for plasma calcination, adapted from (Andersson and Skogström 2020).

In order to quantify the energy demand of the conventional lime kilns and the plasma calcination technology, the two thermodynamic principles of energy and mass conservation are applied for every relevant component. Temperature dependant heat capacities are being considered as well as additional ionization enthalpy when modelling a phase change from gaseous to plasma state within the plasma calcination energy model. For the lime kiln model, different fuels can be investigated based on their molecular composition. Aside from the energetic requirements, both energy models further allow to estimate fuel and electricity consumption as well as corresponding CO<sub>2</sub>-emissions to further assess the decarbonization potential of the plasma calcination technology and its economic competitiveness.

### 3. RESULTS

The developed energy models for conventionally used lime kilns and the novel plasma technology allow for analysis of energetic requirements as well as corresponding fuel consumption and CO<sub>2</sub>-emissions. Both models were developed in Python 3.12.2. A list of relevant process parameters can be found in table 1.

Table 1. Process parameters

Parameter	Value	Unit	Source
Cooling temperature	200	°C	(Bajpai 2018)
Flue gas temperature	200	°C	(Lundqvist 2009)
Solids content	0.75	-	(Vainikainen 2021)
Sintering temperature	1100	°C	(Gulbrandsen and Stenqvist 2016)
Loss factor shell	0.1	-	(Lundqvist 2009)
Plasma temperature	3600	°C	(Andersson and Skogström 2020; Blackman 2024)
Reactor efficiency	1	-	(Bjotveit et al. 2003)

Figure 3 shows the sensitivity analysis results for the energetic analysis of both technologies. Key influential process parameters for the lime kiln model on its heat rate can be identified as the solids content of the lime mud entering the kiln, the flue gas temperature of the combustion products leaving the kiln and the loss factor accounting for heat losses through the shell of the kiln. In the plasma calcination model, partially ionized CO<sub>2</sub> provides the necessary heat for the calcination process. Results indicated that the temperature range of the ionized CO<sub>2</sub> has a significant impact on the electricity demand of the plasma generator, with a steep increase for plasma temperatures lower than 3000 °C. Additionally, variations in the sintering temperature have a higher impact on the specific electricity consumption of the plasma generator than on the heat rate of the lime kiln.

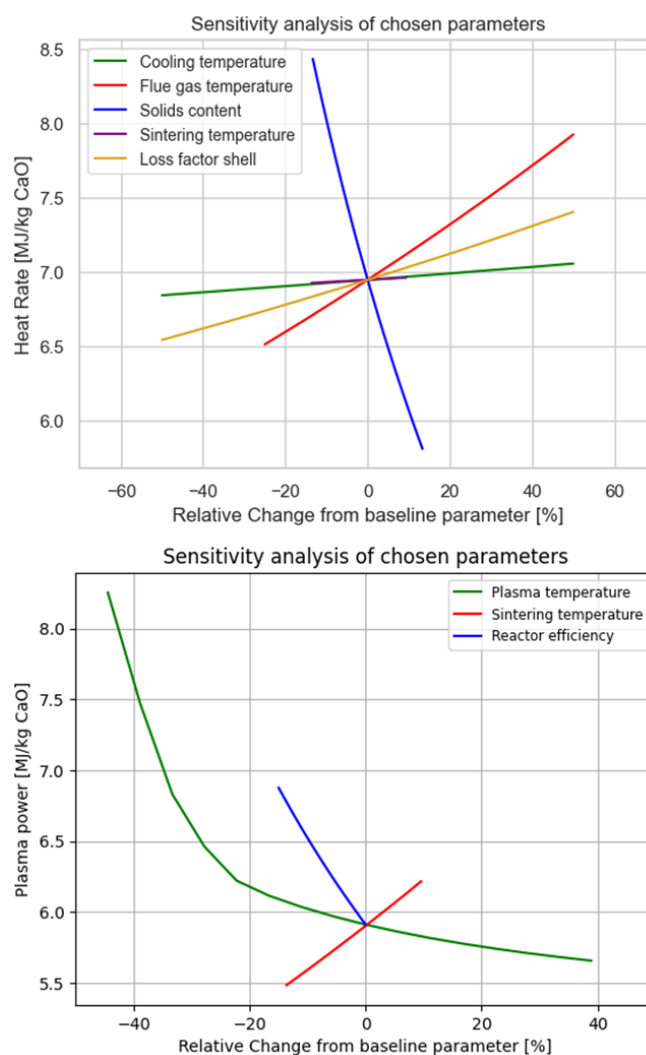


Fig. 3: Sensitivity analysis of chosen process parameters for the lime kiln model (top) and plasma calcination model (bottom). The values in the brackets of the legends refer to the corresponding baseline parameter.

The decarbonization potential of plasma calcination can be estimated with the modelling results generated by an in-depth sensitivity analysis in which the parameters described above

have been cross varied. The conducted emissions analysis takes into account direct CO<sub>2</sub>-emissions from fuel combustion in lime kilns and indirect emissions for the electricity supply of lime production via plasma technology. For the lime kiln, three fuels have been analyzed those being natural gas (NG), tall oil pitch (TOP) and bituminous coal (BC). CO<sub>2</sub>-emissions from plasma calcination are ultimately subject to the carbon intensity of the electricity grid. As carbon intensities vary, three regions have been selected to assess the decarbonization potential of the plasma calcination technology those being the EU, China and Sweden. Current as well as projected carbon intensities for 2030 and 2050 are taken into consideration for the analysis. Figure 4 illustrates the decarbonization potential of the plasma calcination technology. The boxplots show the potential range of specific CO<sub>2</sub>-emissions based on the multi-way sensitivity analysis for the given regions and time spans. The rectangles across the figure represent the interquartile range of CO<sub>2</sub>-emissions for the analyzed lime kiln fuels. It can be concluded that the current decarbonization potential of plasma calcination is heavily dependent on the carbon intensity of the corresponding electricity grid. Countries with a higher carbon intensity, such as China, cause higher CO<sub>2</sub>-emissions from plasma calcination in comparison to the analyzed lime kiln fuels. Countries with a low carbon intensity, such as Sweden, however, cause plasma calcination to outperform conventional lime kilns from a decarbonization perspective. Future trends regarding carbon intensity further reveal the competitiveness of plasma calcination and lead to lower specific CO<sub>2</sub>-emissions for all analyzed regions in comparison to the analyzed lime kiln fuels in 2050.

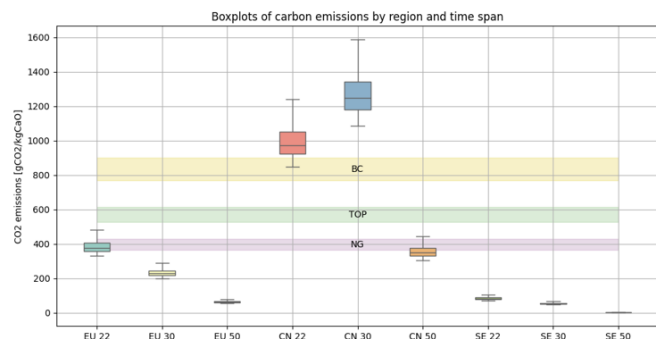


Fig. 4: Decarbonization potential of plasma calcination technology for different regions and time spans.

For the economic analysis, both operational and capital expenditure are being considered using the Net Present Value (NPV). Operational expenditures consider current and projected fuel and industrial electricity prices until 2050. Additionally, carbon prices are applied based on the EU ETS framework. For the capital expenditure, a 15 MW plasma generation system is considered with total investment cost for the plasma generator itself as well as auxiliary equipment. Based on the in-depth sensitivity analysis, this system would yield a lime production capacity of 121 – 245 t/day. This difference creates an upper and lower bound for the total investment costs of a lime kiln with equivalent production capacity. Figure 5 illustrates the NPV over a time span until 2050 as a function of the lime production capacity for two scenarios. The first scenario uses a high carbon price of 500

€/t in 2050 and the second one applies a more conservative carbon price of 200 €/t in 2050. Lower production capacities imply a lower electrical efficiency of the 15 MW plasma generation system while higher production capacities imply an increased electrical efficiency. For the high carbon price scenario both natural gas and bituminous coal can be outperformed by the plasma calcination technology even though higher electrical efficiencies of the plasma calcination system are required to result in a higher NPV than lime kilns fueled with natural gas. A lower carbon price requires higher electrical efficiencies of the plasma calcination system to be economically competitive with lime kilns fueled by bituminous coal while natural gas fired lime kilns always result in a higher NPV.

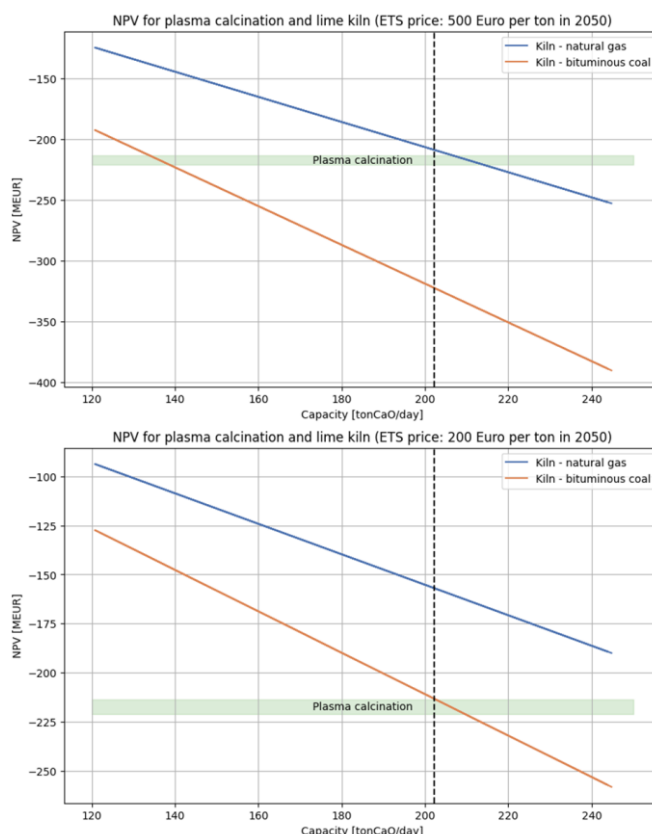


Fig. 5: NPV for plasma calcination system and lime kilns fueled by natural gas and bituminous coal for a carbon price of 500 €/t in 2050 (top) and 200 €/t in 2050 (bottom). The vertical line shows the median value obtained from the sensitivity analysis.

It can be concluded that the economic competitiveness of the plasma calcination technology is highly dependent on the projected EU carbon prices and the time of investment as a general increase of carbon prices is forecasted with a rather high uncertainty regarding the precise extend of this increase.

#### 4. DISCUSSION

The primary objective of this work was to explore the potential of plasma calcination as a sustainable alternative to conventional lime kilns for lime production within the pulp and paper industry. This investigation was motivated by the industry's need to reduce greenhouse gas emissions and align

with European and global decarbonization goals. The study examined the technological, economic, and environmental aspects of plasma calcination and compared those to conventionally used lime kilns. The generated results for the energetic requirements of both models are aligned with literature findings even though a data bottleneck for the plasma calcination technology has been identified. Plasma calcination for lime production in the pulp and paper industry is not yet applied on a commercial level and public availability of process flow sheets or process parameters is very limited causing the range of certain process, such as the plasma temperature, to be most likely wider than for a scenario in which this technology is well established in the industry. The decarbonization potential of plasma calcination is currently heavily dependent on the carbon intensity of the electricity grid. While the current decarbonization potential is geographically limited, the results show that future trends indicate plasma calcination to clearly outperform lime kilns from a decarbonization perspective. The economic competitiveness of plasma calcination is not only governed by the investment costs which are dependent on the selected process parameters but also on the operational expenses. The latter mostly depend on the projected carbon prices and corresponding policies such as the carbon leakage status of the pulp and paper industry or a potential inclusion into the Carbon Border Adjustment Mechanism.

Future work should explore enhancements and additions to the plasma calcination model developed in this work. The results generated in this work are based on simulation models. Especially for the plasma calcination model, for which a data bottleneck has been identified, experimental validation of the modeling results and potential adjustments would be desirable. Based on that, the plasma calcination model can further be enhanced by modelling the plasma generator and calcination reactor component in greater detail under consideration of dynamic heat transfer and different plasma torches. One major advantage of the plasma calcination system are fast ramping times which could be applied to fluctuating electricity prices to optimize lime production from an economic perspective. Additional gases other than carbon dioxide under different thermodynamic properties can further be investigated to increase the performance of the plasma calcination system.

## 5. CONCLUSIONS

This study demonstrates the potential of plasma calcination as a sustainable and efficient alternative to conventional lime kilns for decarbonizing lime production in the pulp and paper industry. While current adoption depends on the carbon intensity of the electricity grid, future projections indicate significant CO<sub>2</sub> emission reductions and competitive economic performance under anticipated carbon pricing scenarios. Plasma calcination offers technological advantages, including faster process times, reduced equipment size, and integration potential within biorefinery frameworks. Future efforts should focus on experimental validation, enhanced modeling, and dynamic optimization to fully realize the benefits of this promising technology.

## REFERENCES

- Andersson, Elin, and Arvid Skogström (2020). *Process Integration of Electric Plasma Calcination in Pulp and Paper Plants*. Chalmers University of technology.
- Bajpai, Pratima (2018). *Biermann's Handbook of Pulp and Paper: Paper and Board Making: Volume 2, Third Edition*. Elsevier.
- Berglin, Niklas, and Von Schenck, Anna (2022). *Biofuels in Lime Kilns*.
- Bjotveit, Sverre, Gordon, Anna, Ohlsson, Kristina, Rutqvist, Elin, Williamsson, Ewa, Westermark, Mats and Westin, Mari (2003). *Förstudie : Eldriven Kalcinerings*.
- Blackman, Corey (2024). *Electrification's Promising Path. International Cement Review*.
- EC (2020). *2050 Long-Term Strategy*.
- Falcke, Heino, Holbrook, Simon, Clenahan, Iain, Lopez Carretero, Alfredo, Sanalan, Teoman, Brinkmann, Thomas, Joze, R. , Benoît, Zerger, Serge, R. and Delgado Sancho, Luis (2017). *Best Available Techniques (BAT) Reference Document for the Production of Large Volume Organic Chemicals. Publications Office of the European Union: Luxembourg*.
- Gulbrandsen, Lars H., and Stenqvist, Christian (2016). *Pulp and Paper Industry*. 1st ed. Elsevier.
- IEA (2020). *Tracking Pulp and Paper 2020*.
- IEA (2024). *Analysis and Forecast to 2026*.
- Lundqvist, Per (2009). *Mass and Energy Balances over the Lime Kiln in a Kraft Pulp Mill*. Uppsala University.
- Madeddu, Silvia, Ueckerdt, Falko, Pehl, Michaja, Peterseim, Juergen, Lord, Michael, Ajith Kumar, Karthik, Krüger, Christoph, and Luderer, Gunnar (2020). *The CO<sub>2</sub> Reduction Potential for the European Industry via Direct Electrification of Heat Supply (Power-to-Heat)*. *Environmental Research Letters* 15(12). doi: 10.1088/1748-9326/abbd02.
- Material Economics (2021). *EU Biomass Use In A Net-Zero Economy - A Course Correction for EU Biomass*.
- Newell, Andrea (2010). *What Is the Carbon Disclosure Project? Triple Pundit*. Retrieved November 11, 2024 (<http://www.triplepundit.com/2010/11/carbon-disclosure-project/>).
- Summanen, Marko (2022). *10 Interesting Facts about the Current State of the Pulp and Paper Industry*. Retrieved November 11, 2024 (<https://www.resourcewise.com/forest-products-blog/10-interesting-facts-about-the-current-state-of-the-pulp-and-paper-industry-0>).
- Svebio (2022). *How Restrictions on "Primary Woody Biomass" Will Impact Swedish Energy and Climate Development*.
- Vainikainen, Olli (2021). *Modelling of a Lime Kiln Using Renewable Fuels and Study of Oxyfuel Combustion*. Aalto University.



# Assessment of Data-Driven Techniques for Flow Rate Predictions in Sub-sea Oil Production<sup>\*</sup>

Neville Aloysius D'Souza, Carlos Pfeiffer,  
Gaurav Mirlekar

*Department of Electrical Engineering, Information Technology and Cybernetics, University of South-eastern Norway (USN), Porsgrunn 3918, Norway (e-mail: gaurav.mirlekar@usn.no).*

**Abstract:** Accurate measurement of flow rate of the multiphase flow of oil, gas and water from the oil wells, is an important part of the oil and gas industry. This enables the safe operation and proper optimization of the production. With the increasing availability of process data, machine learning algorithms are used to create models for various applications. The application of these algorithms for flow rate estimation provides a more accurate representation of the oil and gas production process. In this paper, two oil wells and ten machine learning algorithms are evaluated. Long Short-Term Memory (LSTM) provides the best results with Mean Absolute Percentage Error of 1.96% for Well 1 and 1.56% for Well 2. In addition, the effects of noise on the models are explored. Median filter with window size of three provides good noise reduction. The uncertainty of the predictions are quantified using 95% confidence intervals in XGBoost model.

**Keywords:** Machine learning techniques, Data-driven estimations, Uncertainty quantification, Measurement noise

## 1. INTRODUCTION

The production of oil and gas requires measurements of various process data. This process data is used to ensure optimal production of oil and gas and the safe operation of the production system. One of the most important variables necessary for this is the accurate measurement of oil, gas, and liquid flow rates from the oil wells. Since there is multiphase flow from the oil wells, it is a challenge to obtain the individual flow rates of oil and gas. Typically, a separator is used as shown in Fig 1 to obtain an accurate flow rate of oil, gas and water. Here to measure the individual phases the multiphase mixture are separated physically with a separator. Phase flow meters are used to obtain accurate flow measurements, Bikmukhametov and Jäschke (2020). This process requires a steady state flow from the given oil well. In addition to this, the other oil wells have to be shut down to avoid interference with the results. This is a costly and time consuming process.

To estimate the flow rates without use of separators, Multiphase flow meters (MPFMs) can be used. While MPFMs has many advantages in measurement of multiphase flow, they are very expensive. Also, the accuracy of them gets degraded over time. In addition maintenance of these sensors are important to ensure good working conditions. Oil and gas production systems will already have many sensors installed which monitor certain physical

<sup>\*</sup> We gratefully acknowledge the support from the Research Council of Norway and Equinor ASA through Research Council project "308817 - Digital wells for optimal production and drainage" (Digi-Well). Corresponding author's e-mail: gaurav.mirlekar@usn.no.

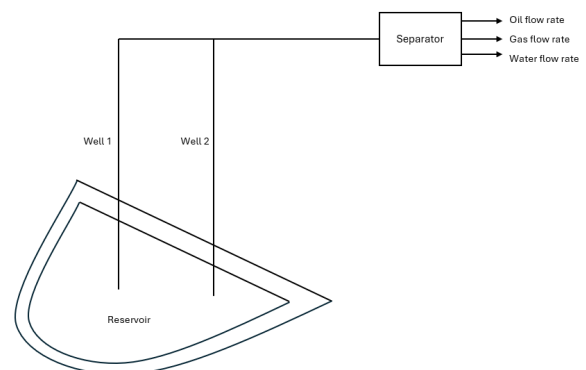


Fig. 1. Example of sub-sea oil production

quantities. The use of data driven modelling (also called machine learning modelling) in the oil and gas industry has been increasing with the availability of and storage of process data. As early as 1993, Qiu and Toral (1993) have used neural networks to estimate the flow rates of multiphase flow. Considerable research is ongoing to improve the application of machine learning models in the oil and gas industry. AL-Qutami et al. (2018) also uses neural networks to estimate phase flow. They used feed forward neural networks with k fold cross validation and early stopping. A more advanced method using LSTM has been used by Andrianov (2018) to forecast forecast the rates for a series of future time instants in addition to reliably estimating the multiphase rates at the current time. For VFM the process data usually collected are:



- Bottomhole pressure and temperature.
- Wellhead pressure and temperature upstream of the choke.
- Wellhead pressure and temperature downstream of the choke.
- Choke opening values.

### 1.1 Objective

The main objective here is to assess the machine learning algorithms for flow rate predictions in a sub sea oil production system. The sub-tasks in this are 1) Data collection and preprocessing, 2) Predictions of flow rates of oil, gas using machine learning, 3) Evaluation of measurement noise on machine learning algorithms performance and 4) Quantification of the uncertainty in the predictions.

## 2. METHODS

### 2.1 Process Description

For the oil well, the simplified schematic is shown in Fig 2. Through the gas lift choke valve, high-pressurized natural gas is continually injected into the wells annulus in this system, which is mostly utilised to extract lighter crude oils. The injected gas finds its way into tubing at some points located at proper depths and mixes with the multi-phase fluid from the reservoir. As a result of this mixing, the density of the fluid in the tubing will be reduced, which means that the flowing pressure losses in the tubing reduce. Consequently, the reservoir pressure will be able to overcome the flowing resistance in the well and push the reservoir fluid to the surface. Each well has its own inflow characteristics. A graphical representation of this project is shown in Fig 3.

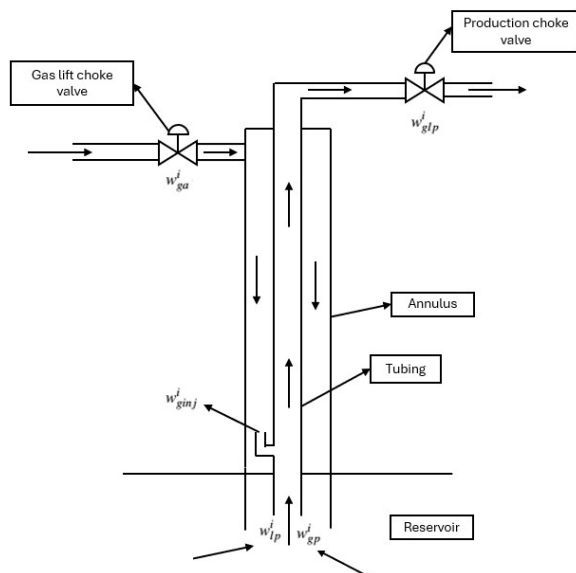


Fig. 2. Single oil well schematic

### 2.2 Machine learning algorithms

A brief description of the machine learning algorithms used in this paper are describe here

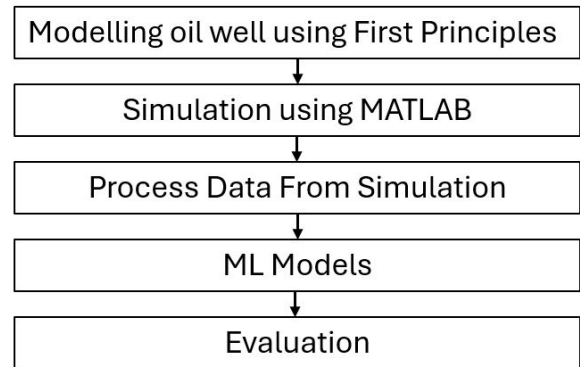


Fig. 3. Process Flow of the Project

*Multivariate Linear Regression* Linear Regression is the simplest machine learning algorithm. It makes a prediction by simply computing a weighted sum of the input features, plus a constant called the bias term, as shown in equation (1).

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \quad (1)$$

Where,

$\hat{y}$  - predicted value,

$n$  - number of features,

$x_i$  -  $i$ th feature value,

$\theta_j$  -  $j$ th model parameter,

$\theta_0$  - bias term.

This can be modified to output multiple  $\hat{y}$  values. Multivariate linear regression is a statistical technique that models the linear relationship between multiple independent variables and a single dependent variable. It extends simple linear regression by allowing for the inclusion of more than one predictor variable. The goal is to find the linear equation that best predicts the dependent variable based on the independent variables Zangl et al. (2014).

*k-Nearest Neighbors Regression* The  $k$ -nearest neighbors (kNN) algorithm is a non-parametric, supervised learning method used for classification and regression tasks. It works by identifying the  $k$  closest training examples to a given data point and assigning a class or value based on the majority vote or average of those neighbors. kNN is a versatile algorithm that can handle both numerical and categorical data without making assumptions about the underlying data distribution. It is commonly used in applications like recommendation systems, pattern recognition, and anomaly detection. The choice of  $k$  is important, as lower values can lead to overfitting while higher values may cause underfitting, Géron (2023).

*Support Vector Regression* Support Vector Regression (SVR) is a non parametric technique that uses kernel functions to estimate a function from a set of training data. The goal is to find a function  $f(x)$  that deviates from the target values  $y$  by no more than  $\varepsilon$ , while being as flat as possible. This is achieved by solving a convex optimization problem that minimizes the norm of  $w$ , subject to the constraint that the regression errors are within  $\varepsilon$ . SVR can handle high-dimensional data and

nonlinear relationships by implicitly mapping the input data into a higher-dimensional feature space using kernel functions. Unlike other regression models that try to minimize the error between the real and predicted values, SVR tries to fit the best line within a threshold value (distance between the hyperplane and boundary line). The data points on either side of the hyperplane that are closest to the hyperplane are called Support Vectors, which are used to predict the output Xu et al. (2012). SVR has several advantages, such as being robust to outliers, having excellent generalization capability, and easy implementation. However, it is not suitable for large datasets, and its performance may degrade when the number of features exceeds the number of training samples Smola and Schölkopf (2004).

*Decision Tree Regression* A decision tree algorithm is a supervised machine learning technique used for both classification and regression tasks. It constructs a tree-like model of decisions based on the data's attributes. The process starts at the root node and splits the data into subsets using the most significant attribute based on selection criteria like information gain or Gini impurity. Each internal node of the tree represents a "test" on an attribute, each branch represents the outcome of that test, and each leaf node represents a class label or a continuous outcome. The paths from root to leaf represent classification rules or regression paths. Decision trees handle both numerical and categorical data and are intuitive, as they mimic human decision-making processes. They are particularly useful in scenarios where relationships between parameters are nonlinear or complex. However, decision trees can suffer from overfitting, especially with very complex trees. Techniques such as pruning are used to remove parts of the tree that do not provide additional power in order to reduce overfitting and improve the model's generalizability. Decision trees are foundational elements in more complex algorithms like Random Forests and boosting methods, enhancing their stability and accuracy Dayev et al. (2023).

*Gradient Boosting Regression* Gradient Boosting Regression is a powerful machine learning algorithm that combines multiple weak models to form a strong learner. It is particularly effective for regression problems where the goal is to predict continuous values. The algorithm works by iteratively training decision trees on the residuals of previous predictions, which are the differences between the actual and predicted values. Each tree is trained to minimize the error of the previous tree, and the learning rate determines the contribution of each tree to the final prediction. The process begins with an initial guess, typically the mean of the target variable. Then, at each iteration, a new tree is trained to predict the residuals from the previous tree. The residuals are the differences between the actual and predicted values. The new tree is added to the previous trees, and the process is repeated until a stopping criterion is reached, such as a maximum number of trees or a minimum improvement in the model's performance. The final prediction is the sum of the predictions from all the trees, weighted by their learning rates. This approach allows the algorithm to capture complex relationships between the input variables and the target variable, making it highly effective for regression problems Kniazev et al. (2023).

*XGBoost Regression* XGBoost is a powerful algorithm for building supervised regression models. It was developed by Chen and Guestrin (2016). It is an implementation of gradient boosting that is designed to be highly efficient and scalable. The algorithm is particularly effective for regression problems where the goal is to predict continuous or real values. XGBoost is based on the concept of ensemble learning, where multiple base learners are trained and combined to produce a single prediction. The core components of XGBoost for regression include the objective function, base learners, and regularization. The objective function is responsible for defining the loss function and the regularization term. The base learners are the individual models that are trained and combined to produce the final prediction. Regularization is used to prevent overfitting by penalizing complex models. XGBoost uses a unique approach to building regression trees. Each tree starts with a single leaf and all residuals go into that leaf. The algorithm then calculates a similarity score for this leaf based on the residuals. The similarity score is used to determine how to split the data into two groups. This process is repeated recursively until a stopping criterion is reached. XGBoost is widely used in various applications due to its high accuracy and efficiency. It is particularly effective for large datasets and can be easily integrated with other tools and packages such as scikit-learn and Apache Spark.

*PC Regression* Principal component regression (PCR) is a regression analysis technique that combines principal component analysis (PCA) and linear regression. The key idea behind PCR is to first perform PCA on the predictor variables to obtain a set of uncorrelated principal components, and then use these principal components as the new predictors in a linear regression model, instead of the original variables. The advantages of PCR are that it can help address issues like multicollinearity and high dimensionality in the predictor variables. By using a subset of the principal components, PCR can reduce the number of predictors in the regression model, which can improve the model's interpretability and generalization performance. However, PCR does not perform feature selection, as each principal component is a linear combination of all the original predictors. While PCR can be a useful technique, it has some limitations. It relies on the assumption that the directions of maximum variance in the predictor variables are also the most predictive of the response variable, which is not always the case. Additionally, PCR can result in information loss, as it discards some of the principal components during the regression step Bello et al. (2014).

*PLS Regression* PLS regression is a powerful statistical technique that is particularly useful for analyzing high-dimensional data with many predictor variables. The key idea behind PLS regression is to find a set of latent components (linear combinations of the original predictors) that maximize the covariance between the predictors and the response variable. Unlike traditional linear regression, PLS does not require the predictors to be orthogonal or the number of predictors to be less than the number of observations. PLS regression works by iteratively extracting latent components that explain as much of the covariance between the predictors and response as possible. The resulting PLS model provides both dimension reduction and

regression coefficients, allowing for accurate prediction of the response variable from the original high-dimensional predictors. PLS regression has several advantages over other regression methods, including its ability to handle multicollinearity, its robustness to noise, and its suitability for datasets with more predictors than observations. As a result, PLS is a widely used technique in fields such as chemometrics, bioinformatics, and marketing research Boulesteix and Strimmer (2006).

**MLP Neural network** A Multilayer Perceptron Neural Network (MLPNN) is a type of artificial neural network that consists of multiple layers of interconnected nodes, or neurons. Unlike a single-layer perceptron, which can only learn linearly separable patterns, an MLP can learn more complex, non-linear relationships in data. The key components of an MLP are the input layer, one or more hidden layers, and an output layer. The input layer receives the data, which is then passed through the hidden layers, where the network learns to represent the data in a more abstract way. Each hidden layer applies a non-linear activation function to the weighted sum of its inputs, allowing the network to learn complex patterns. The final output layer produces the predicted result. MLPs are trained using a supervised learning algorithm, typically back propagation, which adjusts the weights of the connections between neurons to minimize the error between the predicted and actual outputs. This iterative process allows the MLP to learn the underlying structure of the data and make accurate predictions on new, unseen data.

**LSTM** Long Short-Term Memory (LSTM) is a type of recurrent neural network designed to address the vanishing gradient problem in traditional RNNs. The key feature of LSTMs is their memory cell, which can selectively retain or discard information as it flows through the network. LSTMs have three gates that control the flow of information: the input gate, the forget gate, and the output gate. The input gate decides what new information from the current input and previous output should be added to the memory cell. The forget gate determines what information from the previous memory cell should be retained or forgotten. The output gate controls what information from the current memory cell and input should be used to produce the output. This gating mechanism allows LSTMs to learn long-term dependencies in sequential data, making them well-suited for tasks like language modeling, machine translation, speech recognition, and time series forecasting. LSTMs have been widely adopted and have significantly advanced the state-of-the-art in many sequence-to-sequence learning problems.

Data driven VFM (also called machine learning VFM) is the method where a model of the oil and gas production system is created using the available sensor data. Here in depth domain knowledge about the process is not necessary to create a model. A typical schematic for a sub-sea oil and gas production systems which used data driven VFM is shown in Fig 4. Broadly, the steps involved are as follows:

- (1) Data collection.
- (2) Data pre processing.
- (3) Model development.

- (4) Predictions of flow rates.
- (5) Data reconciliation.

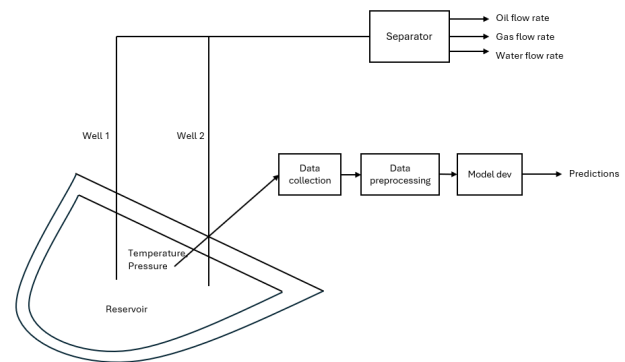


Fig. 4. Data driven VFM.

### 2.3 Data collection

The first step to creating a data driven model is the collection of relevant data. In Virtual Flow Metering systems, information is transmitted from wells and processing facilities and this includes sensor readings. This data may be wireless transmitted using IoT systems or through physical communication wires. It can involve different communication protocols to ensure proper transmission of data. Historical data from the same or analogous fields may also be used as a calibrating data set for fine-tuning the model. Generally, the data collected tends to be unclean, contaminated, and may have missing values, outliers and redundant inputs. Here the data is obtained from the oil well model described by Janatian et al. (2022). Using the equations described in the paper, an open loop simulation of the oil well is obtained. For the oil well 1 and 2, 5762 samples are obtained. These are split in 70% train and 30% test sets.

### 2.4 Data pre processing

Data filtering, where the removal of noise from raw data is performed is part of this step. There exist many filters that can be deployed to clean the raw data. In addition outlier detection, correcting missing values can be included. Preprocessing can also involve data transformation, which might yield new insights about the information the data contains. Feature engineering is the common term for this technique. Since the oil and gas production process is time dependent, care should be taken when splitting the data for training. Time series split from Sci-kit learn library is used for this. Two splits are created as shown in Fig 5. Using this, data leakage in the training stage can be avoided. A standard scaler is used on the inputs of the train dataset. It involves re-scaling each feature such that it has a standard deviation of 1 and a mean of 0. This is necessary since the inputs have different ranges. It also help to reduce training time in certain algorithms like SVR and neural networks.

### 2.5 Model development

In order to create a model, an algorithm that can map input features to output (target) variables must be developed. The mapping process, also known as training or

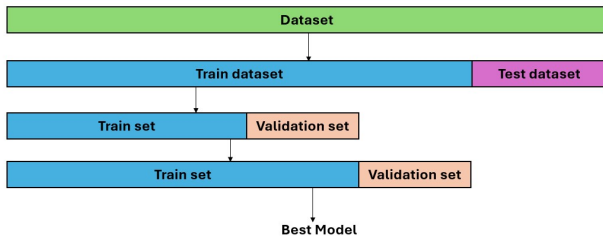


Fig. 5. Time series splitting.

learning, involves the algorithm modifying the parameters so that it can precisely estimate the desired variables. Depending on the algorithm being used, the parameters must be changed. The weights that connect the neurons in a neural network, for example, are the parameters. In regression trees, on the other hand, the parameter may be the tree depth. Reducing the difference between the algorithm's predicted values and the actual (measured) values to minimise a cost function, is how training is accomplished. Mean squared error (MSE) is usually used as a cost function to solve regression problems such as Virtual Flow Metering. Here ten algorithms are used to create models for Well 1 and Well 2. LSTM Regression, Multi-variate Linear Regression, KNN Regression, Decision Tree Regression, Gradient Boosting Regression, XGBoost Regression, Principal Component Regression, Partial Least Squares Regression, and MLP Neural Network Regression are used.

### 2.6 Predictions of flow rates

Once the training and validation for the model is completed, the model is tested on unseen data. New predictions from this data are noted and the effectiveness of model can be determined. For oil and gas flow rate predictions the commonly used performance metric is the Mean absolute percentage error (MAPE). With this the performance across various algorithms can be compared. It is easy to interpret and can be used across different input data scales. MAPE can be found by:

$$\text{MAPE}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)} \cdot 100 \quad (2)$$

where  $\hat{y}$  is the predicted value of the  $i$ th sample,  $y_i$  is the corresponding true value,  $n_{\text{samples}}$  is number of samples,  $\epsilon$  is an arbitrary small yet strictly positive number to avoid undefined results when  $y$  is zero.

### 2.7 Data reconciliation

An optimization algorithm adjusts the model parameters, for instance, flow rates, choke discharge coefficient, gas and water fractions, and friction and heat transfer coefficients such that the model outputs match the validated measured data being constrained to process conditions, for instance, the material balances. The reconciliation procedure in virtual flow metering systems is frequently expressed in the constrained least squares form.

## 3. RESULTS AND DISCUSSIONS

### 3.1 Predictions using algorithms

For Well 1, Figures 6 shows the results of each algorithm. Figure 7 shows the predictions for Well 2.

For LSTM model early stopping is used to prevent overfitting. This is implemented in Tensorflow. For well 1, 32 memory cells are used. For well 2, 40 memory cells were used. The 'adam' optimizer with loss function of mean squared error is used for training both models. A linear activation unit is used in the output layer.

In the Multi-layer perceptron neural network model, for well 1, 2 hidden layers are used, 'identity' activation function, an optimizer in the family of quasi-Newton methods (lbfgs solver) is used, L2 regularization of 0.0005 and the maximum number of iterations is 1000. For well 2 only 'logistic' activation function is changed and other hyper-parameters are same as well 1.

In the Multi variate linear regression, for well 1 and well 2 the intercept is calculated. The copy\_X parameter is true, which means the input features are copied, not overwritten.

In the Support vector regression model, for well 1, radial basis function is used as kernel, with kernel coefficient of 0.1, and regularization of 1 is used. For well 2, radial basis function is used as kernel, with kernel coefficient of 0.01, and regularization of 10 is used.

In the K nearest neighbors model, for well 1 and well 2, 8 neighbors were used. For well 1, the power parameter for the Minkowski metric is used. For well 2, the power parameter for the euclidean distance metric is used. These metrics are used in distance computation.

For the Decision tree model, for well 1, the maximum depth of the tree is 1, the minimum number of samples required to be at a leaf node is 1, the minimum number of samples required to split an internal node is 2 and the number of features to consider when looking for the best split is the sqrt of the number of features. For well 2, the maximum depth of the tree is 10, the minimum number of samples required to be at a leaf node is 1, the minimum number of samples required to split an internal node is 10 and the number of features to consider when looking for the best split is the log2 of the number of features.

For Gradient boosting model, for well 1 and well 2, the loss function used is squared error for regression. The maximum depth of the individual regression estimators is 3. For well 1, the learning rate is 0.1, and for well 2 the learning rate is 0.05. With this algorithm it is possible to find the importance of each input feature. Here  $P_{wh}$ , the wellhead pressure shows the maximum effect.

A more efficient and faster implementation of gradient boost, XGBoost model Zheng et al. (2022) is developed next. Here for well 1, the squared error is used as objective function, 300 estimators, maximum depth of 7 and learning rate of 0.01 is used. For well 2, the number of estimators are reduced to 100, maximum depth of 3 and learning rate of 0.1.

Prediction by multiple models for Well 1

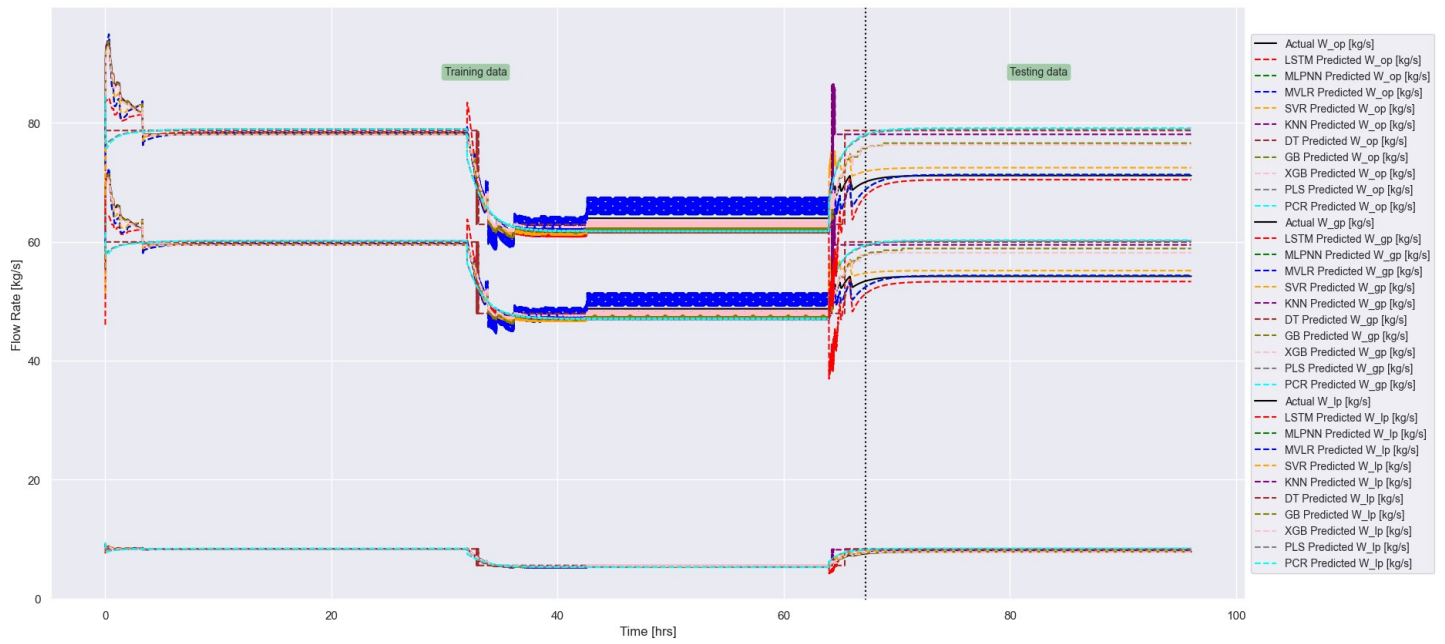


Fig. 6. Flow rate Predictions for Well 1.

Prediction by multiple models for Well 2

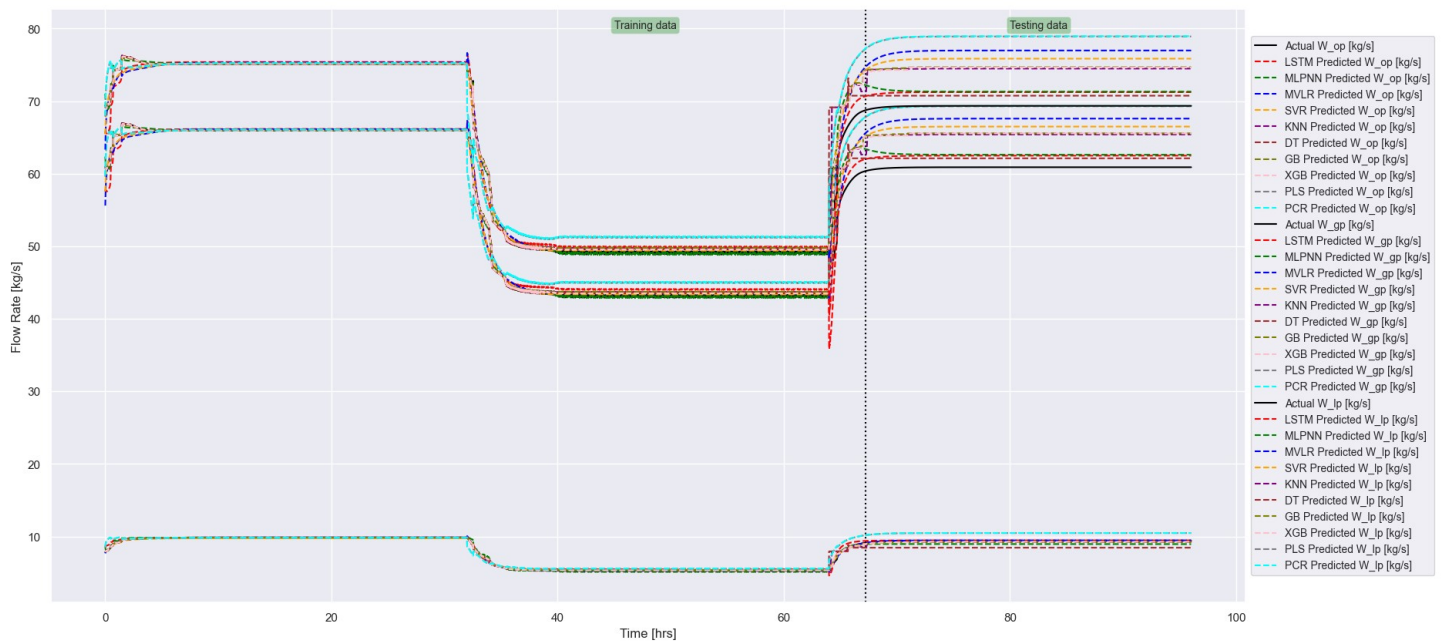


Fig. 7. Flow rate Predictions for Well 2.



For the PLS and PCR models, the number of component used in 1 in both wells. These models are useful when dimension reduction of input feature space is needed.

The MAPE for each model is shown in Table 1

Table 1. MAPE for Well 1 and Well 2.

Algorithm	Well 1 (%)	Well 2 (%)
LSTM	1.96	1.53
MLP NN	2.43	5.49
MV Linear Regression	2.14	7.57
SVR	5.04	4.31
KNN	8.05	5.41
Decision Tree	9.26	5.43
Gradient Boost	4.95	5.55
XGBoost	4.23	5.56
PLS	9.54	7.57
PCR	9.52	16.69

The LSTM model produces the best results. The disadvantages of using this is the training time is longer. Also to find the proper parameters is a time consuming process. It is observed that for each well the hyper-parameters has to be tuned. GridSearchCV helps with this, but it is still a complicated process. For the algorithms that are generally used for classification tasks like SVM, kNN, some modification is required to enabling its use for regression. Many of these algorithms including linear regression, and tree based, require modification to predict multiple outputs. With modifications it is possible to get the results, but the downside is the hyperparameter tuning becomes more complex. Neural networks and the LSTM model can be made more complex, giving better results. This takes more time and computation power. For finding the best hyper-parameters multiple runs are required. Since the programs were executed on a laptop, these take more time. For decrease in computation time a sample size of 5762 was used. If more samples were used in the modelling the results would probably be much better.

3.2 Effects of noise

The effect of random errors is tested on three machine learning models: XGBoost, MLP NN and LSTM. Impulse noise introduces sudden jumps or falls in the data values, simulating real-world data with occasional spikes at random locations. First a noise sample of 3% is created. The values in the sample are uniformly distributed between 20% of the minimum value of the column and 30% of the maximum value of the column. This ensures that the noise added is relative to the range of the data in the column. The noise is randomly distributed across the column and added to the 3 input features. The 3 algorithms are trained and tested. Here the figures are shown of only Well 1, since the effects are can be similarly observed in Well 2. Figures 8 and 9 show the effect of impulse noise on LSTM, MLP NN and XGBoost models respectively. To solve the problem of impulse noise, there are many filters that can be used. For example Median filter, Order statistic filters, and so on. Here the Median Filter is used to reduce the impulse noise. SciPy is used, which has a median filter function that is well-suited for removing impulse noise, as it replaces each data point with the median of the neighboring data points within a specified window size. The results of the

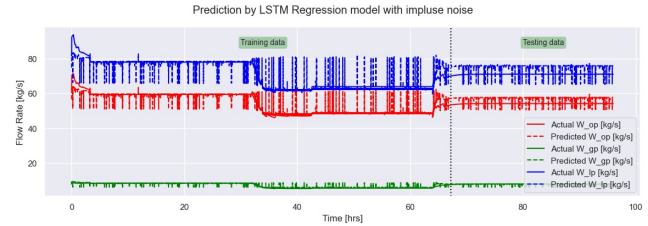


Fig. 8. Effect of Impulse noise on LSTM model.

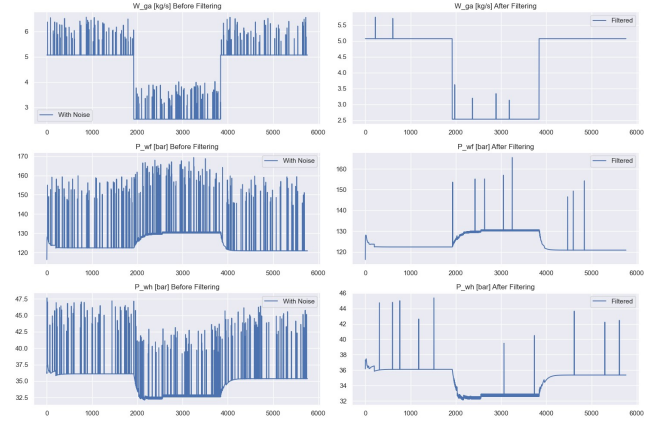


Fig. 9. Median noise filter.

median filter is shown in Fig 9. A window size of 3 is used. Each data point is replaced with the median of itself and its two neighbors. Most of the impulses are filtered out. The prediction accuracy of the 3 models is improved. Tables 2 and 3 shows the effect of impulse noise on MAPE of the 3 models.

Table 2. Impulse noise effects

Well no.	LSTM (%)	MLP NN (%)	XGBoost (%)
Well 1	5.98	8.76	7.51
Well 2	4.67	5.13	5.77

Table 3. Median Filter effects

Well no.	LSTM (%)	MLP NN (%)	XGBoost (%)
Well 1	1.87	4.97	6.47
Well 2	2.86	5.51	5.29

3.3 Uncertainty quantification

Uncertainty refers to a state of limited knowledge or information, where it is impossible to precisely describe an existing state, future outcome, or multiple possible outcomes. There are two main types of uncertainty: Aleatory and Epistemic Uncertainty Pelz et al. (2021). There are many methods to quantify the uncertainty in predictions for machine learning models. Some of them are: Confidence intervals, Quantile regression, Bootstrapping, Ensemble methods and Bayesian optimization.

Using XGBoost the confidence intervals can be added. For other algorithms like LSTM it is a more complicated process. Figs. 10 and 11 show the confidence intervals of 95% for XGBoost model for well 1 and well 2.





Fig. 10. Confidence intervals for XGBoost model (Well 1)

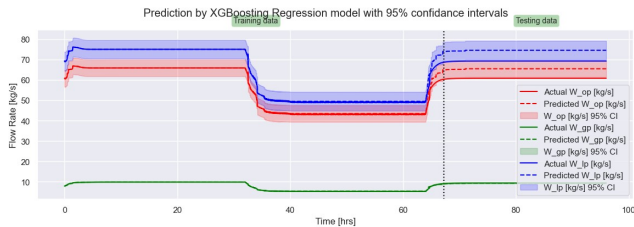


Fig. 11. Confidence intervals for XGBoost model (Well 2)

#### 4. CONCLUSION

Applying machine learning for flow rate estimation in oil and gas productions is a complex process. From the data collection to uncertainty quantification, considerable work has to be done to obtain useful results. The applicability of the results depends on the situation. Long short-term memory (LSTM) provides the best results with Mean absolute percentage error of 1.96% for Well 1 and 1.56% for Well 2. It may be best to use the predictions from the models as a backup for more robust systems. Since each well has its own characteristics, they must be modeled individually. In addition more process data would probably improve the accuracy of the flow rate predictions.

More filters can be tested to remove measurement noise. Different methods of uncertainty quantification can also be tested. The outlier detection and correction can be added in future. Unsupervised techniques like Local Outlier Factor, Isolation Forest, Kernel Density Estimation can be tested. Data reconciliation can also be added. Here the process flow diagram is necessary, the constraints of the each well are also needed.

#### REFERENCES

- AL-Qutami, T.A., Ibrahim, R., Ismail, I., and Ishak, M.A. (2018). Virtual multiphase flow metering using diverse neural network ensemble and adaptive simulated annealing. *Expert Systems with Applications*. doi:10.1016/j.eswa.2017.10.014.
- Andrianov, N. (2018). A machine learning approach for virtual flow metering and forecasting. *IFAC-PapersOnLine*. doi:10.1016/j.ifacol.2018.06.376.
- Bello, O., Ade-Jacob, S., and Yuan, K. (2014). Development of hybrid intelligent system for virtual flow metering in production wells. In *All Days, SPE-167880-MS. SPE*. doi:10.2118/167880-MS.
- Bikmukhametov, T. and Jäschke, J. (2020). First principles and machine learning virtual flow metering: A literature review. *Journal of Petroleum Science and Engineering*, 184, 106487. doi:10.1016/j.petrol.2019.106487.
- Boulesteix, A.L. and Strimmer, K. (2006). Partial least squares: a versatile tool for the analysis of high-

dimensional genomic data. *Briefings in Bioinformatics*, 8(1), 32–44. doi:10.1093/bib/bbl016.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi:10.1145/2939672.2939785. URL <https://arxiv.org/abs/1603.02754>.

Dayev, Z., Yetilmmezsoy, K., Sihag, P., Bahramian, M., and Kiyan, E. (2023). Modeling of the mass flow rate of natural gas flow stream using genetic/decision tree/kernel-based data-intelligent approaches. *Flow Measurement and Instrumentation*, 90, 102331. doi:10.1016/j.flowmeasinst.2023.102331.

Géron, A. (2023). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly, Beijing Boston Farnham Sebastopol Tokyo, third edition edition.

Janatian, N., Jayamanne, K., and Sharma, R. (2022). Model based control and analysis of gas lifted oil field for optimal operation. In *The First SIMS EUROSIM Conference on Modelling and Simulation, SIMS EUROSIM 2021, and 62nd International Conference of Scandinavian Simulation Society, SIMS 2021, September 21-23, Virtual Conference, Finland*, 241–246. doi:10.3384/ecp21185241.

Kniazeev, V., Erofeev, A., Demidov, A., Orlov, D., and Koroteev, D. (2023). Advanced well stimulation selection with gradient boosting. *Geoenergy Science and Engineering*, 228, 212026.

Pelz, P.F., Pfetsch, M.E., Kersting, S., Kohler, M., Matei, A., Melz, T., Platz, R., Schaeffner, M., and Ulbrich, S. (2021). *Mastering Uncertainty in Mechanical Engineering*. Springer International Publishing. doi:10.1007/978-3-030-78354-9\_2.

Qiu, J. and Toral, H. (1993). Three-phase flow-rate measurement by pressure transducers. In *All Days, SPE-26567-MS. SPE*. doi:10.2118/26567-MS.

Smola, A.J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222. doi:10.1023/B:STCO.0000035301.49549.88.

Xu, L., Zhou, W., and Li, X. (2012). Wet gas flow modeling for a vertically mounted venturi meter. *Measurement Science and Technology*, 23(4), 045301. doi:10.1088/0957-0233/23/4/045301.

Zangl, G., Hermann, R., and Schweiger, C. (2014). Comparison of methods for stochastic multiphase flow rate estimation. In *All Days, SPE-170866-MS*. doi:10.2118/170866-MS.

Zheng, H., Mirlekar, G., and Nord, L.O. (2022). Agent-based and stochastic optimization incorporated with machine learning for simulation of postcombustion co2 capture process. *Processes*, 10(12). doi:10.3390/pr10122727. URL <https://www.mdpi.com/2227-9717/10/12/2727>.

#### APPENDIX A. PROGRAM CODES

The Matlab codes for the simulator and the python machine learning code can be accessed here:

<https://github.com/dsouzaneville/FMH606-1-Masters-Thesis>

# Simulation and Cost Estimation of CO<sub>2</sub> Capture with Alternatives for Doubled Capacity

Lars Erik Øi, Masoumeh Dehghanizadeh and Nils Eldrup

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
Norway (Tel: 47-353557; e-mail: [lars.oi@usn.no](mailto:lars.oi@usn.no))*

**Abstract:** This study presents a techno-economic assessment of an amine-based carbon capture technology. The aim is to compare different methods to evaluate the cost effect of doubling the capacity. A base case was established in Aspen HYSYS with 15 m absorber packing height, 6 m desorber packing height, removal efficiency of 85 % and a heat exchanger minimum temperature approach ( $\Delta T_{\min}$ ) of 10 °C. In a first additional case the flue gas flow rate was doubled and in the second case a new absorber in parallel was added. Then dimensioning and cost estimation was carried out using Aspen HYSYS spreadsheets to automatically calculate CAPEX, OPEX and carbon capture cost per ton CO<sub>2</sub> captured. To estimate the Bare Erected Cost (BEC), the Enhanced Detailed Factor (EDF) and the Aspen Process Economic Analyzer (APEA) were employed. The EDF method determines the installation cost of each piece of equipment, while the Nazir-Amini method only offers the Total Plant Cost (TPC) without calculating individual equipment. Applying the EDF method, the TPC for the base case, the doubled feed gas case and the two-absorber case were calculated to 76, 141 and 150 MEuro respectively. This illustrates that cost increase may be less than proportional to the flow rate increase. The estimated annual OPEX for the base case was 42.5 MEuro, while for the two alternatives the OPEX was very close to the double of the base case. The estimated carbon capture cost for the base case, two-absorber case, and double feed gas scenario were 52.4 €/ton, 51.8 €/ton, and 50.5 €/ton, respectively. The study demonstrates that a combination of Aspen HYSYS simulation, Aspen Process Economic Analyzer and the EDF method is an effective method to evaluate different alternatives for increasing the capacity.

*Keywords: Carbon capture, Aspen HYSYS, simulation, dimensioning, cost estimation.*

## 1. INTRODUCTION

### 1.1. Aim

The first aim of this work is to compare different methods to cost estimate a CO<sub>2</sub> capture process based on process simulation. The second aim is to evaluate how efficient the different tools are to calculate different process alternatives and especially evaluate the cost effect of doubling the feed gas capacity of the CO<sub>2</sub> capture process.

### 1.2. Literature

There are several tools available to perform cost estimation of a process simulated in a process simulation tool like Aspen HYSYS or Aspen Plus. The Aspen Process Economic Analyzer (APEA) is a tool that is a part of Aspen HYSYS and Aspen Plus. An alternative is factor based methods, like different detailed factor methods like the Enhanced Detailed Factor (EDF) method developed at USN (Ali, 2019; Aromada et al., 2021).

Much general work has been published on cost estimation of CO<sub>2</sub> capture plants (Rubin et al., 2013; van der Spek et al., 2019; Roussanaly et al., 2019), but in these methods the cost estimation is traditionally performed independent of a process simulation tool. Other publications presenting both process simulation and cost estimation are (Mores et al., 2012; Agbonghae et al., 2014; Manzolini et al., 2015; Luo and Wang,

2016; Eldrup et al., 2019 and Hasan et al., 2020). A traditional limitation for the efficiency of the cost estimation in these references, is that the cost estimation is performed for each case with added specifications for each specific case.

For CO<sub>2</sub> capture, a focus at USN has been on automatic process simulation combined with cost estimation in Aspen HYSYS (Øi et al., 2021; Øi et al., 2022; Shirdel et al., 2022). This work is based on the Master Thesis of Masoumeh Dehghanizadeh (2023). In this work, it is aimed to compare the accuracy and efficiency of different tools for combined simulation and cost estimation.

### 1.3. Process Description

Figure 1 illustrates a typical CO<sub>2</sub> absorption process using amine-based systems. The CO<sub>2</sub> rich gas is first cooled in a direct contact cooler (DCC) and the CO<sub>2</sub> is then absorbed into the monoethanol (MEA) solvent and removed from the gas stream in an absorber. The CO<sub>2</sub> rich solvent is then pre-heated and pumped into a desorber column, where it is heated and the CO<sub>2</sub> is stripped off the CO<sub>2</sub>. The regenerated solvent is recycled to the absorber tower, while the high purity CO<sub>2</sub> stream off the top of the desorber column is sent further to processing for transportation and storage. If the flue gas capacity is doubled, another absorber can be set in parallel to the one in Fig. 1.

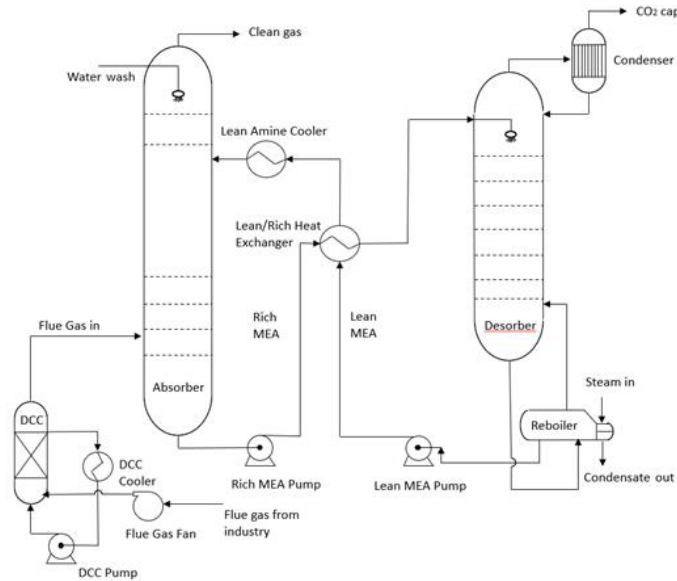


Fig. 1. Process flow diagram of a standard amine-based CO<sub>2</sub> capture process (Aromada et al., 2020).

## 2. SPECIFICATION AND SIMULATIONS

### 2.1. Specifications and simulation of base case CO<sub>2</sub> capture process

The Aspen HYSYS V12 was used to simulate an amine-based CO<sub>2</sub> capture process. The Acid Gas property package was employed, which includes the electrolyte non-random two-liquid (e-NRTL) model for electrolyte thermodynamics and the Peng-Robinson equation of state for the vapor phase.

The absorber and desorber were simulated using equilibrium stages with Murphree stage efficiencies. The Murphree efficiency is defined by dividing the change in CO<sub>2</sub> mole fraction from one stage to the next by the change on the assumption of equilibrium.

The specifications for the base case Aspen HYSYS simulation are given in Table 1. These specifications give a 85 per cent CO<sub>2</sub> removal efficiency and a minimum approach temperature of 10 °C in the lean/rich amine heat exchanger. The simulation is similar to earlier studies (Øi, 2007; Aromada et al., 2021 and Shirdel et al., 2022). The absorber has 15 stages with Murphree efficiency 0.15, while the desorber has 10 stages with Murphree efficiency 0.5. In the columns, the Modified HYSIM Inside-Out numerical solver was selected. The adiabatic efficiency of the pumps was specified to be 75%.

### 2.2. Calculation sequence

The calculation sequence in the Aspen HYSYS flowsheet in Figure 2 is similar to the simulations in Øi et al. (2022) and Shirdel et al. (2022). In the base case there is only one recycle block to check that the amine liquid flow recirculated is equal to the inlet flow to the absorption column. Figure 3 illustrates the process flow diagram of the two-absorber scenario with two simulated absorption columns.

Table 1: Specifications for the base case alternative

<i>Parameter</i>	<i>Value</i>
Inlet flue gas temperature [°C]	40.0
Inlet flue gas pressure [kPa]	110
Inlet flue gas flow rate [kmol/h]	85000
CO <sub>2</sub> content in inlet gas [mole %]	3.73
Water content in inlet gas [mole %]	6.71
Lean amine temperature [°C]	40.0
Lean amine pressure [kPa]	110.0
Lean amine rate [kg/h]	103500
MEA content in lean amine [mass %]	29
CO <sub>2</sub> content in lean amine [mass %]	5.4
Number of stages in absorber [-]	15
Murphree efficiency in absorber	0.15
Rich amine pump pressure [kPa]	200.0
Rich amine temp. out of HEX [°C]	104
Number of stages in desorber [-]	6
Murphree efficiency in desorber	1.0
Reflux ratio in stripper [-]	0.3
Reboiler temperature [°C]	120.0

### 2.3. Equipment dimensioning

The dimensioning of all the equipment (except for the DCC unit) were performed as in previous studies (Øi et al., 2022; Shirdel et al., 2022). The diameters of the absorption and desorption columns were evaluated from the gas volumetric flows and based on a superficial gas velocities of 2.5 m/s for the absorber and 1 m/s for the desorber column. Each packing stage in the absorber and desorber was assumed to be 1 m high. To include the height for packing, liquid distributors, water wash, demister, gas inflow, gas outflow and sump, the total column heights are considerably larger, and set to 30 m and 16 m for the tangent-to-tangent heights of the absorber and desorber column.

The overall heat transfer coefficients specified are 1.20 kW/(m<sup>2</sup>·K) for the reboiler, 0.73 kW/(m<sup>2</sup>·K) for the lean/rich heat exchanger, 0.80 kW/(m<sup>2</sup>·K) for the amine cooler, and 1.00 kW/(m<sup>2</sup>·K) for the condenser as in Aromada et al. (2022). The pumps had an efficiency of 0.75. It was assumed that the maximum heat exchanger size is 1000 m<sup>2</sup>, and in case of the need for larger heat exchanger area, more units are necessary.

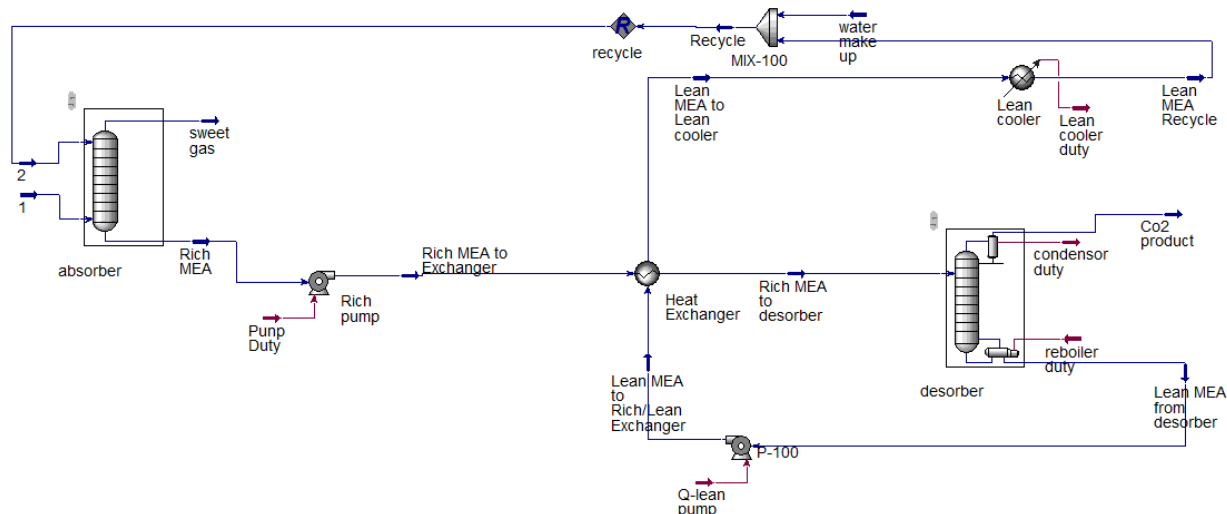


Fig. 2. Aspen HYSYS flowsheet for the base case (Dehghanizadeh, 2023)

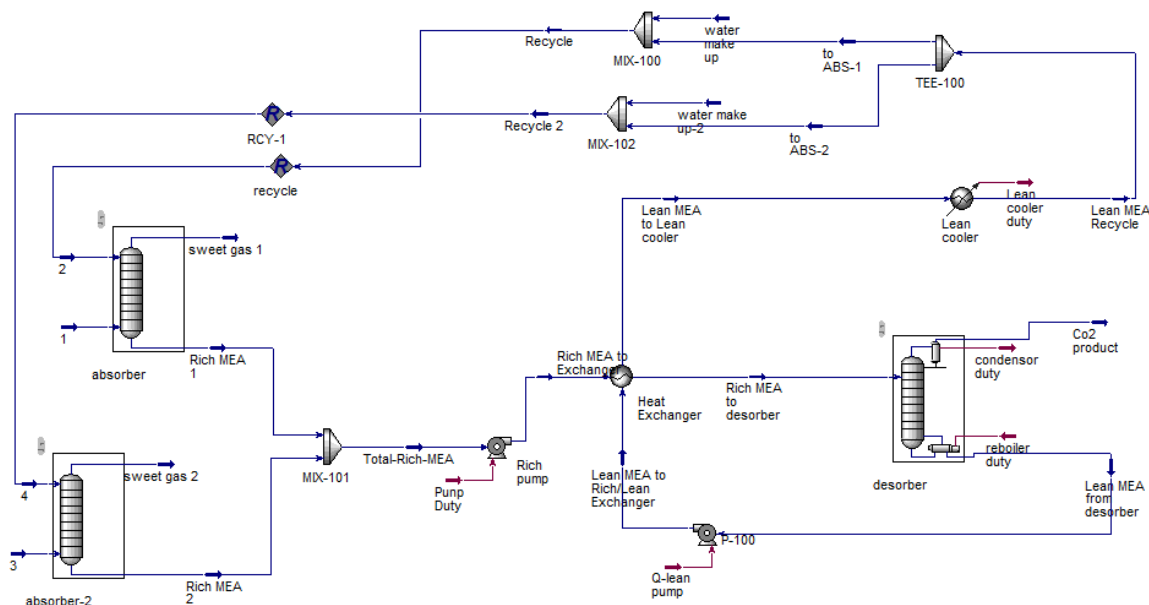


Fig. 3. Aspen HYSYS flowsheet for the case with two simulated absorption columns (Dehghanizadeh, 2023)

### 3. COST ESTIMATION PROCEDURES AND ASSUMPTIONS

#### 3.1. Capital cost estimation method

The purchased cost of each equipment unit is estimated in this work with Aspen In-Plant Cost Estimator based on the dimensioning.

After estimating the cost of each part of equipment, cost factors are added to obtain the quantities Bare Erected Cost (BEC) and Total Plant Cost (TPC). A description of what is traditionally included in the BEC, and what is traditionally

included in the TPC is presented and discussed in Rubin et al. (2013). The BEC and TPC are defined differently in different literature (Rubin et al, 2013). In this work the BEC is determined using the EDF method by creating a detailed list of all the process equipment, obtaining estimates on purchased equipment cost and estimating all the cost of material and labour required to complete the installation. It includes the cost of equipment, erection, piping, electro, instrument, ground work, steel and concrete, insulation and engineering. To obtain the TPC, contractor services, process contingency and project contingency is also included. Then the cost of the equipment is adjusted to the correct size, year,

and material of construction. This method is documented in (Ali, 2019 and Aromada et al., 2021).

Another approach to estimate BEC employed in this study is the Aspen Process Economic Analyzer (APEA). It relies on model-based estimation to generate project cost estimates. The APEA can calculate not only the equipment cost but also the installed direct cost (piping, civil, structural steel, insulation, etc.) for each process equipment. The equipment cost calculated using APEA and Aspen In-Plant has in this work been compared, and the results were very similar.

To calculate the TPC, the Nazir-Amini method (Nazir et al., 2018) was used as an alternative to the EDF method. In this work 10 % of BEC is added for engineering procurement construction cost, 10 % for process contingency and 15 % for project contingency is added to obtain the TPC.

The cost currency and cost year were Euro (€) and 2019 for Aspen In-Plant and 2020 for the detailed factor table. The default location in Aspen In-Plant Cost Estimator, Rotterdam, was assumed in this work. The equipment units were assumed to be constructed from stainless steel SS316. The material factor for welded equipment was 1.75 and 1.30 for machined equipment.

The total installation cost factor includes the sub-factors for direct costs, engineering costs, administration costs, and commissioning and contingency costs. Equation (1) is used to calculate the total installation factor in carbon steel ( $F_{T,CS}$ ). The procedure of utilizing the EDF method for TPC calculation corresponds to the methodology outlined in Ali (2019).

$$F_{T,CS} = f_{direct} + f_{eng} + f_{adm} + f_{comm} + f_{cont} \quad (1)$$

where the subscripts in the factors means direct installation cost, engineering, administration, commissioning and contingency. The individual factors are in this work from an EDF table sheet in (Aromada et al., 2021).

The total equipment installed cost (EIC) for each unit in carbon steel can be calculated from Equation (2).

$$EIC_{CS} = F_{T,CS} \times Equipment\ cost_{CS} \quad (2)$$

Total plant cost is the sum of the total installation costs for each equipment unit and is calculated by Equation (3). In the case of calculating BEC, the factors for administration, commissioning and contingency is omitted in Equation (1).

$$TPC(2019) = \sum EIC(all\ equipment) \quad (3)$$

If the equipment is to be made of a material other than carbon steel, the installation factor must be adjusted accordingly. Equation (4) is used to make this correction:

$$F_T = [F_{T,CS} + (f_{mat} - 1) \times (1 + f_{T,pp,CS})] \quad (4)$$

where  $f_{mat}$  is the material factor which is the ratio between the unit cost and the unit cost in carbon steel, and  $f_{pp,CS}$  is the piping factor (for carbon steel) in the EDF table sheet.

The capital cost of the CO<sub>2</sub> capture plant is then escalated from 2019 using a consumer cost index from Statistisk Sentralbyrå (SSB).

During optimization or sensitivity analysis, where a parameter is varied, the capacities/sizes of some equipment will change. Therefore, there is a need to estimate new cost for the equipment units due to the resulting changes in size/capacity. This is automatically estimated based on the Power law using an exponent of typically 0.65 based on the previous cost obtained from Aspen In-Plant Cost Estimator as done in (Aromada et al., 2022; Øi et al., 2022).

$$Cost_{NEW} = Cost_{OLD} \times \left(\frac{Capacity_{NEW}}{Capacity_{OLD}}\right)^{0.65} \quad (5)$$

### 3.2. Operating cost estimation and assumptions

The annual operating cost in this work is the sum of the fixed operating cost and variable operating costs estimated as in Øi et al. (2022):

$$Annual\ cost = Consumption \times Unit\ cost \quad (6)$$

The assumptions used for estimating the annual operating cost are presented in Table 2. The values are similar to values used in earlier work like Aromada et al. (2021).

Table 2: Annual operating cost assumptions

<i>Item</i>	<i>Unit</i>	<i>Value</i>
Operating lifetime	[Year]	22(2+20)
Annual hours of operation	[h/year]	8000
Electricity cost	[€/kWh]	0.132
Steam cost	[€/kWh]	0.032
MEA cost	[€/ton]	1450
Maintenance cost	[€/year]	4% of CAPEX
Operator cost (6 oper)	[€/year]	85350(*6)
Engineer cost (1 eng)	[€/year]	166400

### 3.3. CO<sub>2</sub> capture annualized cost

An economic key performance indicator in this work is CO<sub>2</sub> captured cost. This was estimated using Equations (7) to (10), as shown in (Aromada et al., 2021):

$$CO_2\ captured\ cost = \frac{Total\ annual\ cost}{Mass\ of\ Captured\ CO_2/year} \quad (7)$$



$$Total\ annual\ cost = Annualized\ CAPEX + Yearly\ OPEX \quad (8)$$

$$Annualized\ CAPEX = \frac{CAPEX}{Annualized\ factor} \quad (9)$$

$$Annualized\ factor = \sum_{i=1}^n \left[ \frac{1}{(1+r)^i} \right] \quad (10)$$

where n is the plant lifetime, 22 years which includes 2 years for the plant’s construction. And r is the discount rate and was assumed to be 7.5 %.

#### 4. RESULTS AND DISCUSSION

##### 4.1. BEC and TPC for the base case

Figure 4 shows the Bare Erective Cost (BEC) calculated with the EDF method and the Aspen Economic Analyzer. The contribution from each equipment unit is also shown in the figure. It shows that the absorber is the dominating part, and then the main heat exchanger. It also shows that the EDF method and the results from the Aspen Economic Analyzer (APEA) give reasonable close results (within 5-10 %).

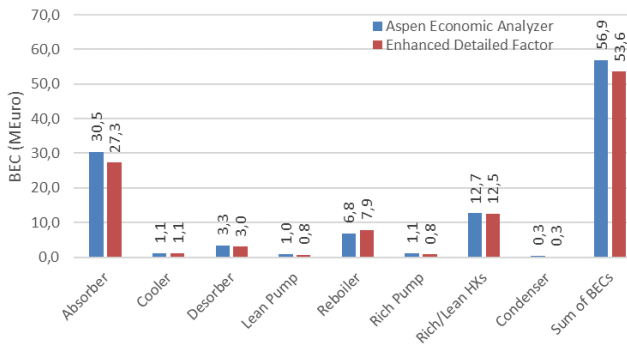


Fig. 4. BEC comparison for the Base Case applying the APEA and EDF method (Dehghanizadeh, 2023)

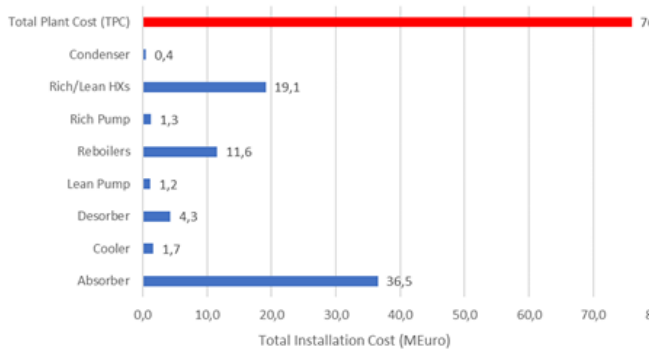


Fig. 5. Total installation cost (TPC) applying the EDF method (Dehghanizadeh, 2023)

Figures 5 and 6 show Total Project Cost (TPC) and OPEX for the base case. The values for the TPC are higher, but the

cost distribution is similar to the BEC values. For OPEX, steam for heating is as expected the most significant part.

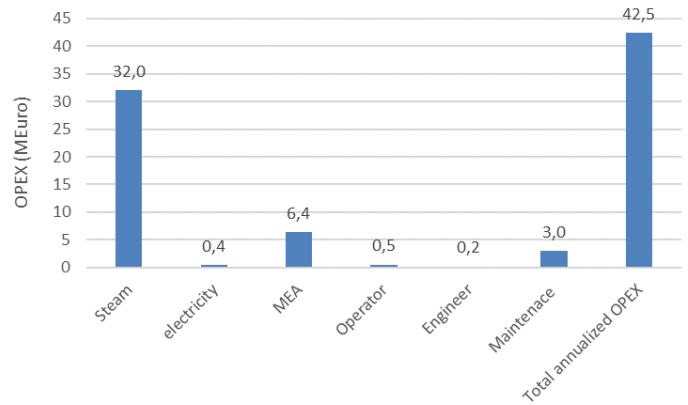


Fig. 6. OPEX estimated for the Base Case (Dehghanizadeh, 2023)

##### 4.2. Results for doubled capacity

Figures 7 and 8 show a BEC comparison for the Base case and Doubled feed gas applying the EDF and APEA methods. It shows that the absorber cost and total cost increases a little less than to the double cost.

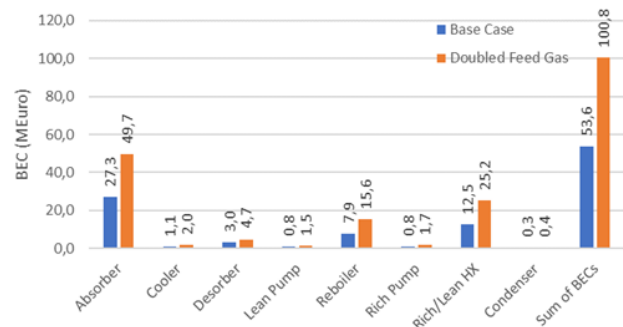


Fig. 7. BEC comparison for the Base case and Doubled feed gas applying the EDF method (Dehghanizadeh, 2023)

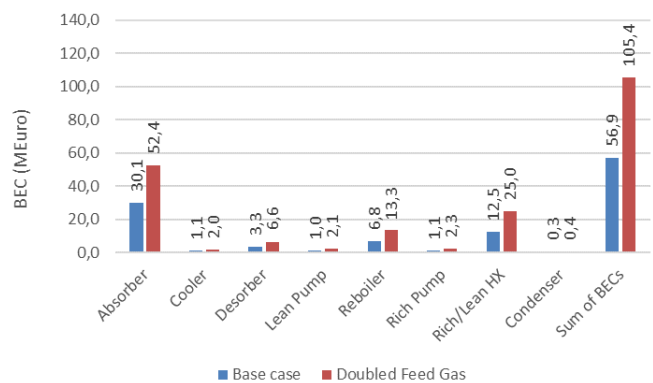


Fig. 8. BEC comparison for the Base case and Doubled feed gas case applying the APEA method (Dehghanizadeh, 2023)



Figure 9 is showing the TPC for the Doubled feed gas by the EDF method. It shows that the increase is lower than the double of the TPC for the base case, 140.4 compared to 76 MEuro.

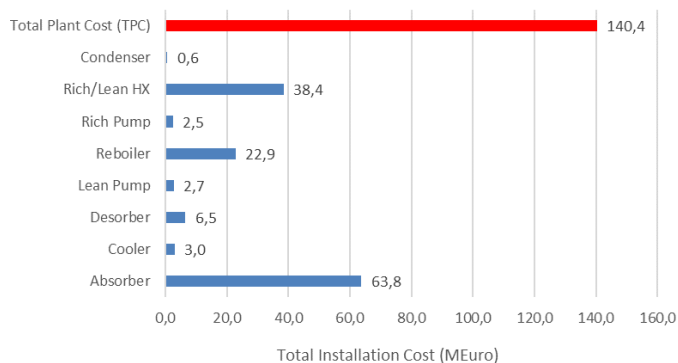


Fig. 9. TPC for Doubled feed gas using the EDF method (Dehghanizadeh, 2023)

Figure 10 shows a comparison of TPC calculated by the EDF method, the Nazir-Amini method and use of the power law. It shows that the EDF method and Nazir-Amini method are close, while applying the power law is considerably lower. It is expected that the use of the EDF method is the most accurate because it is more detailed.

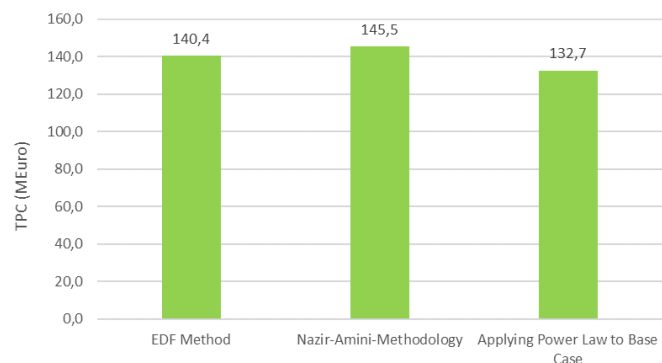


Fig. 10. TPC for Doubled Feed gas applying the EDF method (Dehghanizadeh, 2023)

### 4.3. Results for two-absorber scenario

Figures 11 and 12 show calculated BEC for the Doubled feed case and the Two-absorber case calculated by the EDF method (Fig. 11) and the APEA method (Fig. 12). Both methods show that the Two-Absorber case is slightly more costly.

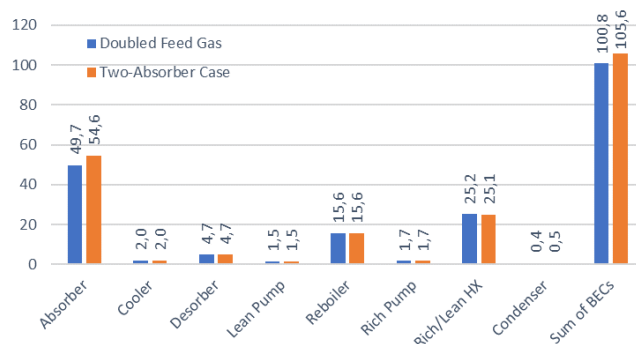


Fig. 11. Comparison of BEC for Doubled feed gas case and Two-absorber case applying the EDF method (Dehghanizadeh, 2023)

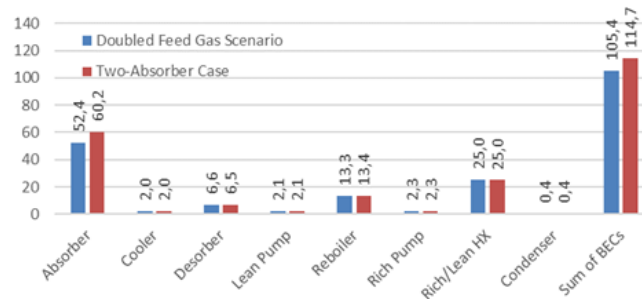


Fig. 12. Comparison of BEC for Doubled feed gas case and Two-absorber case applying the APEA method (Dehghanizadeh, 2023)

Figure 13 shows the TPC calculated for the Two-absorber case. Compared to the calculation for the Doubled feed case in Fig. 9, it is considerably more expensive (149.7 compared to 140.4).

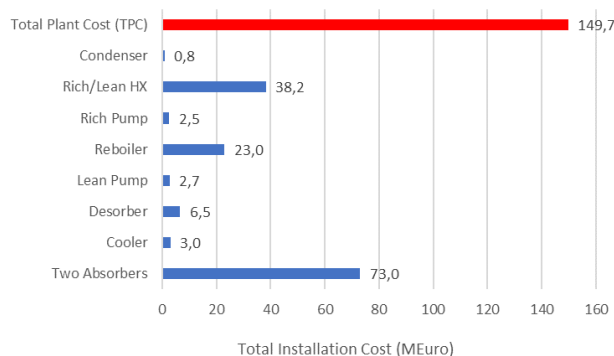


Fig. 13. TPC for Doubled feed gas applying the EDF method (Dehghanizadeh, 2023)

Figure 14 shows a comparison of TPC calculated by the EDF method, the Nazir-Amini method and use of the power law. It shows that the methods give close to the same results, but the Nazir-Amini method gives the highest value. It is expected that the EDF method is most accurate for sensitivity calculations of different parameters because it is more detailed. The Nazir-Amini additional factors are very general.

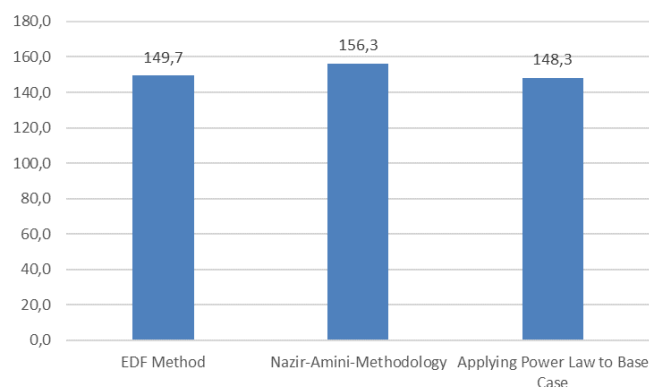


Fig. 14. TPC for Two-absorber case calculated by different methods (Dehghanizadeh, 2023)

Using the power law gives approximately the same result compared to the more detailed EDF method. The cost is as expected higher (about 8 %) than for the Doubled feed scenario.

Figure 15 shows the TPC as a function of gas flow, using the EDF method. The sensitivity calculation is performed automatically by using a Case study function in Aspen HYSYS. It shows that the cost is increasing a little less than proportional to the gas flow. This is as expected because the power law is used, which specifically calculates a cost for the absorber less than proportional to the gas flow. It shows that it is straightforward to use the EDF method using the power law for a fast and automatic sensitivity analysis.

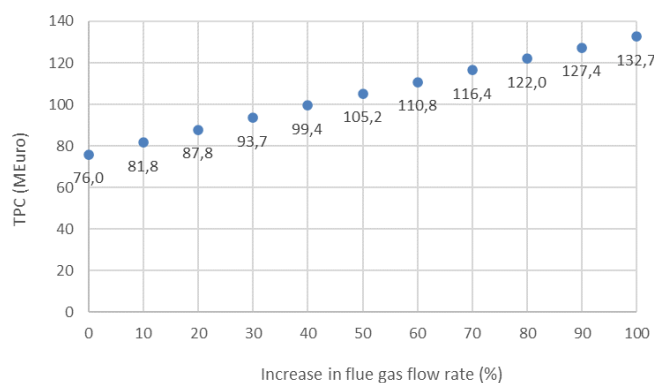


Fig. 15. Impact of flue gas increase on the TPC using the EDF method and the power law for equipment cost estimation at changed capacity (Dehghanizadeh, 2023)

#### 4.4. Uncertainties in the cost estimation and cost comparison

The uncertainties in the absolute value of calculated CAPEX and OPEX for each case is large. An uncertainty for the CAPEX of +/- 50 % has been suggested (Ali, 2019). The uncertainty of the OPEX is even larger, especially due to the

high uncertainty in heat and electricity cost (Aromada et al, 2021). However, in this work accurate absolute values of CAPEX and OPEX was not the main aim.

The main aim in this work was to make efficient cost comparisons of specific process alternatives using different tools. The results indicate that the results from the different calculation methods, especially the EDF and APEA methods, gave quite similar values. When comparing the process alternatives, the EDF and the APEA methods estimate similar cost differences between the double feed gas and two-absorber case. The higher cost of the two-absorber alternative is regarded to be significant, even though the absolute values of the cost estimates are inaccurate.

## 5. CONCLUSIONS

A base case was established in Aspen HYSYS with 15 m absorber packing height, 6 m desorber packing height, removal efficiency of 85 % and a heat exchanger minimum temperature approach ( $\Delta T_{min}$ ) of 10 °C. In a first additional case the flue gas flow rate was doubled and in the second case a new absorber in parallel was added. Then dimensioning and cost estimation was carried out using Aspen HYSYS spreadsheets to automatically calculate CAPEX and OPEX and carbon capture cost per ton CO<sub>2</sub> captured. To estimate the Bare Erected Cost (BEC), the Enhanced Detailed Factor (EDF) and the Aspen Process Economic Analyzer (APEA) were employed. The EDF method determines the installation cost of each piece of equipment, while the Nazir-Amini method only offers the TPC without calculating individual equipment. Applying the EDF method, the TPC (CAPEX) for the base case, the doubled feed gas case and the two-absorber case were calculated to 76, 141 and 150 MEuro respectively. This illustrates that the cost increase may be less than proportional to the flow rate increase. The estimated annual OPEX for the base case is about 42.5 MEuro, while for the two alternatives the OPEX was very close to the double of the base case. The estimated carbon capture costs for the base case, two-absorber case, and double feed gas scenario were 52.4 €/ton, 51.8 €/ton, and 50.5 €/ton, respectively. The study demonstrates that a combination of Aspen HYSYS simulation, Aspen Process Economic Analyzer and the EDF method is an effective method to evaluate different alternatives for increasing the capacity.

## REFERENCES

- Agbonghae, E.O., Hughes, K.J., Ingham, B., Ma, L. and Pourkashanian, M. (2014). Optimal Process Design of Commercial-Scale Amine-Based CO<sub>2</sub> Capture Plants. *Industrial & Engineering Chemistry Research*, 53, 14815-14829.
- Ali, H. (2019). *Techno-economic analysis of CO<sub>2</sub> capture concepts*. PhD Thesis, University of South-Eastern Norway.

- Aromada, S.A. and Øi, L.E. (2017). Energy and economic analysis of improved absorption configurations for CO<sub>2</sub> capture. *Energy Procedia*, 114, 1342-51.
- Aromada, S.A., Eldrup, N.H., Normann, F. and Øi, L.E. (2020). Simulation and Cost Optimization of different Heat Exchangers for CO<sub>2</sub> Capture. Linköping Electronic Conference Proceedings, SIMS 61, pp. 22-24. doi:10.3384/ecp20176318
- Aromada, S.A., Eldrup, N.H. and Øi, L.E. (2021). Capital cost estimation of CO<sub>2</sub> capture plant using Enhanced Detailed Factor (EDF) method: Installation factors and plant construction characteristic factors. *International Journal of Greenhouse Gas Control*. 110, 103394.
- Aromada, S.A., Eldrup, N.H. and Øi, L.E. (2022). Cost and Emissions Reduction in CO<sub>2</sub> Capture Plant Dependent on Heat Exchanger Type and Different Process Configurations: Optimum Temperature Approach Analysis. *Energies*, 15(2), 425.
- Dehghanizadeh, S. (2023). *Simulation and Cost Estimation of CO<sub>2</sub> Capture Using Aspen HYSYS or Aspen Plus*. Master Thesis, University of South-Eastern Norway.
- Eldrup, N., Mathisen, A., Skagestad, R. and Haugen, H.A. (2019). A Cost Estimation Tool for CO<sub>2</sub> Capture Technologies. SSRN Scholarly paper ID 3366036, doi: 10.2139/ssrn.3366036
- Hasan, S., Abbas, A.J. and Nasr, G.G. (2020). Improving the Carbon Capture Efficiency for Gas Power Plants through Amine-Based Absorbents. *Sustainability*, 13(1), doi: 10.3390/su13010072
- Luo, X and Wang, M. (2016). Optimal operation of MEA-based post-combustion carbon capture for natural gas combined cycle power plants under different market conditions. *International Journal of Greenhouse Gas Control*, 48, 312-320. doi: 10.1016/j.ijggc.2015.11.014.
- Manzolini, G., Sanches Fernandez, E., Rezvani, S., Macchi, E., Goetheer, E.I.V. and Vlught, T.J.H. (2015). Economic assessment of novel amine based CO<sub>2</sub> capture technologies integrated in power plants based on European Benchmarking Task Force methodology. *Applied Energy*, 138, 546-548.
- Mores, P., Rodriguez, N., Scenna, N. and Mussati, S. (2012). CO<sub>2</sub> capture in power plants: Minimization of the investment and operating cost of the post-combustion process using MEA aqueous solution. *International Journal of Greenhouse Gas Control*, 10, 148-163.
- Nazir, S.M., Morgado, J.F., Bolland, O., Quinta-Ferreira, R. and Amini, S. (2018). Techno-economic assessment of chemical looping reforming of natural gas for hydrogen production and power generation with integrated CO<sub>2</sub> capture. *International Journal of Greenhouse Gas Control*, 78, 7–20. doi: 10.1016/j.ijggc.2018.07.022.
- Roussanaly, S. (2019). Calculating CO<sub>2</sub> avoidance costs of Carbon Capture and Storage from industry. *Carbon Management.*, 10(1), 105–112. doi: 10.1080/17583004.2018.1553435.
- Rubin, E.S., Short, C., Booras, G., Davison, J., Ekstrom, C., Matuszewski, M. and McCoy, S. (2013). A proposed methodology for CO<sub>2</sub> capture and storage cost estimates. *International Journal of Greenhouse Gas Control*, 17, 488-503.
- Shirdel, S., Valand, S., Fazli, F., Winther-Sørensen, B., Aromada, S.O., Karunaratne, S. and Øi, L.E. (2022). Sensitivity Analysis and Cost Estimation of a CO<sub>2</sub> Capture Plant in Aspen HYSYS. *Chem. Engineering*, 6(2), 28.
- Van der Spek, M., Roussanaly, S. and Rubin, E.S. (2019). Best practices and recent advances in CCS cost engineering and economic analysis. *International Journal of Greenhouse Gas Control*, 83, 91-104.
- Øi, L.E. (2007). Aspen HYSYS simulation of CO<sub>2</sub> removal by amine absorption from a gas based power plant. In The 48th Scandinavian Conference on Simulation and Modeling (SIMS 2007); *Linköping University Electronic Press*, pp. 73-81.
- Øi, L.E., Haukås, A., Aromada, S.A. and Eldrup, N. (2021). Automated Cost Optimization of CO<sub>2</sub> Capture Using Aspen HYSYS. Linköping Electronic Conference Proceedings SIMS 62, pp. 293-300. doi: 10.3384/ecp21185293
- Øi, L.E., Eldrup, N., Aromada, S., Haukås, A., Helvig, J., Hæstad, I. and Lande, A.M. (2022). Process simulation and automated Cost Optimization of CO<sub>2</sub> Capture Using Aspen HYSYS. Linköping Electronic Conference Proceedings SIMS 63, pp. 65-71.

# Simulation of Biogenic Carbon Capture and Utilization Process Chain

K. Tiiro. M. Ohenoja. O. Ruusunen. R. L. Keiski. M. Ruusunen

*Environmental and Chemical Engineering Research Unit, University of Oulu, 90570 Oulu, Finland,  
(e-mail: forename.surname@oulu.fi).*

**Abstract:** Carbon capture and utilization (CCU) is a growing field in chemical engineering with high expectations to replace fossil carbon. This paper focuses on modeling and simulation of a CCU process chain utilizing biogenic CO<sub>2</sub>. A scenario with a pulp mill recovery boiler effluent is assumed. CO<sub>2</sub> capture is performed with a membrane-based system. This is followed by methanol synthesis, and the majority of produced methanol is directed to dimethyl carbonate (DMC) synthesis.

The process chain with fixed process design was simulated for different scenarios of the flue gas properties. The key process indicators were observed. Further, the flexibility of the processes was evaluated to mitigate the changes in process indicators due to fluctuating flue gas properties. Finally, model parameter uncertainties and modeling assumptions were discussed. The results indicate the level of uncertainties of CCU models and their key process indicators that should be considered when moving on to the system level simulations and techno-economic or life cycle analyses.

*Keywords: Process modeling, Membrane separation, Methanol synthesis, Dimethyl carbonate production, Sensitivity analysis*

## 1. INTRODUCTION

Carbon capture has been identified as an important tool for managing carbon emissions for a long time (Reichle et al., 1999). Early studies have focused on CO<sub>2</sub> capture from fossil-based power generation and industrial point sources, such as steel mills. More recently, emphasis has also been given to CO<sub>2</sub> sources with biogenic origin, such as from biogas upgrading processes, fermentation, pulp mills and biomass fueled power plants (Rodin et al., 2020). Future possibilities also involve direct carbon capture from air as the capture technologies are developing to be more feasible also with minimal CO<sub>2</sub> contents and availability of green energy is increasing (Akimoto et al., 2021).

There are many possibilities for utilizing the captured CO<sub>2</sub> with the highest market potential existing in oil and chemical industry sectors (Koytsoumpa et al., 2018). Methanol is of particular interest in this study due to its possible use as both a fuel and as a chemical feedstock. As a fuel, methanol has advantages over directly combusting hydrogen, due to methanol's easier storability, better volumetric energy density, and compatibility to existing internal combustion engines and infrastructure (Gumber and Gurumoorthy, 2018).

Methanol can be further refined into a multitude of different hydrocarbon products including formaldehyde, methyl tert-butyl ether, acetic acid, methyl methacrylate, and dimethyl ether (Gumber and Gurumoorthy, 2018). Another derivative from methanol is dimethyl carbonate (DMC). DMC serves a multitude of purposes across diverse industries, such as a solvent for paints, coatings, and cleaning agents. DMC can be used as an additive in gasoline and diesel fuels, enhancing combustion efficiency. Moreover, DMC plays a pivotal role in the realm of energy storage, being utilized as an electrolyte component in advanced lithium-ion batteries (Kohli et al.,

2022). DMC also offers a non-carcinogenic alternative to commonly used chemicals like dichloromethane and dimethyl sulfate in carbonylation and transesterification reactions. (Wei et al., 2023)

Stemming from the choice of methanol and DMC as products of interest, the modeled process chain in this work includes a membrane-based CO<sub>2</sub> capture unit, methanol synthesis plant unit, and a DMC reactor unit. Membranes were chosen over the more popular amine-based approach, as it shows high energy efficiency and small physical and chemical footprint (Hou et al., 2022). High-level illustration of the process chain is visible in Fig. 1. The figure depicts that the process chain takes flue gas and hydrogen as raw material feeds, and outputs methanol and DMC as the main products. It is visible in Fig. 1 that the captured CO<sub>2</sub> stream is divided between methanol and DMC syntheses units and that some of the produced methanol is considered as direct product along with DMC. One advantage of this kind of synthesis chain is the ability to vary the ratio of selling methanol directly to refining it into DMC, depending on current market prices.

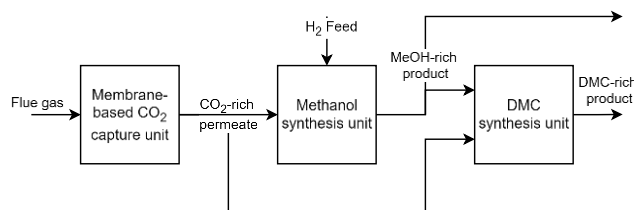


Fig. 1. Simplified diagram of the modeled CCU process chain.

The feasibility studies at system level are typically focusing on the economic profit of the production. Historical data is often available for estimating the market prices of raw materials,

products, heat, and especially the dynamical fluctuation of electricity (Karjunen et al., 2023). However, without incorporating unit level models, the system level modeling can often be limited to using constant estimates for product yield, net energy consumption, or side stream properties (Karjunen et al., 2023). Instead, these key process indicators (KPI) should preferably be treated as variables, with reasonable uncertainty ranges.

In order to assess the process performance under varying feed properties, technological limitations and uncertainties, this research aims to develop a simulation tool for the studied CCU process chain and to indicate ranges for the KPIs that can be used in system level modeling and risk assessment of an integrated CCU process. Incorporating unit level models into the system level modeling can increase the accuracy of the economical assessment by considering the effect of operation point dependencies and other sources of uncertainty in process KPIs.

## 2. MATERIAL AND METHODS

### 2.1 CO<sub>2</sub> capture

For the membrane-based CO<sub>2</sub> capture, the two-stage, pressure-driven process configuration and membrane material reported in (Asadi and Kazempoor, 2021) is considered. Namely, Polaris gen 1 membrane is assumed with permeability of CO<sub>2</sub> of 1000 GPU (gas permeation units, 1 GPU = 3.35·10<sup>-10</sup> mol/m<sup>2</sup>/s/Pa). The selectivity between CO<sub>2</sub> and N<sub>2</sub> is 50, and the selectivity between CO<sub>2</sub> and O<sub>2</sub> is 20 (Khalilpour et al., 2012). Hollow-fiber membranes with constant dimensions are assumed, and they are operated in counter-current flow. The feed is introduced to the shell side of the membrane (outside of the fibers, retentate), and the permeate is collected from inside the membranes (bore side). The operating temperature (*T*), feed pressure (*P*), initial permeate pressure, inlet molar flowrate, recycle ratio between the membrane stages (*RR*), and the feed gas composition can be manipulated as well.

The modeling approach follows the reported model in (Asadi and Kazempoor, 2021) with the following exceptions:

- In addition to CO<sub>2</sub> and N<sub>2</sub>, also O<sub>2</sub> balance is modeled.
- Feed (retentate) pressure is assumed to be constant.
- The dynamic viscosity of the gas mixture is calculated following (Wilke, 1950) applying absolute viscosities of the gas components at the operation temperature.

The process flow diagram is depicted in Fig. 2. The model also comprises the calculation of electricity consumption by the two compressors, which are modeled according to adiabatic compression equations in (Green and Southard, 2019). The process was sized by performing a constrained optimization with MATLAB<sup>®</sup> *fmincon*-function where the relative membrane area for the first stage, the recycle ratio of the second stage retentate to the first stage feed, and the operating pressure were determined. The second stage membrane area is assumed to be 3% of the first stage membrane area. The optimization objective was to minimize the specific energy consumption, with a penalty for deviating CO<sub>2</sub> purity from target 98%. This objective balances the CO<sub>2</sub> capture rate and

energy consumption while keeping the CO<sub>2</sub> quality deviation within a small tolerance.

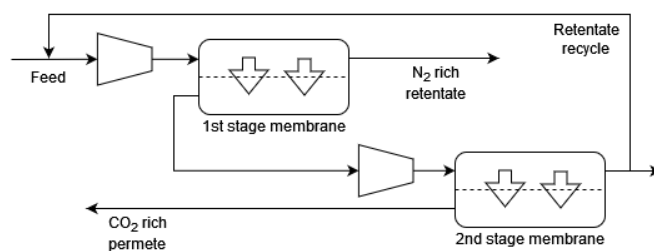
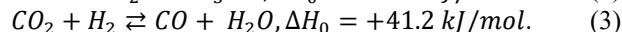
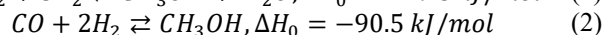
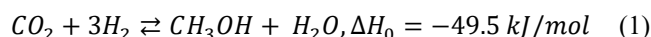


Fig. 2. Membrane-based CO<sub>2</sub> capture process.

### 2.2 Methanol synthesis

Methanol (MeOH) is generally possible to be synthesized through two overall reactions, which are the hydrogenation of CO<sub>2</sub> and CO (Poto et al., 2022), see (1) and (2). Reverse Water Gas Shift reaction (RWGS) is also present in the same reaction conditions (3).



Heterogeneous catalysis using a Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst, in either adiabatic or isothermal reactors, is the predominately used MeOH production method in industry (Bozzano and Manenti, 2016). Due to its dominant role, Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst was also chosen for the model of this study. Even though Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst has been researched for decades, the exact roles of different reaction pathways are still debated in the literature to this day (Azhari et al., 2022). Nevertheless, for this work, the kinetic model developed in (Bussche and Froment, 1996) was chosen, due to a recommendation in (Bozzano and Manenti, 2016). The chosen kinetic model dismisses CO hydrogenation (2) to form MeOH and does not include any side reactions besides RWGS (Bussche and Froment, 1996).

For the model of this work, a Lurgi-type tube-and-shell reactor was chosen. The gas phase reactants flow through the reactor tubes that are filled with solid catalyst pellets. The shell side of the reactor has pressurized cooling water to control the reaction temperature. The purpose of cooling is to prevent the maximum temperature inside the reactor tubes from exceeding 280 °C (553 K) as the catalyst deactivates faster at high temperatures (Hartig and Keil, 1993). A steady-state pseudo-homogeneous reactor model based on mass fractions was chosen from (Manenti et al., 2011) with the following assumptions:

- Ideal plug flow (constant axial velocity, negligible axial diffusion, perfect radial mixing) (Manenti et al., 2011).
- Homogeneous gas and solid phases inside the tubes (no temperature, pressure, or composition gradient within a catalyst particle or in the surrounding gas in radial direction) (Manenti et al., 2011).
- Catalyst particle efficiency modeled by the modified Thiele modulus (Lommerts et al., 2000).



- Pressure loss inside the reactor tube modeled by the Ergun equation (Manenti et al., 2011).
- Shell side at constant bulk temperature (Manenti et al., 2011).
- Heat transfer between tube and shell modeled as in (Hartig and Keil, 1993).
- Gas density inside the reactor tube given by the Peng-Robinson equation of state (PR-EoS) (Peng and Robinson, 1976).
- Knock out drum separator 1 (KO1) is assumed to be sized so that thermodynamic equilibrium between the liquid and gas phases is always reached. Phase equilibrium is solved using PR-EoS (Peng and Robinson, 1976), and the Rachford-Rice method (Green and Southard, 2019).
- Divisor 1 (DIV1) determines the recycle and purge stream flows based on the set maximum mass-based recycle ratio (mass flow of recycle stream divided by mass flow of feed stream). Any portion of KO1 gas effluent that exceeds the maximum recycle ratio is sent to purge stream.
- Compressor 1 (CP1) outlet pressure is fixed to the feed stream pressure. The temperature rise and electrical energy consumption of CP1 are modeled by adiabatic compression equations in (Green and Southard, 2019).

A process flow diagram of the modeled MeOH synthesis unit is presented in Fig. 3. In the figure, the sections surrounded by red dashed line representing the feed compression and crude methanol purification are not modeled in detail. Instead, results from (Van-Dal and Bouallou, 2013) are used to estimate the electrical and thermal energy consumptions of both unmodeled sections. Additionally, the compositions of exit streams from the crude methanol section are also estimated based on (Van-Dal and Bouallou, 2013). Most notably, this includes the product stream purity, which is assumed constant 99.9931 wt-% MeOH (Van-Dal and Bouallou, 2013). Further assumptions in the MeOH synthesis model are:

- Mixer 1 (MX1) is assumed to achieve ideal and perfect mixing of streams.
- Heat exchangers (HE) consist of bulk models, where the synthesis stream perfectly reaches the desired temperature setpoint and the model only considers the required amount of heat flow (kW). Thus, heating/cooling medium flows and temperatures are dismissed. Since sizing of heat exchangers is not considered, there are no pressure losses modeled in heat exchangers. (Parvasi et al., 2008)

The methanol synthesis reactor was sized so that the number of reactor tubes would result the weight hourly space velocity (*WHSV*) to be equal to 4 h<sup>-1</sup> at the nominal operation point, which is typical in the industry (Arab et al., 2014). In general, the model parameters were selected based on the previous studies described in the literature (Hartig and Keil, 1993; Lommerts et al., 2000; Manenti et al., 2011, 2014; Parvasi et al., 2008). The methanol synthesis model was programmed in MATLAB<sup>®</sup> and the model is publicly available at (Tiiri, 2024).

### 2.3 DMC synthesis

Among the DMC synthesis routes, the direct synthesis of DMC from CO<sub>2</sub> and methanol offers a compelling eco-friendly alternative to traditional methods. However, this pathway encounters thermodynamic hurdles stemming from the equilibrium constraints of the reaction (4):

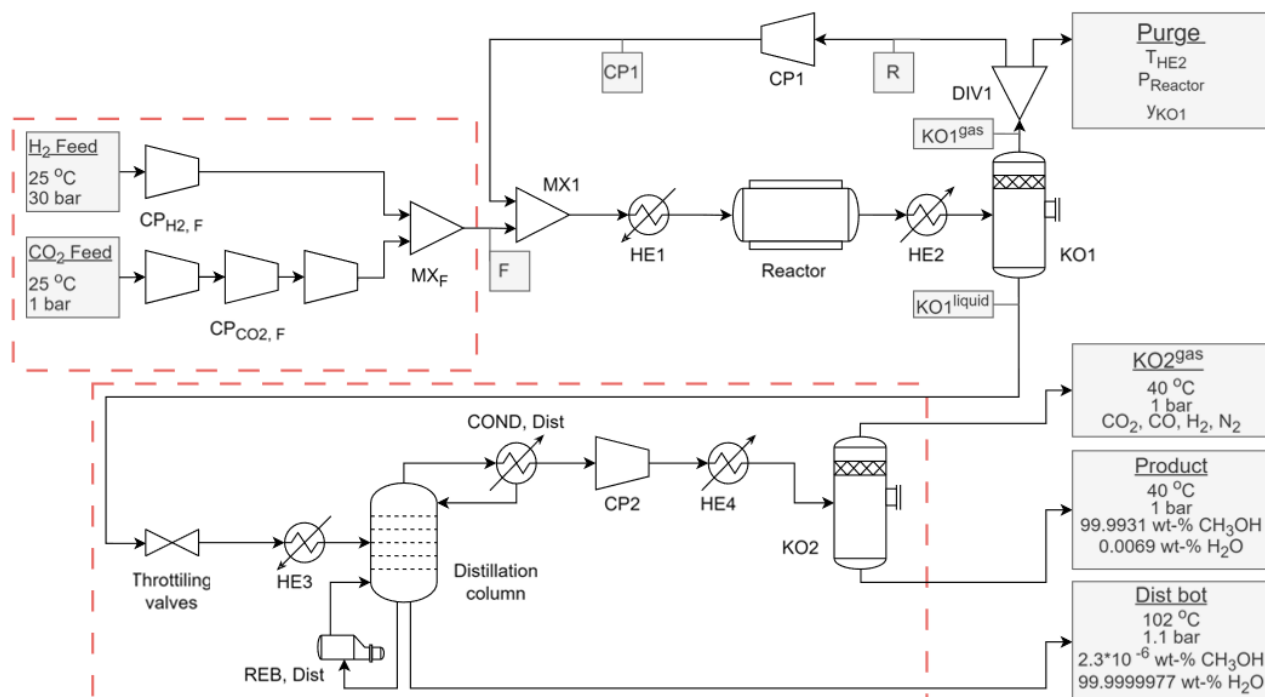
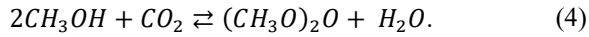


Fig. 3. Methanol synthesis process flowchart.





In addition to thermodynamic limitations, the direct DMC synthesis possesses significant challenges due to the low reactivities of both MeOH and CO<sub>2</sub>. Consequently, catalysts play a crucial role in enhancing the efficiency of the process. In (Zheng et al., 2022), recent catalyst research is summarized, emphasizing the pivotal role of heterogeneous zirconia-based catalysts, including ZrO<sub>2</sub> and solid solutions of ZrO<sub>2</sub> with other metal oxides (M/ZrO<sub>2</sub>).

Equations (5)–(8) detail the mass balances for each component involved in the DMC simulations. In this research, the kinetic parameters,  $k$ , were accurately fitted with the experimental data reported by (Camy et al., 2003). Differential Evolution optimization algorithm was used for the fitting.

$$\frac{dC_{CO_2}}{dt} = -k_1 C_{CO_2} \cdot C_{MeOH}^2 + k_2 C_{DMC} \cdot C_{H_2O} \quad (5)$$

$$\frac{dC_{MeOH}}{dt} = -k_3 C_{CO_2} \cdot C_{MeOH}^2 + k_4 C_{DMC} \cdot C_{H_2O} \quad (6)$$

$$\frac{dC_{H_2O}}{dt} = k_5 C_{CO_2} \cdot C_{MeOH}^2 - k_6 C_{DMC} \cdot C_{H_2O} \quad (7)$$

$$\frac{dC_{DMC}}{dt} = k_7 C_{CO_2} \cdot C_{MeOH}^2 - k_8 C_{DMC} \cdot C_{H_2O} \quad (8)$$

The reaction time for the assumed sizing of the reactor in nominal operation conditions was set to 2 h. The pressure was assumed being constant, and the temperature dependence of the kinetic parameters in (5)–(8) were assumed to follow the Arrhenius equation (9):

$$k = k_0 e^{-\frac{E_A}{RT}}. \quad (9)$$

where  $k_0$  [1/s] is the pre-exponential factor,  $E_A$  [J/mol] is the activation energy,  $R$  is the universal gas constant, and  $T$  is process temperature. The parameters are given in Table 1. The energy consumption of the direct DMC synthesis was not modeled.

Table 1. Estimated kinetic parameters for DMC synthesis.

Forward reaction		Reverse reaction	
$k_{0\_1}=11.31$	$E_{A1}=2.18 \cdot 10^4$	$k_{0\_5}=8.23$	$E_{A5}=1.93 \cdot 10^4$
$k_{0\_2}=3.50$	$E_{A2}=1.11 \cdot 10^6$	$k_{0\_6}=12.66$	$E_{A6}=5.58 \cdot 10^4$
$k_{0\_3}=4.05$	$E_{A3}=1.57 \cdot 10^4$	$k_{0\_7}=1.70$	$E_{A7}=1.37 \cdot 10^4$
$k_{0\_4}=3.06$	$E_{A4}=7.09 \cdot 10^5$	$k_{0\_8}=1.44$	$E_{A8}=2.41 \cdot 10^5$

### 3. SIMULATIONS, RESULTS, AND DISCUSSION

#### 3.1 Simulation scenarios

For demonstrating the developed process model chain, the simulation scenarios feature CCU of pulp mill recovery boiler effluent based on (Gardarsdottir et al., 2014). In all simulation scenarios the flue gas is assumed to be first dehydrated (e.g. vapor-liquid separation, adsorption) from all water and cooled to 50 °C. The nominal simulation scenario (Case I), and two other scenarios to simulate the effect of feed variability to the process chain, are considered. For Case II, the flue gas has +10% molar flow and +5% CO<sub>2</sub> fraction. For Case III the flue gas has -10% molar flow and -5% CO<sub>2</sub> fraction. The flue gas

flows of dried effluent, and CO<sub>2</sub> contents for the three scenarios are presented in Table 2.

The process designs and choices for dividing streams between units are determined and fixed based on the nominal simulation case (Case I). Fixing some of the flows in the process chain streams represents constraints that might be present in a plant due to fixed orders from customers, or due to instrument sizing. The higher flue gas flow and CO<sub>2</sub> content in Case II represent an undersized process design for the process chain, and the opposite, Case III represents partial load circumstances. The KPIs of especial interest in the process chain are yields and specific energy consumptions (*SEC*) of the process units.

Table 2. Flue gas properties for the different simulation scenarios.

	Case I	Case II	Case III
Flue gas flow (mol/s)	3734	4107	3361
CO <sub>2</sub> content (mol-%)	16.4	17.2	15.6
N <sub>2</sub> content (mol-%)	78.2	77.6	79.0
O <sub>2</sub> content (mol-%)	5.4	5.2	5.4

The process chain simulations assume the following:

- CO<sub>2</sub> capture is performed with a membrane-based process as described in Section 2.1. The total membrane area is  $2.12 \cdot 10^5$  m<sup>2</sup>, the membrane fiber outer radius is 600 μm and the inner radius is 400 μm. The operation pressure is 5 bar, temperature 50 °C and the recycle ratio is 1.
- The CO<sub>2</sub> rich permeate stream is divided into methanol and DMC syntheses so that 125 mol/s of permeate is always directed to the DMC synthesis and rest to the MeOH synthesis.
- Hydrogen production is not simulated, and the hydrogen feed is assumed to be perfectly pure. If the permeate stream contains oxygen, excess hydrogen is included in the MeOH synthesis feed to convert all oxygen into water. Hydrogen flow is controlled so that the MeOH synthesis feed contains 3:1 molar ratio of H<sub>2</sub> to CO<sub>2</sub> and no free oxygen.
- The MeOH synthesis is simulated as described in Section 2.2. For design parameters, the reactor tube length is 7 m, the tube inner radius is 4.2 cm, and the number of tubes is 1900. The catalyst is assumed to be fresh, and thus at 100% activity. For operational variables, the reactor inlet temperature ( $T_{inlet}$ ) is 510 K, the reactor inlet pressure is 78 bar, the reactor shell side temperature ( $T_{shell}$ ) is 533 K, temperature for the crude methanol separation at KO1 is 308 K, and the mass based recycle ratio (*RR*) is 5.
- Produced MeOH is divided into two streams. The first stream is a fixed product stream with a flow of 45 mol/s. The rest of MeOH flows to the DMC synthesis to combine with the dedicated portion of CO<sub>2</sub>-rich permeate stream.
- The DMC synthesis is simulated according to the model in Section 2.3. The pressure is assumed to be 126 bar and the temperature 453.15 K.

### 3.2 Sensitivity to feed fluctuations

The resulting molar flows of products and KPIs from the three different cases are reported in Table 3. In the table, the total CO<sub>2</sub> efficiency is given (10):

$$\frac{\dot{n}_{MeOH,sold} + 3 \dot{n}_{DMC,sold}}{\dot{n}_{CO_2,flue\ gas}}, \quad (10)$$

where  $\dot{n}_{MeOH,sold}$  combines the constant stream of 45 mol/s, and the flow of leftover MeOH after the DMC synthesis.  $\dot{n}_{DMC,sold}$  represents the molar flow of produced DMC and is multiplied by 3 to account for reaction stoichiometry from CO<sub>2</sub>.  $\dot{n}_{CO_2,flue\ gas}$  is the molar flow of CO<sub>2</sub> in flue gas stream.

In the nominal case, membrane-based CO<sub>2</sub> capture shows a capture efficiency of 70.1% while the CO<sub>2</sub> content is increased from 16.4 mol-% to 97.8 mol-%. According to Table 3, the relative change in the quality of captured CO<sub>2</sub> is small, less than 0.5 mol-%. The change in CO<sub>2</sub> capture rate in Table 3 corresponds approximately  $\pm 5\%$  change with respect to the nominal case.

From the perspective of the MeOH synthesis, the different simulation cases have similar captured CO<sub>2</sub> qualities, while the permeate flowrates vary more significantly. Compared to Case I, the resulting changes to MeOH production rate in Case II and Case III are in-line with the changes to permeate flow rates ( $> \pm 10\%$ ), while changes to yield and *SEC* are much smaller ( $< \pm 1\%$ , and  $< \pm 2\%$ , respectively). The differences in MeOH yield and *SEC* between the cases can be explained by changes in *WHSV*. With less feed flow in Case III compared to Case I, the reactants have longer residence time in the reactor, and the pressure losses are smaller, thus allowing improved yield and *SEC*. The effects to yield and *SEC* are reversed for Case II, as *WHSV* increases compared to Case I. However, in general, it can be concluded that the MeOH production KPIs are not drastically affected by fluctuations in the feed conditions.

Nominally, the DMC production feed ratio for MeOH and CO<sub>2</sub> is equal to 2. With this feed composition, the DMC process shows a yield of 53.4% (on MeOH basis). For DMC, the upstream changes cause the MeOH-to-CO<sub>2</sub> ratio to change between 1.7 and 2.3. In addition, the fixed reactor size also means that the residence time changes due to fluctuations in the total feed flow. In Case II, the residence time is 1.83 h and in Case III 2.23 h instead of nominal value of 2 h (Case I). As presented in Table 3, longer residence time increases the DMC yield.

With respect to the total CO<sub>2</sub> efficiency, the partial load case represents 4% better relative efficiency than in the nominal case, as molar yields and CO<sub>2</sub> capture rate are all improved. On the contrary, in Case II, the overall efficiency is decreased by 5.2%.

### 3.3 Process chain flexibility

To study the flexibility of the process chain, it is examined if modifying the operational variables of CO<sub>2</sub> capture and MeOH synthesis units can compensate the lower flue gas flow and quality in Case III. The aim is to maintain the process chain throughput at the same level as in Case I, while the process design parameters remain fixed. For CO<sub>2</sub> capture, the goal is to provide an equal amount of CO<sub>2</sub> flow to downstream processes as in the nominal case, while achieving as high purity as possible. For MeOH synthesis unit, the aim is to produce an equal flow of methanol as in Case I.

The membrane-based CO<sub>2</sub> capture can be operated at different pressures and 2<sup>nd</sup> stage retentate recycle ratios. The increment of 1<sup>st</sup> stage operation pressure would have a significant negative effect on the *SEC* due to the very large amount of gas (flue gas and recycle) needed to be elevated into higher pressure. Thus, adjusting only the 2<sup>nd</sup> stage operation pressure, and recycle ratio are considered in this work.

Table 3. Resulting molar flows and key performance indicators in cases I to III.

	Case I	Case II	Case III
Molar flow of captured CO <sub>2</sub> (mol <sub>CO<sub>2</sub></sub> / s)	430	470	388
Captured CO <sub>2</sub> quality (mol-%)	97.8	98.2	97.4
CO <sub>2</sub> capture rate (mol-%)	70.1	66.4	74.0
<i>SEC</i> , CO <sub>2</sub> capture (kJ <sub>el</sub> / mol <sub>CO<sub>2</sub></sub> )	78.4	79.0	78.4
Ratio of captured CO <sub>2</sub> flowing to MeOH feed and to DMC feed (-)	2.5	2.8	2.2
Molar flow of captured CO <sub>2</sub> to MeOH synthesis (mol <sub>CO<sub>2</sub></sub> / s)	308	347	266
Molar flow of produced MeOH (mol <sub>MeOH</sub> / s)	289	324	251
Molar yield of MeOH from CO <sub>2</sub> (mol-%)	94.1	93.3	94.3
<i>SEC</i> , MeOH synthesis (kJ <sub>el</sub> / mol <sub>MeOH</sub> )	33.9	34.4	33.6
Molar flow of MeOH to DMC synthesis (mol <sub>MeOH</sub> / s)	245	279	206
Molar ratio of MeOH and CO <sub>2</sub> in DMC feed (-)	2	2.3	1.7
Molar flow of produced DMC (mol <sub>DMC</sub> / s)	65	69	60
Molar yield of DMC from MeOH (mol-%)	53.4	49.2	58.6
Molar flow of unreacted MeOH from DMC synthesis (mol <sub>MeOH</sub> / s)	81	99	61
Total CO <sub>2</sub> efficiency (mol-%)	52.4	49.7	54.5

A set of simulations were performed to assess the flexibility of CO<sub>2</sub> capture in Case III by altering the recycle ratio and 2<sup>nd</sup> stage operation pressure. High recycle ratio is preferred to increase both the CO<sub>2</sub> purity and recovery. Recycle increases the feed CO<sub>2</sub> content and thus generates a higher driving force for the membrane separation. However, the operation pressure has opposite effects to the two KPIs, as can be seen in Fig. 4. Increasing the pressure leads to a higher recovery, but due to improved overall permeation, the quality of permeate decreases linearly. In terms of energy consumption, the higher 2<sup>nd</sup> stage pressure, and thus capture rate, results in lower *SEC* (in 9 bar, 73.6 kW/molCO<sub>2</sub>). Finally, in the partial load case, the target CO<sub>2</sub> flow (equal to Case I) could be achieved with the 2<sup>nd</sup> stage operation pressure of 11.5 bar (extending from the range seen in Fig. 4). In these conditions, the CO<sub>2</sub> purity was only 91.3 mol-%, the CO<sub>2</sub> recovery 82.0%, and the total molar flow rate 471 mol/s. The resulting total flow rate would be 7.2% higher than in Case I, meaning that the downstream processes will need to handle a larger amount of gas than expected by the nominal design, requiring more flexibility from them.

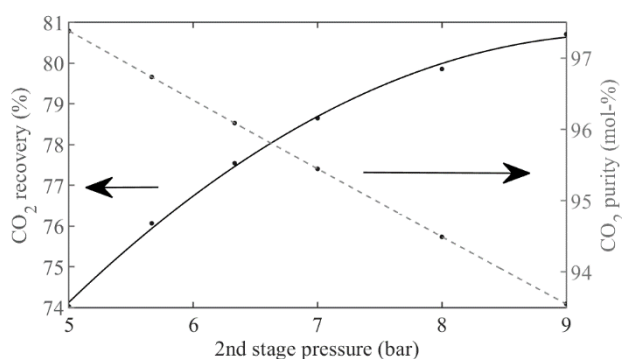


Fig. 4. The effect of operation pressure on the CO<sub>2</sub> capture performance. The recycle ratio is 1.

Thus, it can be concluded that the selected operation variables in the CO<sub>2</sub> capture cannot solely compensate for the flue gas variations in the partial load case. The disturbances are partly propagated to the downstream processes resulting in 337 mol/s of permeate flowing to MeOH synthesis and 134 mol/s of permeate to the DMC synthesis with decreased purity as described earlier. It should be noted that the gas membrane systems can be designed as modular units. Thus, increased flexibility can be achieved by altering the effective membrane area (number of active modules) to maintain the performance characteristics.

For flexibility in MeOH synthesis unit, it was seen that there are three operational variables, that can be most easily manipulated to adjust the production rate: Reactor inlet temperature ( $T_{inlet}$ ), reactor shell side temperature ( $T_{shell}$ ), and recycle ratio ( $RR$ ). There were no constraints set to the operational variables, other than that the resulting temperature should not exceed the earlier mentioned temperature limit of 553 K at any point along the reactor tubes.

By performing global optimization using differential evolution algorithm from (Buehren, 2024), a set of operational variables was found that results in practically equal MeOH production

rate (289 mol/s) as in the nominal case:  $T_{inlet} = 505.2$  K,  $T_{shell} = 534.1$  K, and  $RR = 11.19$ . Expectedly, with such a high recycle stream, *SEC* is noticeably higher (40.0 kJ<sub>el</sub> / mol<sub>MeOH</sub>) compared to the nominal case (33.9 kJ<sub>el</sub> / mol<sub>MeOH</sub>). The found operational point does demonstrate that the MeOH synthesis plant can be operated flexibly to compensate even for major changes in the captured CO<sub>2</sub> purity. Thus, the membrane-based CO<sub>2</sub> capture unit and the MeOH synthesis unit combined can fully compensate for the feed variability of the studied scenarios, so that the DMC production and the overall productivity of the process chain is unaffected.

However, it is worth mentioning that the operational variable values for the MeOH synthesis in the flexibility case are considerably above their normal ranges that can be found in the literature. The reason why using so large temperatures and a recycle ratio does not cause operational issues within the simulations, can likely be attributed to the simplifications present in the MeOH synthesis model. By including currently unmodeled phenomena, such as limitations to radial heat transfer inside reactor tubes (Hartig and Keil, 1993), or pressure losses taking place in heat exchangers, such an extreme operation point might become unviable.

### 3.4 Implications to techno-economic analyses

The ongoing debate in the scientific literature for the kinetics of methanol synthesis reactions is one major source of uncertainty for modeling the MeOH production. To examine sensitivity of the MeOH synthesis towards chosen reaction kinetics, Case I was simulated with an updated kinetic parameter set reported in (Mignard and Pritchard, 2008). The simulation resulted in the flow of produced methanol rising from 289 to 298 mol/s. Methanol yield from CO<sub>2</sub> rose from 94.1 to 96.7 mol-%. *SEC* decreased from 33.9 to 33.0 kJ<sub>el</sub> / mol<sub>MeOH</sub>. The changes in these KPIs are significant even though the update from (Mignard and Pritchard, 2008) is only altering two of the parameters in the kinetic model of (Bussche and Froment, 1996). It could be speculated that modeling the reaction kinetics after a completely different model structure, such as in (Graaf et al., 1988), could cause even more significant changes to the KPIs. In general, the same principle applies for the kinetics in the DMC production, and the permeabilities and selectivities of different membrane materials.

As demonstrated by the flexibility simulations, the choice of operational variables carries a significant effect on the product purity, yield and *SEC* in the different unit processes, meaning that choosing an optimal operation point is also important at system level analyses. For example, methanol synthesis in Case I was simulated with the highest reactor inlet and shell side temperatures found from the used literature. If the same case is instead simulated with significantly lower temperatures ( $T_{inlet} = 484$  K,  $T_{shell} = 520$  K), from (Manenti et al., 2014), then the flow of produced methanol lowers by 5.5% to 273 mol/s, the methanol yield from CO<sub>2</sub> drops to 88.7 mol-%, whilst *SEC* rises to 35.7 kJ<sub>el</sub> / mol<sub>MeOH</sub>.

For the DMC production, the estimates are uncertain due to the low technology readiness level and lab-scale data. In general, the direct DMC synthesis is very energy intensive (*e.g.* 19.2

MJ/mol under process conditions  $T = 323.15$  K and  $P = 150$  bar (Saavalainen et al., 2015)). To enhance the yield of DMC, Zheng et al. (2022) implemented a natural convection circulation system specifically to adsorb and remove water.

Models can also give overly optimistic estimates due to their inability to account for uncertainties in geometry. For example, the hollow-fiber membrane model assumed a fixed inner and outer radius, although real systems with thousands or millions of fibers might have variability in their properties. It has been shown that the standard deviation over 10% in fiber geometry can have a significant impact on the CO<sub>2</sub> recovery (Bocciardo, 2015). The sensitivity of the geometry parameters in CO<sub>2</sub> capture was also observed in this study. For instance, simulating Case I with a 2.5% decrement in both membrane fiber radius values leads to CO<sub>2</sub> purity lowering from 97.8 to 97.6%, capture rate lowering from 70.1 to 69.2%, and *SEC* rising from 78.4 to 82.0 kJ<sub>el</sub> / mol<sub>CO<sub>2</sub></sub>. Moreover, MeOH and DMC production efficiencies are also greatly dependent on the reactor dimensions.

Yet another factor to be accounted for is the long-term stability of the processes. The membrane-based CO<sub>2</sub> capture has shown good stability in extended periods of operation for coal-fired combustion flue gases (Cui et al., 2021). On the other hand, an industrial data-based study suggests that the activity of Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst in methanol synthesis can drop to 65% when the catalyst has operated for 100 days, and further reduce to 50% after a year of operation (Parvasi et al., 2008). Thus, focusing solely on the performance of fresh catalyst, as has been done in this study, can result in overly optimistic KPIs. If simulating Case I otherwise unchanged, but the catalyst activity is at 50%, the flow of produced methanol drops to 228 mol/s. The methanol yield from CO<sub>2</sub> lowers to 74.2 mol-%, and *SEC* rises to 42.6 kJ<sub>el</sub> / mol<sub>MeOH</sub>. The matter is relevant considering that the average lifetime of Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst is 3 to 4 years (Bozzano and Manenti, 2016).

#### 4. CONCLUSIONS

In this work a CCU process chain involving membrane-based CO<sub>2</sub> capture, MeOH synthesis and DMC synthesis was modeled. In simulations it was found that the studied CCU process chain is quite robust against feed fluctuations, and by operating the membrane-based CO<sub>2</sub> capture unit and the MeOH synthesis flexibly together, it was possible to fully compensate the decrease in the production rate in the studied worst case feed conditions. More importantly for techno-economic analysis, it was discovered that the CCU process chain is significantly more sensitive to model parameters than to feed variations.

Furthermore, the significant changes to yields and specific electricity consumptions from varying operational variables and design parameters imply that considering their optimal choice should play a significant role also in system level studies. Thus, for techno-economic analyses of CCU processes it is crucial to find profit-wise optimum solutions that balance various factors, such as investment costs versus operating costs when sizing equipment, or yield versus catalyst life when choosing synthesis temperatures. Access to industrial data

from relevant operation conditions would be highly valuable for increasing the reliability of the process models.

#### ACKNOWLEDGEMENTS

This work was funded by Business Finland via project 'Bio-CCU - Creating sustainable value of the bio-based CO<sub>2</sub>' (2352/31/2022).

#### REFERENCES

- Akimoto, K., Sano, F., Oda, J., Kanaboshi, H., and Nakano, Y. (2021). Climate change mitigation measures for global net-zero emissions and the roles of CO<sub>2</sub> capture and utilization and direct air capture. *Energy and Climate Change*, 2, 100057.
- Arab, S., Commenge, J.-M., Portha, J.-F., and Falk, L. (2014). Methanol synthesis from CO<sub>2</sub> and H<sub>2</sub> in multi-tubular fixed-bed reactor and multi-tubular reactor filled with monoliths. *Chemical Engineering Research and Design*, 92(11), 2598–2608.
- Asadi, J., and Kazempoor, P. (2021). Techno-economic analysis of membrane-based processes for flexible CO<sub>2</sub> capturing from power plants. *Energy Conversion and Management*, 246, 114633.
- Azhari, N. J., Erika, D., Mardiana, S., Ilmi, T., Gunawan, M. L., Makertihartha, I. G. B. N., and Kadja, G. T. M. (2022). Methanol synthesis from CO<sub>2</sub>: A mechanistic overview. *Results in Engineering*, 16, 100711.
- Bocciardo, D. (2015). *Optimisation and integration of membrane processes in coal-fired power plants with carbon capture and storage*. University of Edinburgh.
- Bozzano, G., and Manenti, F. (2016). Efficient methanol synthesis: Perspectives, technologies and optimization strategies. *Progress in Energy and Combustion Science*, 56, 71–105.
- Buehren, M. (2024). *Differential Evolution* (<https://www.mathworks.com/matlabcentral/fileexchange/18593-differential-evolution; 1.16.0.1>) [MATLAB]. MATLAB Central File Exchange.
- Bussche, K. M. V., and Froment, G. F. (1996). A Steady-State Kinetic Model for Methanol Synthesis and the Water Gas Shift Reaction on a Commercial Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> Catalyst. *Journal of Catalysis*, 161(1), 1–10.
- Camy, S., Pic, J.-S., Badens, E., and Condoret, J.-S. (2003). Fluid phase equilibria of the reacting mixture in the dimethyl carbonate synthesis from supercritical CO<sub>2</sub>. *The Journal of Supercritical Fluids*, 25(1), 19–32.
- Cui, Q., Wang, B., Zhao, X., Zhang, G., He, Z., Long, Y., Sun, Y., and Ku, A. Y. (2021). Post-combustion slipstream CO<sub>2</sub>-capture test facility at Jiangyou Power Plant, Sichuan, China: Performance of a membrane separation module under dynamic power-plant operations. *Clean Energy*, 5(4), 742–755.

- Gardarsdottir, S. O., Normann, F., Andersson, K., and Johnsson, F. (2014). Process Evaluation of CO<sub>2</sub> Capture in three Industrial case Studies. *Energy Procedia*, 63, 6565–6575.
- Graaf, G. H., Stamhuis, E. J., and Beenackers, A. A. C. M. (1988). Kinetics of low-pressure methanol synthesis. *Chemical Engineering Science*, 43(12), 3185–3195.
- Green, D. W., and Southard, M. Z. (Eds.). (2019). *Perry's Chemical Engineers' Handbook 9th Edition* (Ninth edition). McGraw Hill Education.
- Gumber, S., and Gurumoorthy, A. V. P. (2018). Methanol Economy Versus Hydrogen Economy. In *Methanol* (pp. 661–674). Elsevier.
- Hartig, F., and Keil, F. J. (1993). Large-scale spherical fixed bed reactors: Modeling and optimization. *Industrial and Engineering Chemistry Research*, 32(3), 424–437.
- Hou, R., Fong, C., Freeman, B. D., Hill, M. R., and Xie, Z. (2022). Current status and advances in membrane technology for carbon capture. *Separation and Purification Technology*, 300, 121863.
- Karjunen, H., Tynjälä, T., Kuparinen, K., Vakkilainen, E., and Joronen, T. (2023). Value creation by converting pulp mill flue gas streams to green fuels. *TAPPI Journal*, 22(3), 193–205.
- Khalilpour, R., Abbas, A., Lai, Z., and Pinnau, I. (2012). Modeling and parametric analysis of hollow fiber membrane system for carbon capture from multicomponent flue gas. *AIChE Journal*, 58(5), 1550–1561.
- Kohli, K., Sharma, B. K., and Panchal, C. B. (2022). Dimethyl Carbonate: Review of Synthesis Routes and Catalysts Used. *Energies*, 15(14), 5133.
- Koytsoumpa, E. I., Bergins, C., and Kakaras, E. (2018). The CO<sub>2</sub> economy: Review of CO<sub>2</sub> capture and reuse technologies. *The Journal of Supercritical Fluids*, 132, 3–16.
- Lommerts, B. J., Graaf, G. H., and Beenackers, A. A. C. M. (2000). Mathematical modeling of internal mass transport limitations in methanol synthesis. *Chemical Engineering Science*, 55(23), 5589–5598.
- Manenti, F., Cieri, S., and Restelli, M. (2011). Considerations on the steady-state modeling of methanol synthesis fixed-bed reactor. *Chemical Engineering Science*, 66(2), 152–162.
- Manenti, F., Leon-Garzon, A. R., Ravaghi-Ardebili, Z., and Pirola, C. (2014). Systematic staging design applied to the fixed-bed reactor series for methanol and one-step methanol/dimethyl ether synthesis. *Applied Thermal Engineering*, 70(2), 1228–1237.
- Mignard, D., and Pritchard, C. (2008). On the use of electrolytic hydrogen from variable renewable energies for the enhanced conversion of biomass to fuels. *Chemical Engineering Research and Design*, 86(5), 473–487.
- Parvasi, P., Rahimpour, M. R., and Jahanmiri, A. (2008). Incorporation of Dynamic Flexibility in the Design of a Methanol Synthesis Loop in the Presence of Catalyst Deactivation. *Chemical Engineering and Technology*, 31(1), 116–132.
- Peng, D.-Y., and Robinson, D. B. (1976). A New Two-Constant Equation of State. *Industrial and Engineering Chemistry Fundamentals*, 15(1), 59–64.
- Poto, S., Vico Van Berkel, D., Gallucci, F., and Fernanda Neira d'Angelo, M. (2022). Kinetic modelling of the methanol synthesis from CO<sub>2</sub> and H<sub>2</sub> over a CuO/CeO<sub>2</sub>/ZrO<sub>2</sub> catalyst: The role of CO<sub>2</sub> and CO hydrogenation. *Chemical Engineering Journal*, 435, 134946.
- Reichle, D., Houghton, J., Kane, B., Ekmann, J., and Others, A. (1999). *Carbon sequestration research and development* (DOE/SC/FE-1). Oak Ridge National Lab. (ORNL), Oak Ridge, TN (United States); National Energy Technology Lab., Pittsburgh, PA (US); National Energy Technology Lab., Morgantown, WV (US).
- Rodin, V., Lindorfer, J., Böhm, H., and Vieira, L. (2020). Assessing the potential of carbon dioxide valorisation in Europe with focus on biogenic CO<sub>2</sub>. *Journal of CO<sub>2</sub> Utilization*, 41, 101219.
- Saavalainen, P., Kabra, S., Turpeinen, E., Oravisjärvi, K., Yadav, G. D., Keiski, R. L., and Pongrácz, E. (2015). Sustainability Assessment of Chemical Processes: Evaluation of Three Synthesis Routes of DMC. *Journal of Chemistry*, 2015, 1–12.
- Tiiri, K. (2024). *krti01github/Modeling-methanol-synthesis-plant-in-Matlab: V1.0.0 (v1.0.0) [MATLAB]*. doi:10.5281/ZENODO.11121416
- Van-Dal, É. S., and Bouallou, C. (2013). Design and simulation of a methanol production plant from CO<sub>2</sub> hydrogenation. *Journal of Cleaner Production*, 57, 38–45.
- Wei, W., Wang, Y., Yan, Z., Hou, J., Xu, G., and Shi, L. (2023). One-step DMC synthesis from CO<sub>2</sub> under catalysis of ionic liquids prepared with 1,2-propylene glycol. *Catalysis Today*, 418, 114052.
- Wilke, C. R. (1950). A Viscosity Equation for Gas Mixtures. *The Journal of Chemical Physics*, 18(4), 517–519.
- Zheng, Q., Nishimura, R., Sato, Y., Inomata, H., Ota, M., Watanabe, M., and Camy, S. (2022). Dimethyl carbonate (DMC) synthesis from methanol and carbon dioxide in the presence of ZrO<sub>2</sub> solid solutions and yield improvement by applying a natural convection circulation system. *Chemical Engineering Journal*, 429, 132378.



## Simulation Model for an Amine-based CO<sub>2</sub> Capture Rig

Soudeh Shamshiri, Neda Razi and Lars Erik Øi

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, 3901 Porsgrunn, Norway; e-mail: ([lars.oi@usn.no](mailto:lars.oi@usn.no)).*

---

**Abstract:** The amine-based CO<sub>2</sub> capture rig at USN in Porsgrunn has been operating since 2007. In this study, the main aim was to develop predictive models in Aspen HYSYS and Aspen Plus for the CO<sub>2</sub> test rig. The models accuracy were verified by comparing different test scenarios with results from the models. Aspen HYSYS and Aspen Plus have simulated eleven scenarios (test series) with varying process parameters. In Aspen HYSYS, Murphree efficiencies (stage efficiencies) were fitted, and in Aspen Plus two approaches were used, fitting the interfacial area or the holdup factor to minimize the deviation between the model and experimental data. The Aspen HYSYS model with the fitted Murphree efficiencies (from top to bottom 0.11, 0.1, 0.09 and 0.07) predicted seven scenarios with an average deviation of 12-24 % from experimental data. In the Aspen Plus rate-based model with interfacial area fitted, most of the scenarios were predicted by a model with correlation Brf-85 (mass transfer), Brf-85 (heat transfer) and an interfacial area factor of 0.5. Minimum and maximum deviations for different scenarios were 2.1 and 9 %. In the approach with fitting of the holdup factor, the Brf-92 holdup method with a holdup factor of 0.5 gave the best fit, resulting in an average deviation of 1.4-9 % from the test results across all scenarios.

*Keywords:* CO<sub>2</sub>, amine, rate-based, absorption, simulation

---

### 1. INTRODUCTION

The amine-based CO<sub>2</sub> capture rig at the University of South-Eastern Norway (USN) in Porsgrunn has been operating since 2007. It is a small-scale CO<sub>2</sub> capture plant that includes an absorber, a desorber with reboiler and condenser, heat exchangers and pumps. Performance data measuring especially CO<sub>2</sub> capture rate at different conditions using 30 weight-% MEA have been presented in several papers (Øi et al., 2013; Øi et al., 2015; Øi et al., 2017; Øi et al., 2021).

There have been attempts to fit the experimental data from the CO<sub>2</sub> capture rig to simulation models with limited success by fitting data series to a constant Murphree efficiency and using rate-based models. However, the models fitted to experimental data have shown poor predictive properties for conditions outside the experimental range (Karunarathne and Øi, 2019).

Early work on comparing test data with different simulation tools was by Luo et al. (2009) who tested Aspen RadFrac, ProTreat, ProMax, Aspen RateSep, CHEMASIM from BASF and CO<sub>2</sub>SIM from SINTEF/NTNU using different pilot plant data. The conclusion was that all the models could fit the CO<sub>2</sub> capture rate, but the temperature and concentration profiles were not well predicted.

Much work with fitting CO<sub>2</sub> absorption performance data to Aspen HYSYS or Aspen Plus models have been performed with data from the Test Center Mongstad (TCM). In previous research at USN, data from TCM was used to fit equilibrium models (in Aspen Plus and Aspen HYSYS) to various scenarios by adjusting each stage's Murphree efficiency ( $E_M$ ), and rate-based models (in Aspen Plus) were fitted by varying

the interfacial area factor or the holdup factor (Øi et al., 2018; Øi and Fagerheim, 2020; Øi et al., 2022). Except for this work at TCM, there is very little published information about comparing different models in simulation tools to fit experimental data (Kvamsdal et al., 2011; Razi et al., 2012; Razi et al. 2013a).

In rate-based models, it is traditional to adjust the interfacial area factor to fit the model to performance data, but to use the holdup factor (multiplied with the holdup estimated by in-built correlations) has been used for the same purpose (Øi et al., 2022). Two of the estimation methods in Aspen Plus for several of the correlations, especially the liquid holdup, is Brf-85 and Brf-92 which is short for the Bravo/Rocha/Fair model (Rocha et al., 1992).

This work is based on the Master Thesis of Soudeh Shamshiri (2023) at USN which aims to fit different models in Aspen HYSYS and Aspen Plus to experimental data in the USN test rig. The ultimate goal is to assess which models provide the most accurate predictions for different conditions, and giving the most accurate dependence on the varying process parameters.

This paper starts with an introduction with background literature in section 1. In section 2 different correlations for fitting experimental CO<sub>2</sub> absorption data to models and correlations are presented. In section 3 specifications for the simulations and the data fitting to experimental data from the CO<sub>2</sub> absorption rig are presented. In section 4 the deviations between the simulations and performance data are presented, and the different approaches for modelling and fitting are discussed.



## 2. METHODS FOR FITTING EXPERIMENTAL CO<sub>2</sub> ABSORPTION DATA TO MODELS

### 2.1 Equilibrium-based models

Simulation tools include column models for distillation and absorption, which assume equilibrium at each stage. Changing the number of simulation stages until the desired CO<sub>2</sub> capture efficiency is achieved is an easy method of adjusting an absorption equilibrium model to match experimental CO<sub>2</sub> capture efficiency.

Equilibrium based models can be refined by defining a Murphree efficiency at each column stage. This efficiency is calculated as the ratio of the change in mole fraction from a stage to the next divided by the change assuming equilibrium. An absorption column can be defined with a number of stages, e.g. equal to the packed column height, and the Murphree stage efficiency can be specified to be equal for each stage. The constant Murphree efficiency can be fitted to experimental CO<sub>2</sub> efficiency in performance data.

At USN equilibrium models in Aspen Plus and Aspen HYSYS have been fitted to different scenarios (experimental data) at Test Center Mongstad (TCM) by adjusting the Murphree efficiency ( $E_M$ ) for each stage (Øi et al., 2018; Øi and Fagerheim, 2020; Øi et al., 2022).

### 2.2 Rate-based models

The process simulation tool Aspen Plus has a rate-based model to describe the mechanisms in the absorption process. Rate-based models calculate rate of mass transfer, rate of heat transfer, pressure drop and equilibrium. Unlike the equilibrium model, the rate-based approach assumes that separation is achieved through mass and energy transfer between gas and liquid which is in equilibrium at the interface. The rate-based modelling is increasingly accepted over traditional equilibrium-stage modelling for CO<sub>2</sub> capture processes.

A rate-based model for CO<sub>2</sub> absorption in MEA is available in the Aspen Plus program. This model was developed and fitted based on data from a CO<sub>2</sub> capture pilot plant at the University of Texas (Zhang et al., 2009). The parameter values in this Aspen Plus model can be changed by the user.

### 2.3 Correlations in rate-based models

A rate-based model includes correlations (submodels) for mass transfer rates through the gas film and the liquid film, rate of heat transfer, pressure drop, interfacial area and liquid holdup. Aspen Plus has several available correlations for each of these quantities. A review of mass transfer models for both random and packed columns is given by Wang et al. (2005). Reviews of rate-based models for CO<sub>2</sub> capture are presented by Kvamsdal et al. (2011), Razi et al. (2012) and Amirkhosrow et al. (2021). Razi et al. (2013) have evaluated how different correlations in rate-based simulations work in Aspen Plus to predict performance from the Esbjerg pilot plant. The prediction of performance data is very dependent on the

correlations used. The models differ in accuracy and correct description of sensitivity of operating parameters like temperature, gas flow and liquid flow.

In the Aspen Plus rate-based model, examples of mass transfer correlations are named Brf-85 (Bravo et al, 1985), Hanley (2012) and Brf-92 (Bravo et al, 1993) and interfacial area methods are named Brf-85, Hanley, Brf-92 and Mod-Tsai. Inaccuracy in one of the correlations can be compensated by adjusting that correlation or other correlations by adjusting model parameters. In Aspen Plus, it is possible to compensate in some of the correlations by adding an adjustment parameter. For the interfacial area, the interfacial area factor is used, and for the liquid holdup, the liquid holdup factor is used.

These correlations normally require physical data (like densities and viscosities) as input. The effect of this is discussed in Nookuea et al. (2015) and Karunaratne and Øi (2019). In this work, the default property methods from the sample file from Aspen Plus are used.

### 2.4 Equilibrium models

Both an equilibrium-based model and a rate-based model is dependent on the vapour/liquid equilibrium model. In earlier work, different equilibrium models have been used (Øi et al., 2018). Both Aspen Plus and the new Aspen HYSYS acid gas model now use an Electrolyte-NRTL equilibrium model (Austgen et al., 1989). This is used in this work.

## 3. DATA AND SPECIFICATIONS

### 3.1. Performance data from test rig at USN

Data were taken from experimental work that has been done for several years in the test rig at USN, a picture is shown in Fig. 1 and a simplified P&ID is shown in Fig. 2. Eleven scenarios have been defined in this work in which different process parameters like gas and liquid flow rate, solvent inlet temperature and CO<sub>2</sub> inlet concentration were varying. Figures 3, 4 and 6 are from Øi et al. (2017), Figure 5 is from Øi et al. (2013) and Figure 7 is from Øi et al. (2021).



Fig. 1. Picture of CO<sub>2</sub> test rig at USN.

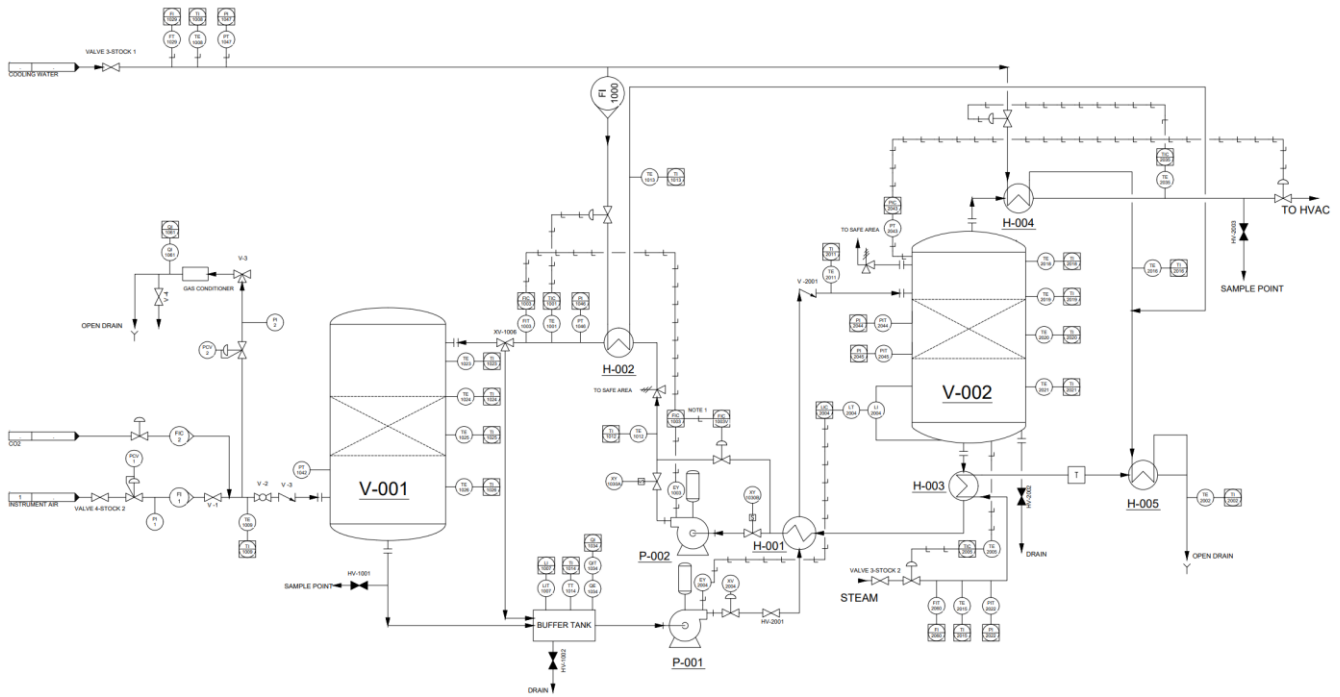


Fig. 2. Simplified P&ID for USN test rig.

Figures 3 to 7 show data as a function of only one parameter. The CO<sub>2</sub> capture rate is dependent on many parameters. Detailed information can be found and the exact numerical values for the data points can be found in the original references or in Shamshiri (2023).

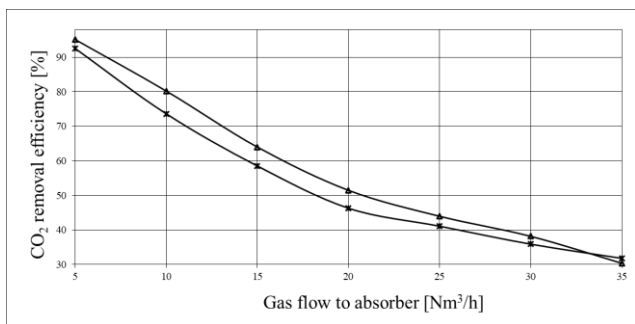


Fig. 3. Data for varying the gas flow at a high liquid flow in scenario 1-1 and 1-2. The upper curve is for liquid flow 150 l/h and the lower curve is for 140 l/h. There is 10 % CO<sub>2</sub> in the inlet gas.

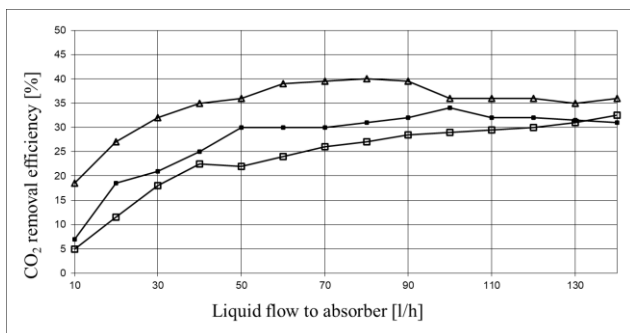


Fig. 4. Data for varying liquid flow in scenario 2-1, 2-2 and 2-3. The upper curve is for 5 % CO<sub>2</sub> in inlet gas, the middle curve for 10 % CO<sub>2</sub> and the lower curve is for 15 % CO<sub>2</sub>.

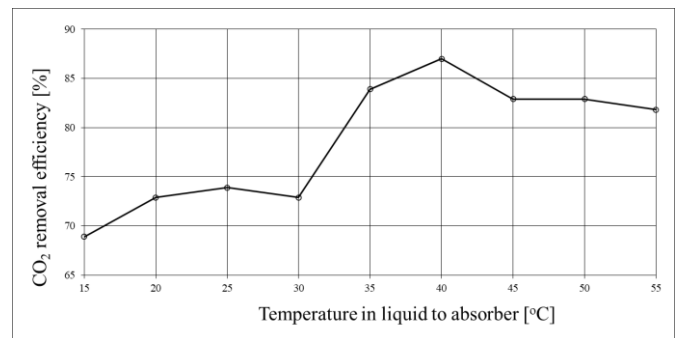


Fig. 5. Data for varying the temperature in scenario 3-1. The data are for constant gas flow 14 Nm<sup>3</sup>/h and liquid flow 50 kg/h. There is 10 % CO<sub>2</sub> in the inlet gas.

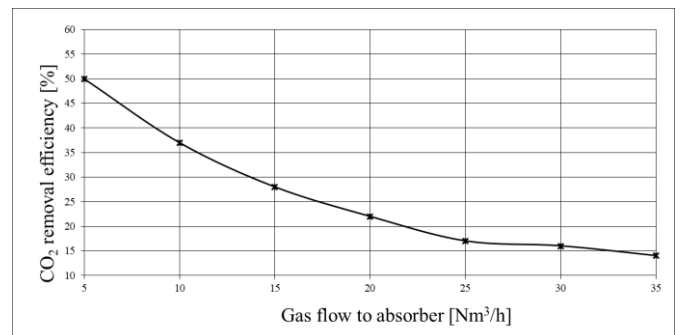


Fig. 6. Data for varying the gas flow at a low liquid flow for scenario 4-1. Liquid flow is 20 l/h and there is 10 % CO<sub>2</sub> in inlet gas.

The selection of scenarios is tried to be representative for the available data for the different variables and for the variable ranges. The selection have not been systematical, and in this work there has not been any evaluation about which experimental data for the test rig which are of highest quality.

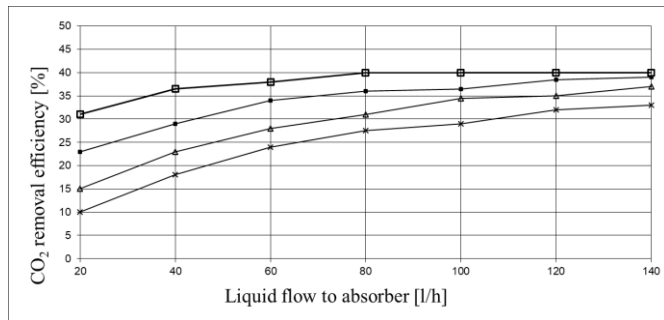


Fig. 7. Data for varying the liquid flow for scenario 5-1, 5-2, 5-3 and 5-4. The curves are from top to bottom for 2.5 % CO<sub>2</sub> in inlet gas, 5 %, 10 % and 15 %. Gas Flow is 30 Nm<sup>3</sup>/h.

### 3.2. Specifications for the equilibrium-based simulation tools

The input specifications in Table 1 for the Aspen HYSYS absorber model are similar to earlier simulations (Øi et al., 2018; Øi and Fagerheim, 2020). A flowsheet of the process is shown in Fig. 8.

Table 1. Specifications for equilibrium-based calculations.

Specifications – Aspen HYSYS equilibrium	
Calculation Type	equilibrium
Property method	Acid gas- Chemical solvents
Valid phases	Vapor-Liquid
Number of stages	4
Nominal pressure	110 kPa
Uniform section	yes
Internal type	Packed
Solving method	Modified HYSIM Inside-out

### 3.3. Specifications for the rate-based tool

The input specifications for the rate-based absorber calculations are given in Table 2. A flowsheet is shown in Fig. 9. In the rate-based simulation in Aspen Plus, the local model example "ENRTL-RK\_Rate\_Based\_MEA\_Model" from the

Aspen library was used. The Elec-NRTL thermodynamic package with Redlich-Kwong equation of state (RK) was chosen.

Table 2. Specifications of the model for rate-based calculations

Specifications – Aspen Plus rate-based	
Calculation Type	Rate-based
Property method	ENRTL-RK
Henry comp ID	Global
Chemistry ID	MEA-CHEM
Valid phases	Vapor-Liquid
Number of stages	4
Hold up factor	0.5-1
Reaction condition factor	0.9
Packing type	Mellapak 250Y
Section diameter (m)	0.1
Section packed height (m)	1.5
Flow model	VPlug*
Interfacial area factor	0.5-1-1.3-1.6
Film liquid phase	Discretize film
Film vapor phase	Consider film
Mass transfer coefficient method	Brf-85/Hanley/Brf-92
Heat transfer coefficient method	Chilton and Colburn
Interfacial area method	Brf-85/Hanley/Brf-92/Mod-Tsai
Holdup method	Brf-92
Add. Discretize points liquid	5

The specifications for the Aspen Plus model are mostly as in earlier work (Øi et al., 2021), but the chosen correlations, and especially the combination of correlations, are mostly a result of trial and error. A combination of correlations obtaining good prediction over a large range of the varied parameter is regarded as a useful method. The choice of interfacial area is traditional (Zhang et al., 2009; Øi et al., 2018; Øi and Fagerheim, 2020), but the use of the holdup factor as fitting variable is only found in earlier work by Øi et al. (2022).

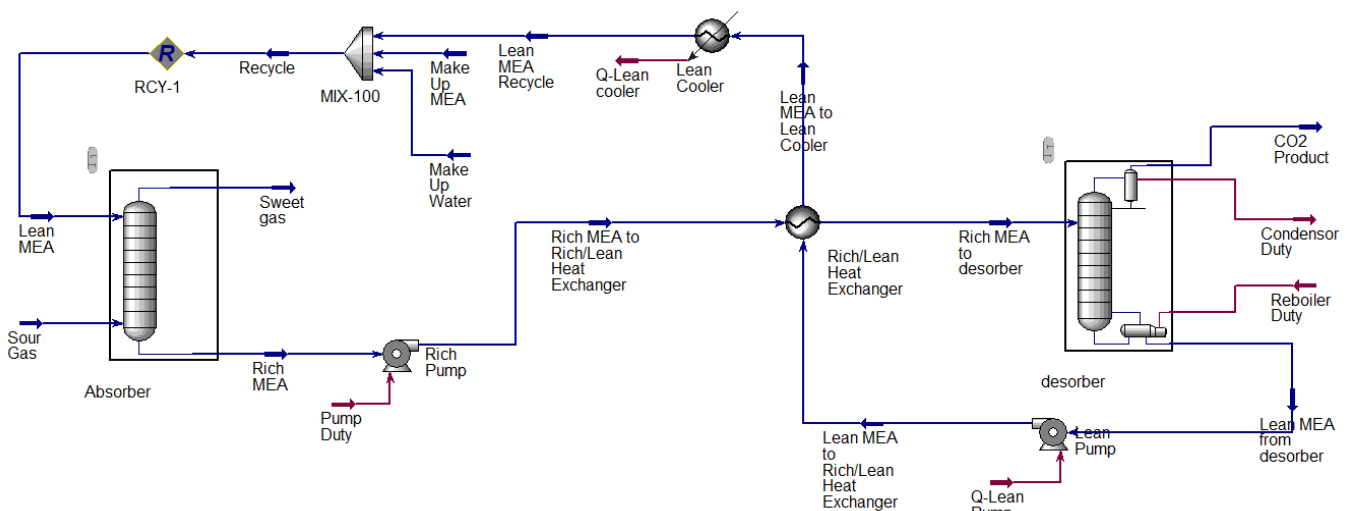


Fig. 8. Process simulation model of the CO<sub>2</sub> capture rig in Aspen HYSYS.

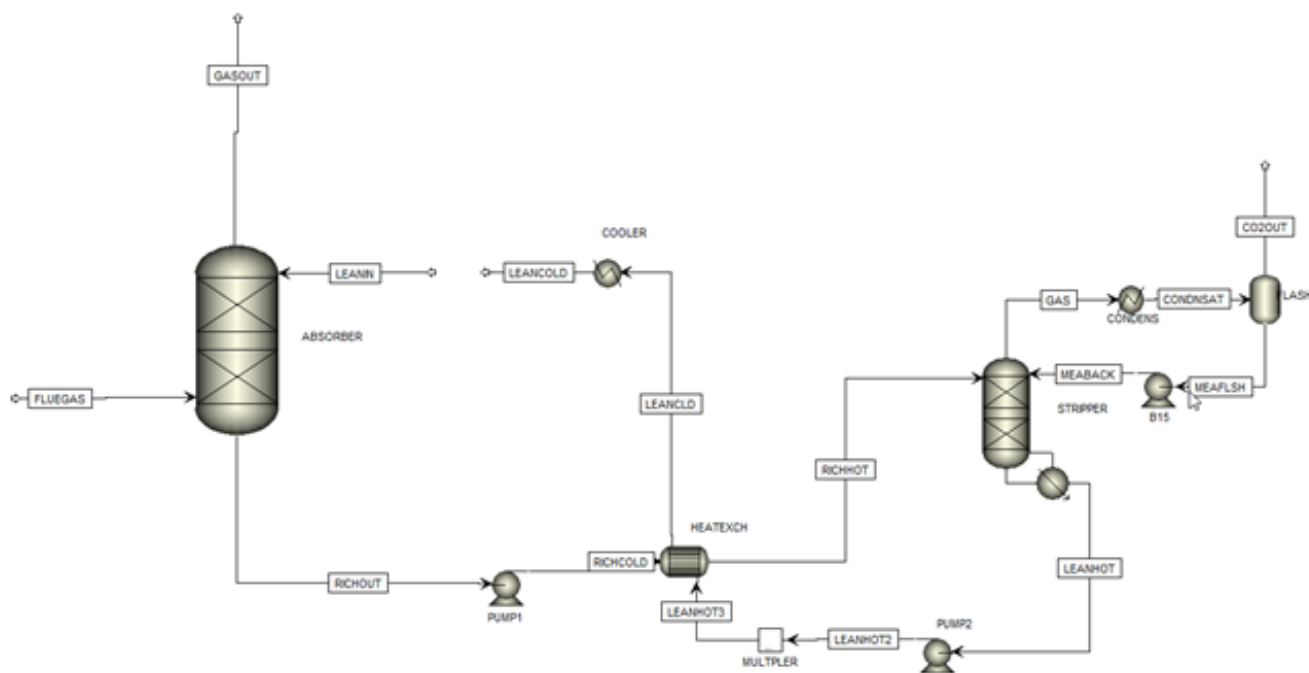


Fig. 9. Process simulation model of the CO<sub>2</sub> capture rig in Aspen Plus.

#### 3.4. Simulations to fit experimental values

The rate-based model developed in the simulation focused on the absorber column. In Aspen HYSYS, the Murphree efficiency was adjusted to determine the experimental CO<sub>2</sub> removal efficiency for a given scenario. Using Aspen Plus, the interfacial area factor or the holdup factor were fitted to obtain the experimental CO<sub>2</sub> removal efficiency.

#### 3.5. Simulations to predict performance

Based on experimental data, the fitted models in Aspen HYSYS or Aspen Plus were used to predict conditions at different operating conditions. The deviation between predicted and measured CO<sub>2</sub> capture rate was calculated in %.

### 4. SIMULATION RESULTS AND DISCUSSION

#### 4.1. General Results

One fitted simulation case is the result for each data point in each of the eleven scenarios. The main result from each simulation is the CO<sub>2</sub> removal rate which is found as the CO<sub>2</sub> amount in the gas feed to the absorber minus the CO<sub>2</sub> amount in the gas from the absorber. The deviation in this work is the simulated CO<sub>2</sub> removal minus the experimental CO<sub>2</sub> removal divided by the simulated value.

#### 4.2. Simulation model for equilibrium-based model

The equilibrium-based model with defined parameters and properties has been simulated for the eleven scenarios in the test rig USN. The results in Fig. 10 show that the Aspen

HYSYS can predict the CO<sub>2</sub> removal rate for the sets of scenarios 2-1, 2-2, 2-3, 5-1, 5-2, 5-3 and 5-4 where the deviation of predicted model from experimental data is between 12-24%. In these simulations the Murphree efficiency for all stages were defined to be 0.11, 0.1, 0.09 and 0.07 from top to bottom.

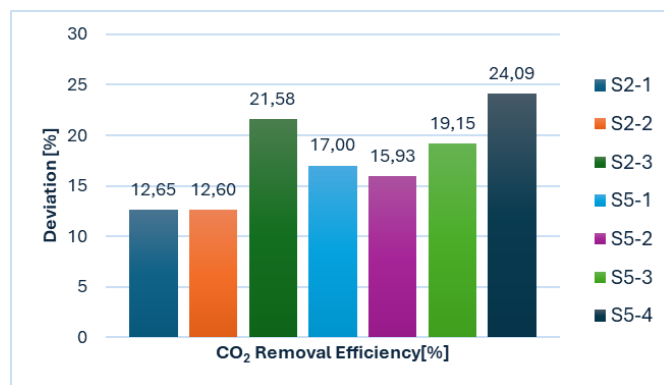


Fig. 10. Deviation between the experimental data and the simulation results from Aspen HYSYS.

#### 4.3. Simulation model for rate-based model

The model was used to simulate the performance data using various mass transfer and interfacial area correlations. Figure 11 shows that the Brf-85 method for both mass and heat transfer and an interfacial area factor of 0.5 or 1.0 gives smaller deviation than when the Hanley method is used. When the interfacial area factor is changed from 1 to 0.5, most of the experimental data will be reasonably predicted when varying the process parameters.



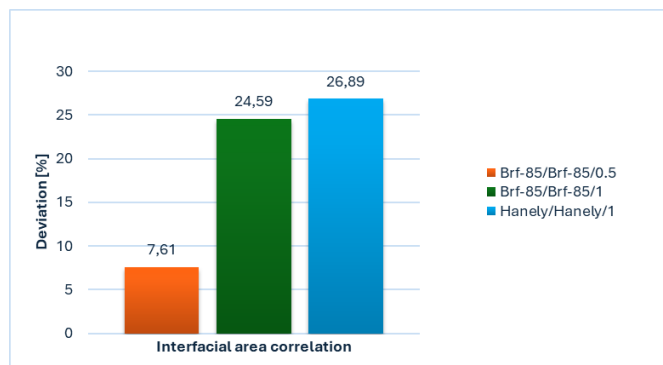


Fig. 11. Deviation of simulated data in Aspen Plus/Interfacial area from experimental data with different correlations for scenario 2-1.

#### 4.4 Deviation results for rate-based model

If the Brf-85 correlation could not predict all experimental data, another method would have simulated all scenarios, such as changing the liquid holdup factor. According to Fig. 12, Brf-85 as a mass transfer and interfacial area method with factor 1 beside Brf-92 as a holdup correlation with holdup factor 0.5 provides results with less error than the Hanley method and the same holdup correlation and factor.

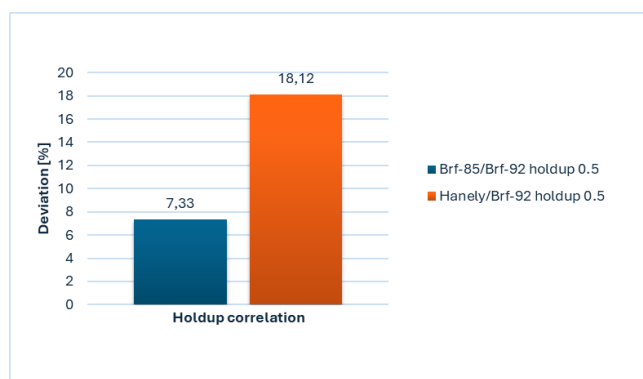


Fig. 12. Deviation of simulated data in Aspen Plus/Holdup from experimental data with different correlations for scenario 2-1.

Figures 11 and 12 indicate that using the interfacial area or holdup factor as the variable parameter gives similar sum of deviations at least for this case.

Only a limited number of different correlations and especially combination of correlations were tested. Evaluation of other correlations is discussed in the general discussion.

#### 4.5 Deviation results for rate-based model with varying interfacial area

In Fig. 13, the deviations are shown for all scenarios when using a model with Brf-85 for both the heat transfer, mass transfer and the interfacial area and an interfacial area factor 0.5, except for scenarios S3-1 and S4-1. In S3-1 the Hanley model for interfacial area and an interfacial area factor of 1.3 was necessary to fit the data and in S4-1 the Brf-92 model for heat and mass transfer, the ModTsai model for interfacial area and an interfacial area factor of 1.0 were necessary to fit the data satisfactory.

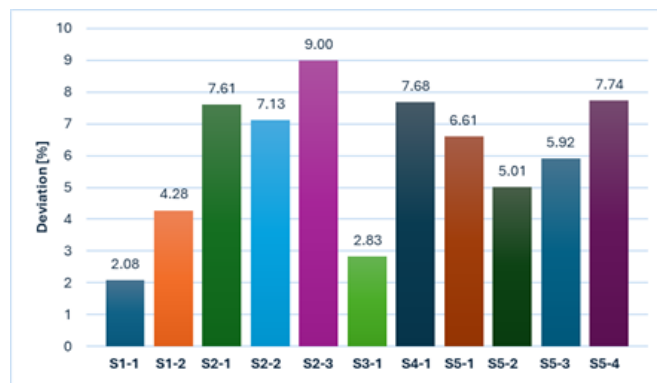


Fig. 13. Deviation between the experimental data and the simulation results from Aspen Plus/Interfacial area method with different correlations for S3-1 and S4-1.

For the scenarios S3-1 and S4-1, using the model with Brf-85 and interfacial area = 0.5 the fit was not satisfactory.  $\text{Øi}$  and Fagerheim (2020) and  $\text{Øi}$  et al. (2022) experienced that when fitting TCM data to a similar model, some experimental data was difficult to fit using the Brf-85 model for interfacial area, so the Brf-92 was used by  $\text{Øi}$  et al (2022).

#### 4.6 Deviation results for rate-based model with varying holdup factor

Figure 14 shows the deviation results for all the eleven data sets or scenarios. The same combination of correlations models were used in these simulations, with Brf-85 for heat and mass transfer, Brf-92 for holdup and  $h = 0.5$ .

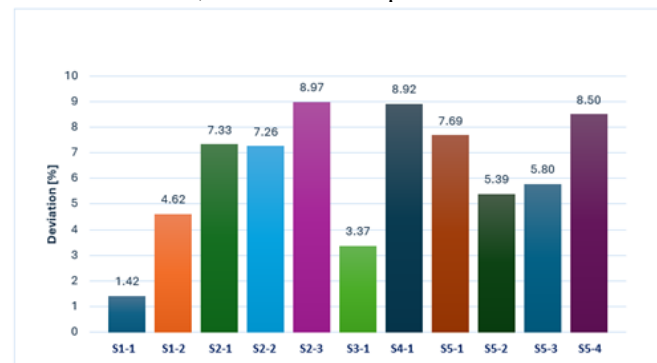


Fig. 14. Deviation between the experimental data and the simulation results from Aspen Plus/Holdup method with the same correlations also for S3-1 and S4-1.

By running the simulation for all the experiments, the average deviation from all sets of scenarios is below 9 %.

#### 4.7 General discussion

The results from this and earlier work show that it is possible to fit the  $\text{CO}_2$  removal rate for a performance data set using either a rate-based or an equilibrium-based model. For most cases, it is enough to adjust only one parameter. This can be the interfacial area factor or the liquid holdup factor in a rate-based model or an  $E_M$ -factor (adjusting all the stage efficiencies in an  $E_M$  profile) for an equilibrium-based model.

In Luo et al. (2009), Øi et al. (2012) and in Øi and Fagerheim (2020) it was claimed that at their conditions (low inlet CO<sub>2</sub> concentration) the equilibrium-based models and rate-based models performed about equally well in fitting the available performance data. This work indicates that for predicting performance at very different conditions, a rate-based model performs better than an equilibrium-based model.

This work and the work by Øi et al. (2022) show that a promising approach to obtain a predictive model for a CO<sub>2</sub> capture plant based on performance data, is to vary the holdup factor to fit performance data. It is recommended for future work to compare the approach of fitting the interfacial area factor and the holdup factor in more detail.

There is little in the literature about which combination of correlations for mass transfer, heat transfer, interfacial area and holdup that gives the most accurate prediction as a function of varied temperature, gas flow and liquid flow. This topic is discussed by Kvamsdal et al. (2011), Kvamsdal and Hillestad (2012), Razi et al. (2012) and Reza et al. (2013b). To find the most accurate or most convenient combination of correlations for rate-based simulation of amine-based CO<sub>2</sub> capture is a challenge for future work. One strategy is to test all (or most of) the possible combinations of models. Then it is a challenge to treat large amounts of experience data to obtain an optimum choice of models. Another strategy is to study the correlations more in detail to select the most accurate correlations. Then it is a challenge whether to choose the most accurate correlations, or the correlations which give the most accurate dependence on the varying process parameters.

## 5. CONCLUSIONS

In this study, the main aim was to develop predictive models in Aspen HYSYS and Aspen Plus for the CO<sub>2</sub> test rig. Aspen HYSYS and Aspen Plus have been used to simulate eleven scenarios (test series) with varying process parameters. In Aspen Plus two approaches were used, fitting either the interfacial area factor or the holdup factor to minimize the deviation between model and experimental values. An Aspen HYSYS model could not predict all experimental data when varying liquid temperature, inlet gas temperature and flow rates. A fitted rate-based model in Aspen Plus gave less error than the equilibrium-based Aspen HYSYS model. In the rate-based model with interfacial area fitted, most scenarios can be predicted by a model with correlation Brf-85 (mass transfer), Brf-85 (heat transfer) and an interfacial factor of 0.5. Minimum and maximum deviations for different scenarios were 2.1 and 9 %. In the approach with fitting the holdup factor, the Brf-92 holdup method with a holdup factor of 0.5 gave the best fit, resulting in an average deviation of 1.4-9 % from the test results for all scenarios. Testing and evaluation of combinations of correlations for rate-based models is a future challenge.

## REFERENCES

- Amirkhosrow, M., Perez Calvo, J.F., Gazzani, M., Mazotti, M., and Lay, E.N. (2021). Rigorous rate-based model for CO<sub>2</sub> capture via monoethanolamine-based solutions: effect of kinetic models, mass transfer, and holdup correlations on prediction accuracy. *Separation science and technology*, 56, pp. 1491–1509.
- Austgen, D.M., Rochelle, G.T., Peng, X. and Chen, C. (1989). Model of Vapor-Liquid Equilibria for Aqueous Acid Gas-Alkanolamine Systems Using the Electrolyte-NRTL Equation. *Industrial & Engineering Chemistry Research*, 28, pp. 1060-73.
- Bravo, J.L., and Fair, J. (1985). Mass-Transfer in Gauze Packings. *Hydrocarbon Processing*, 64, pp. 91-95.
- Garcia, M., Knuutila, H.K., and Gu, S. (2017). ASPEN Plus simulation model for CO<sub>2</sub> removal with MEA: Validation of desorption model with experimental data. *Journal of Environmental Chemical Engineering*, 5(5), pp. 4693-4701.
- Hanley, B. and Chen, C. (2012). New Mass-Transfer Correlations for Packed Towers. *AIChE Journal*, 58(1), pp. 132-152.
- Karunarathne, S.S. and Øi, L.E. (2019). Aspen HYSYS and Aspen Plus simulations for amine based absorption process compared to results from experiments in CO<sub>2</sub> rig. *10th Trondheim Conference on CO<sub>2</sub> Capture, Transport and Storage* (NTNU).
- Kvamsdal, H.M., Chikukwa, A., Hillestad, M., Zakeri, A. and Einbu, A. (2011). A comparison of different parameter correlation models and the validation of an MEA-based absorber model. *Energy Procedia*, 4, 1526-33.
- Kvamsdal, H.M., and Hillestad, M. (2012) Selection of model parameter correlations in a rate-based CO<sub>2</sub> absorber model aimed for process simulation. *International Journal of Greenhouse Gas Control*, 11, pp. 11-20.
- Luo, X., Knudsen, J.N., de Montigny, D., Sanpasertparnich, T., Idem, R., Gelowitz, D., Notz, R., Hoch, S., Hasse, H., Lemaire, E., Alix, P., Tobiesen, F.A., Juliussen, O., Köpcke, M. and Svendsen, H.F. (2009). Comparison and validation of simulation codes against sixteen sets of data from four different pilot plants. *Energy Procedia*, 1, 1249-56.
- Nookuea, W., Tan, Y., Li, H., Thorin, E. and Yan, J. (2015). Sensitivity study of thermo-physical properties of gas phase on absorber design for CO<sub>2</sub> capture using monoethanolamine. *Energy Procedia*, 75, 2305-10.



- Razi, N., Bolland, O., and Svendsen, H.S. (2012). Review of design correlations for CO<sub>2</sub> absorption into MEA using structured packings, *International Journal of Greenhouse Gas Control*, 9, 193-219.
- Razi, N., Svendsen, H.F. and Bolland, O. (2013a). Validation of mass transfer correlations for CO<sub>2</sub> absorption with MEA using pilot data, *International Journal of Greenhouse Gas Control*, 19, 478-491.
- Razi, N., Svendsen, H.F., and Bolland, O. (2013b). The impact of design correlations on rate-based modeling of large scale CO<sub>2</sub> capture with MEA. *Energy Procedia*, 37, 1977-86.
- Rocha, J.A., Bravo, J.L., and Fair, J.L. (1992). Distillation Columns Containing Structured Packings: A comprehensive Model for their Performance. 1. Hydraulic Models. *Industrial & Engineering Chemistry Research*, 32, 641-651.
- Shamshiri, S. (2023). Simulation model for the amine-based CO<sub>2</sub> capture rig in Porsgrunn. Master Thesis, University of South-Eastern Norway
- Wang, G.Q., Yuan, X.G. and Yu, K.T. (2005). Review of Mass-Transfer Correlations for Packed Columns. *Industrial & Engineering Chemistry Research*, 44, 8715-99.
- Zhang, Y., Chen, H., Chen, C., Plaza, J., Dugas, R., and Rochelle, G.T. (2009). Rate-based Process Modelling Study of CO<sub>2</sub> Capture with Aqueous Monoethanolamine Solution. *Industrial & Engineering Chemistry Research*, 48, pp. 9233-46.
- Zhang, Y., and Chen, C. (2013). Modeling CO<sub>2</sub> absorption and desorption by aqueous monoethanolamine solution with Aspen rate-based model. *Energy Procedia*, 37, 1584-96.
- Øi, L. E. (2012). Comparison of Aspen HYSYS and Aspen Plus simulation of CO<sub>2</sub> absorption into MEA from atmospheric gas. *Energy Procedia*, 23, 360-369.
- Øi, L.E., Lundberg, J., Pedersen, M., Hansen, P.M., and Melaaen, M.C. (2013). Laboratory rig for atmospheric CO<sub>2</sub> absorption and desorption under pressure. GHGT-11 conference. *Energy Procedia*, 37, 1933-40.
- Øi, L.E., Hansen, P.M., and Henriksen, M. (2015). Measurements of CO<sub>2</sub> absorption and heat consumption in laboratory rig. GHGT-12 conference, *Energy Procedia*, 63, 1273-81.
- Øi, L.E., Hansen, P.M., and Henriksen, M. (2017). CO<sub>2</sub> absorption efficiency and heat consumption measured at high gas to liquid ratio. GHGT-13 conference, *Energy Procedia*, 1273-81.
- Øi, L.E., Sætre, K.A., and Hamborg, E.S. (2018). Comparison of Simulation Tools to Fit and Predict Performance data of CO<sub>2</sub> Absorption into Monoethanol Amine. *Linköping Electronic Conference Proceedings SIMS 59*, pp.230-235. doi: [10.3384/ecp18153230](https://doi.org/10.3384/ecp18153230)
- Øi, L.E., and Fagerheim, S. (2020). Simulation of CO<sub>2</sub> Absorption at TCM Mongstad for Performance Data Fitting and Prediction. *Linköping Electronic Conference Proceedings SIMS 61*, pp. 332-337. doi: [10.3384/ecp20176332](https://doi.org/10.3384/ecp20176332)
- Øi, L.E., Hansen, P.M., and Henriksen, M. (2021). Laboratory rig for CO<sub>2</sub>-absorption and desorption at USN. *15th International conference on greenhouse gas control technologies*, GHGT-15, Abu Dhabi, UAE.
- Øi, L.E., Sæter, N., and Razi, N. (2022). Process simulation of CO<sub>2</sub> absorption data fitted to performance efficiency at TCM Mongstad. *Linköping Electronic Conference SIMS 63, Vol 192*, pp. 109-116.

## Design of electrified fluidized bed calciner for direct capture of CO<sub>2</sub> from cement raw meal

Ladan Samaei\*, Lars-Andre Tokheim\*\*, Christoffer Moen\*\*\*

\* Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
(e-mail: [258911@usn.no](mailto:258911@usn.no))

\*\* Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
(e-mail: [Lars.A.Tokheim@usn.no](mailto:Lars.A.Tokheim@usn.no))

\*\*\* Heidelberg Materials Norway (e-mail: [christoffer.moen@heidelbergmaterials.com](mailto:christoffer.moen@heidelbergmaterials.com))

**Abstract:** In this study, a fluidized bed electrified calciner was designed, and various operating conditions were investigated by CPFD simulations. It was found that maintaining a constant fluidization velocity while increasing the temperature of hot cylinders or preheating the raw meal significantly enhances the calcination degree. Altering the fluidization velocity while keeping temperatures constant also greatly affects the calcination degree and particle entrainment. A fluidization velocity of 0.3 m/s appears to be optimal for the reactor, whereas 0.8 m/s resulted in complete entrainment of the bed. The maximum calcination degree achieved was 90% with preheated meal. The average meal residence time was found to be 24–26 s.

**Keywords:** CO<sub>2</sub> capture, Electrification, Fluidized bed, Calciner, CPFD

### 1. INTRODUCTION

With an estimated 2600 million tons of CO<sub>2</sub> emissions in 2020, the cement industry holds the second-largest share of direct industrial CO<sub>2</sub> emissions worldwide (IEA, 2020). Furthermore, because of urbanization and the growing global population, cement production is predicted to increase by 12–23% by 2050 (IEA., 2018). By 2050, the European Union (EU) wants to cut its greenhouse gas emissions by 80–95% from 1990 levels (Dupont and Oberthür, 2015). Achieving this objective will largely depend on decarbonizing the cement sector. About two thirds of the CO<sub>2</sub> emissions from the cement industry are from calcination of limestone, whereas one third comes from the burning of fuels (Nikolakopoulos et al., 2024). However, using green electricity to calcine the raw materials and combining this with storage of the pure CO<sub>2</sub> generated in the calcination process can vary significantly reduce CO<sub>2</sub> emissions in the cement industry (Tokheim et al., 2019).

By switching from fuel combustion to electrical energy for the cement process both in calciner and in the rotary kiln, two things happen at once. Firstly, the CO<sub>2</sub> produced from burning fuel in the calciner is eliminated, reducing the overall CO<sub>2</sub> emissions from making clinker. Secondly, the exhaust gas from the calciner is almost pure CO<sub>2</sub>, so it can go straight to a CO<sub>2</sub> processing unit without needing a separate CO<sub>2</sub> separation plant. Having said that, if only the calciner uses the clean electricity instead of burning fuels while the kiln still uses the fuel combustion for providing heat, a significant reduction (around 70%) in emissions can be achieved. This happens while only one of the main equipment units in the kiln system (the calciner) needs to be altered. (Tokheim et al., 2019). An illustration of such a concept is shown in Fig. 1.

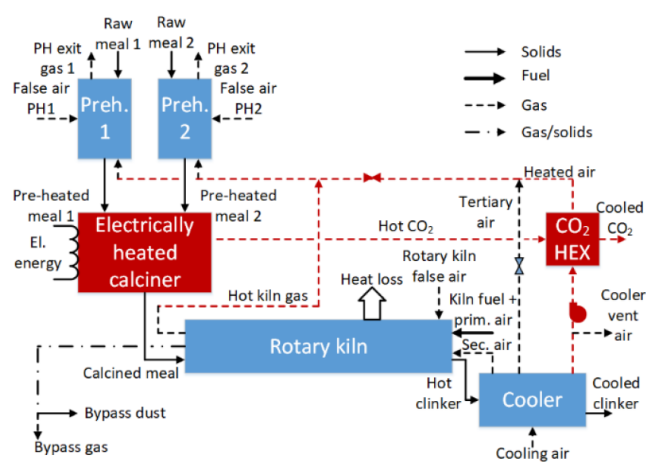


Fig. 1. Block flow diagram of a cement kiln system with electrified calciner (Tokheim et al., 2019).

Insisting on the benefits of using an electrified calciner, the question is which type of reactor is the best for designing an electrified calciner? This has been investigated in previous studies, and a fluidized bed (FB) calciner was found to be a good candidate as it provides good mixing and high heat and mass transfer (Samani et al., 2020). The raw meal feed is usually a fine powder (median size less than 30  $\mu\text{m}$ ) which can be classified as Geldart C particles. Geldart C particles are difficult to fluidize due to large inter-particle forces in comparison with the drag forces applied by the gas (Geldart, 1973). One way to tackle this problem is to improve the fluidization properties by mixing the small particles with coarse inert particles. After calcination of the meal, separation and entrainment of Geldart C particles is obtained by a change in the gas velocity. The reactor should be designed in such a way that the coarse particles remain in the bed, whereas the fine uncalcined particles are continuously fed close to the

bottom of the bed and fine calcined particles are continuously discharged at top of the bed. The CO<sub>2</sub> evolving from the calcination reaction will contribute in the entrainment of the calcined particles.

The FB operating with a binary particle mixture has previously been studied in several publications. Kato et al. (1991) investigated the residence time distribution of fine particles in a powder-particle FB. In another study, Tashimo et al. (1999) investigated the calcination of fine powder raw meal (Geldart C) in a powder-particle FB experimentally. Other studies and experiments have also been conducted in investigation of the fluidization conditions of binary particles FB (Kim et al., 2014). More recently a study was conducted by Jacob et al. (2022), who made an experimental and numerical investigation of using coarse particles for fluidization of fine limestone in the calciner. That study found that a fluidization velocity of 0.3 m/s and a weight fraction of 25 % fine particles gave a stable bed (Jacob et al., 2022).

A recent study by Jacob and Tokheim (2023) designed an electrified calciner with vertical heating channels and investigated this by CPFD simulations. While the designed calciner in that study showed promising results, in the industry it may be difficult to distribute the meal evenly between the heating channels, and therefore, there may be a risk that local overheating may occur.

The current study investigates an alternative heater arrangement, in the form of horizontally immersed cylindrical heating elements, as this will likely not prevent transverse mixing of the raw meal. A lab-scale electrified unit of about 100 kW is designed and simulated using the commercial CPFD software Barracuda. The main aims are to determine conditions for obtaining both fluidization, calcination, and separation in the designed reactor. Key process parameters, such as heater temperature, fluidization velocity and meal temperature are varied in different simulation cases.

## 2. METHODS

### 2.1 Design and sizing

A binary particle FB reactor with electrically heated immersed horizontal cylinders should be designed. A heat and mass balance is made, and FB calculations are performed to ensure good mixing and entrainment of fine calcined particles while keeping other particles in the bed. A trial-and-error procedure is required to find an appropriate size for the reactor and cylinders as well as an appropriate cylinder arrangement. The design calculations are done based on the assumption that the bottom cross section of the reactor is rectangular, and the length of the cylinders are all the same equal to one side of the bottom cross section. This assumption is needed because the cylinders should be electrically heated. Also, it is assumed that the reactor wall temperature is constant and that there is no heat loss.

The decomposition of calcium carbonate ( $CaCO_3 \rightarrow CaO + CO_2$ ) is an endothermic reaction requiring  $1.7 \frac{MJ}{kg_{CaCO_3}}$  (Jacob, 2023). Also, the raw meal should be heated up to reach the calcination temperature ( $T_{cal}$ ) before the reaction happens.

Equation (1) shows the energy balance of the reactor where the first part shows the heat needed for the reaction and the second part shows the sensible heat needed to heat up the meal. In the equation,  $\dot{m}_{rm}$ ,  $w_{CaCO_3}$  and  $C_{rm}$  are the mass flow rate of feeding raw meal to the reactor, the weight fraction of limestone in the raw meal and the specific heat capacity of the raw meal, respectively.  $\eta$  is the degree of calcination,  $H_{cal}$  is the enthalpy of calcination and  $\dot{E}$  is the electric energy supply.

$$\dot{m}_{rm} w_{CaCO_3} \eta H_{cal} + \dot{m}_{rm} C_{rm} (T_{cal} - T_{in}) = \dot{E} \quad (1)$$

Using the energy balance, the mass flow rate of the raw meal is calculated. So, the required heat transfer area for the calcination and for heating up the meal can be calculated using (2) and (3), respectively. In both equations,  $U$  is the overall heat transfer coefficient which is assumed to be  $0.3 \frac{kW}{m^2 K}$ .  $\Delta T$  in (2) shows the temperature difference between the hot cylinder wall and the calcination temperature, whereas  $\Delta T_{LMTD}$  in (3) is the logarithmic mean temperature between the cylinder wall and the raw meal being heated.

$$A_{cal} U \Delta T = \dot{m}_{rm} w_{CaCO_3} \eta H_{cal} \quad (2)$$

$$A_{sensible} U \Delta T_{LMTD} = \dot{m}_{rm} C_{rm} (T_{cal} - T_{in}) \quad (3)$$

The overall heat transfer area is the sum of the areas in the equations above and is used to calculate the number of cylinders. This will be calculated after some trial-and-error procedure for defining the size of the bottom cross section of the reactor.

The procedure begins with the assumption of two values, L1 and L2, for the rectangular bottom cross section of the reactor, aiming to keep the area as small as possible. The mass of gas required to achieve the fluidization velocity is then calculated using (4) where  $u_F$  is the fluidization velocity. The produced CO<sub>2</sub> is calculated based on a mass balance (5)-(8), where  $M$  shows the molecular weight, and  $\dot{m}$  and  $\dot{n}$  show the mass flow rate and molar flow rate, respectively.

$$\dot{m}_{CO_2, inj} = \rho_{CO_2} u_F (L_1 L_2) \quad (4)$$

$$\dot{m}_{CaCO_3} = \dot{m}_{rm} w_{CaCO_3} \quad (5)$$

$$\dot{n}_{CaCO_3} = \frac{\dot{m}_{CaCO_3}}{M_{CaCO_3}} \quad (6)$$

$$\dot{n}_{CaCO_3} = \dot{n}_{CO_2} \quad (7)$$

$$\dot{m}_{CO_2, prod} = \dot{n}_{CO_2, prod} M_{CO_2} \quad (8)$$

The total CO<sub>2</sub>, i.e., the sum of the injected and produced CO<sub>2</sub>, determines the outlet velocity. If this velocity is below the required entrainment velocity, the calculation is repeated. As a design constraint the exit is cylindrical with a fixed diameter.

$$\dot{m}_{CO_2, out} = \dot{m}_{CO_2, inj} + \dot{m}_{CO_2, prod} \quad (9)$$

$$u_{out, CO_2} = \frac{\dot{m}_{CO_2, out}}{A_{top} \rho_{CO_2}} \quad (10)$$

Next, the length of each hot cylinder is set to  $L_1$ , and the surface area for one cylinder is computed. Dividing the total heat transfer area ( $A_{cal} + A_{sensible}$ ) by the surface area of one cylinder ( $\pi D_{cyl} L_1$ ) gives the number of cylinders. The number of rows of cylinders needed are calculated. The height of the area including cylinders should be almost equal to the height of the bed. Adjustments are made to the cross-section dimensions until the desired conditions are met. Finally, the total reactor height is ensured to be adequate to retain coarse particles but short enough for proper calcination and entrainment of fine particles.

## 2.2 CPFD Simulation

Barracuda version 21.1.1 is used for the CPFD simulations. The simulation is conducted to investigate different operational conditions. The investigated operating conditions include variation in the temperature of the raw meal feed to the reactor, the wall temperature of the hot cylinders and the fluidization gas velocity. The reaction rate, amount of particle entrainment, entrainment velocity and pressure drop in the reactor are then compared for different simulation cases.

To do the simulations, after preparing the geometry and a proper grid the governing equations should be solved for the fluid and particle phases. Barracuda technology uses the 3D Multiphase Particle-in-Cell (3D-MP-PIC) method developed by CPFD Software. The MP-PIC method solves equations for both the gas and solids phases. It uses a Eulerian approach for the gas and a mix of the Eulerian and Lagrangian approach for the solids. Instead of tracking each physical particle, it groups similar ones into numerical parcels. This makes calculating properties like particle stresses faster in the gas phase. The method can handle various particle phases, sizes, and materials efficiently (Snider, 2001), (Jacob, 2023).

To define the fluid behavior in a CPFD simulation, equations for continuity, momentum, energy and species transport should be solved. In the continuity equation (11),  $\theta_f$ ,  $\rho_f$  and  $u_f$  are volume fraction, density and velocity of the fluid. Equation (12) shows the momentum equation where  $\nabla p$  denotes the pressure gradient in the gas flow,  $\tau_f$  symbolizes the stress exerted by the fluid,  $F$  is the inter-phase momentum transfer rate per unit, and  $g$  represents the gravity constant. Equation (13) is the species transport equation, where  $Y_{f,i}$  and  $\delta\dot{m}_{i,chem}$  are the mass fraction of species  $i$  in the fluid and the chemical source term, respectively. The term  $D_t$  is the turbulent mass diffusivity, which can be related to the flow viscosity by the Schmidt number correlation ( $Sc = \frac{\mu}{\rho D}$ ). For a non-isothermal simulation, the energy equation (14) should also be solved.  $H_f$  is the enthalpy of the fluid,  $\phi$  is the viscous dissipation rate,  $\dot{q}_f''$  is the fluid heat flux,  $\dot{Q}$  represents the energy source and  $S_h$  is the energy exchange between fluid and particles (Snider et al., 2011), (Jacob, 2023).

$$\frac{\partial(\theta_f \rho_f)}{\partial t} + \nabla \cdot (\theta_f u_f \rho_f) = \delta\dot{m}_p \quad (11)$$

$$\frac{\partial(\theta_f \rho_f u_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f u_f u_f) = -\nabla p + F + \theta_f \rho_f g + \nabla \cdot (\theta_f \tau_f) \quad (12)$$

$$\frac{\partial(\theta_f \rho_f Y_{f,i})}{\partial t} + \nabla \cdot (\theta_f u_f \rho_f Y_{f,i}) = \nabla \cdot (\rho_f D_t \theta_f \nabla Y_{f,i}) + \delta\dot{m}_{i,chem} \quad (13)$$

$$\frac{\partial(\theta_f \rho_f H_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f H_f u_f) = \theta_f \left( \frac{\partial p}{\partial t} + u_f \nabla p \right) + \phi - \nabla \cdot (\theta_f \dot{q}_f'') + \dot{Q} + S_h + \dot{q}_D + \dot{q}_w \quad (14)$$

Following the particle movement equation the term  $\delta\dot{m}_p$  in the fluid continuity equation can be calculated in (15), where  $\frac{dm_p}{dt}$  is the rate of change of the particle mass producing gas through chemistry and  $f$  is a function for particle movement, which depends on the particle spatial location  $x_p$ , velocity  $u_p$ , mass  $m_p$ , temperature  $T_p$  and time  $t$ . The particle volume fraction is defined in (16), which can also be used for calculating fluid volume fraction ( $\theta_f$ ) as the sum of particle volume fraction and fluid volume fraction is one. The particle acceleration is calculated using (17), where  $\rho_p$  is the particle density,  $D_p$  is the drag function,  $\tau_p$  is the particle contact stress and  $\tau_D$  is the particle collision damping time (Snider et al., 2011), (Jacob, 2023). The drag function selected for the simulations is the Wen-Yu/Ergun model, which is a blend of two drag models from Wen-Yu and Ergun. The former is more appropriate for dilute flows, while the latter works better for dense flows (Jacob and Tokheim, 2023). The energy equation for the particle lumped heat transfer assumes that the temperature within a particle has no spatial gradients, further, there is no heat release from chemical reactions inside the particles, and any heat released from reactions on the particle surface does not significantly affect the surface energy balance shown in (18), where  $C_v$  is the particle specific heat capacity,  $\lambda_f$  is the fluid thermal conductivity,  $Nu_{f-p}$  represents the Nusselt number of heat transfer between particle and fluid and  $A_{sp}$  is the particle surface area (Snider et al., 2011), (Jacob, 2023).

$$\delta\dot{m}_p = - \iiint f \frac{dm_p}{dt} dm_p du_p dT_p \quad (15)$$

$$\theta_p = - \iiint f \frac{m_p}{\rho_p} dm_p du_p dT_p \quad (16)$$

$$\frac{du_p}{dt} = D_p(u_f - u_p) - \frac{1}{\rho_p} \nabla p - \frac{1}{\theta_p \rho_p} \nabla \tau_p + g + \frac{u_p - u_p}{\tau_D} \quad (17)$$

$$C_v \frac{dT_p}{dt} = \frac{1}{m_p} \frac{\lambda_f Nu_{f-p}}{2r_p} A_{sp} (T_f - T_p) \quad (18)$$

Considering the heat transfer between fluid and particles, in CPFD simulation the heat transfer coefficient between the two phases can be calculated in (19) Where  $Re$  and  $Pr$  are the Reynolds and Prandtl number, respectively,  $k_f$  is the fluid thermal conductivity and  $d_p$  is the particle diameter.

$$h_p = \left( 0.37 Re^{0.6} Pr^{0.33} + 0.1 \right) \frac{k_f}{d_p} \quad (19)$$

The heat transfer between the fluid and the wall assuming isothermal wall can be calculated in (20) Where  $h_{fw}$  shows the heat transfer coefficient between fluid and wall,  $h_l$  is the lean gas phase heat transfer coefficient and  $h_d$  represents the dense particle phase heat transfer coefficient. It is noted that  $f_d$ , which shows the fraction of contact time by the dense particle phase, depends on the particle volume fraction at the wall,  $\theta_p$ ,

and the close pack value fraction,  $\theta_{cp}$ . (Barracuda-Virtual-Reactor, 2022).

$$h_{fw} = h_l + f_d h_d \quad (20)$$

The radiation heat transfer from hot walls to the particles, which is calculated using “near wall model”, is shown in (21), (22), where  $T_w$  and  $T_p$  are the temperature of the wall and the particle, respectively.  $\sigma$  is the Stefan-Boltzmann constant,  $\varepsilon_{wp}$  is the (volume averaged) particle emissivity,  $F_{wp}$  represents the wall-particle view factor and  $A_w$  is the wall area (Barracuda-Virtual-Reactor, 2022).

$$q_{wp} = A_w F_{wp} \varepsilon_{wp} \sigma (T_w^4 - T_p^4) \quad (21)$$

$$\varepsilon_{wp} = \left( \frac{1}{\varepsilon_p} + \frac{1}{\varepsilon_w} - 1 \right)^{-1} \quad (22)$$

The only reaction in this reactor is the  $\text{CaCO}_3$  decomposition ( $\text{CaCO}_3 \rightarrow \text{CaO} + \text{CO}_2$ ). Equations (23) to (25) show the reaction kinetics. In (23),  $\frac{dm_j}{dt}$  represents the rate of decomposition or formation of each component  $j$  within the meal,  $M_j$  denotes the molecular mass of component  $j$ , while  $\vartheta_j$  signifies the stoichiometric coefficient related to the calcination reaction.  $p_{\text{CO}_2}$  stands for the partial pressure of  $\text{CO}_2$  present in the calciner,  $A_{sp}$  represents the surface area of the meal particles, and  $A_{eff}$  denotes the excess area fraction, accounting for voids present within the particles.  $k_D$  is the rate kinetics and  $p_{eq}$  is the equilibrium pressure (Mikulčić et al., 2012).

$$\frac{dm_j}{dt} = M_j \vartheta_j k_D (p_{eq} - p_{\text{CO}_2}) A_{sp} A_{eff} \quad (23)$$

$$k_D = 1.22 \times 10^{-5} \exp\left(-\frac{4026}{T_p}\right) \quad (24)$$

$$p_{eq} = 4.192 \times 10^{12} \exp\left(-\frac{20474}{T_p}\right) \quad (25)$$

### 3. SYSTEM DESCRIPTION

The designed FB reactor, which has a bottom cross section area of  $0.286 \text{ m}^2$  and a total height of  $1.880 \text{ m}$ , is shown in Fig. 2. This geometry corresponds to the ‘Electrically heated calciner’ shown in Fig. 1. The top cross section is cylindrical, and a transition piece connects these two sections, as shown in Fig. 2. Based on the design calculations the number of required electrically heated cylinders is 17, and each cylinder has a length of  $0.65 \text{ m}$ . The cylinders are arranged as shown in Fig. 2, and the bed height before expansion is about the same as the height level of the upper cylinder row ( $0.61 \text{ m}$ ). Raw meal is injected via rectangular inlets at the reactor sidewalls, below the first row of the hot cylinders.

The designed electrified calciner has two main sections. In the dense mixing section, at the bottom, fine raw meal particles are mixed with coarse  $\text{CaO}$  particles. The latter are easily fluidized and also act as a thermal reservoir for efficient heat transfer to the fine meal particles. The hot cylinders, which are

electrically heated, heat up both the coarse particles and the raw meal. The fine particles are calcined mainly in this dense section of the FB.

The  $\text{CO}_2$  resulting from the calcination and the injected fluidization gas is meant to entrain the calcined meal but leave the coarse particles in the dense bed. Hence, this upper dilute section is called the segregation section.

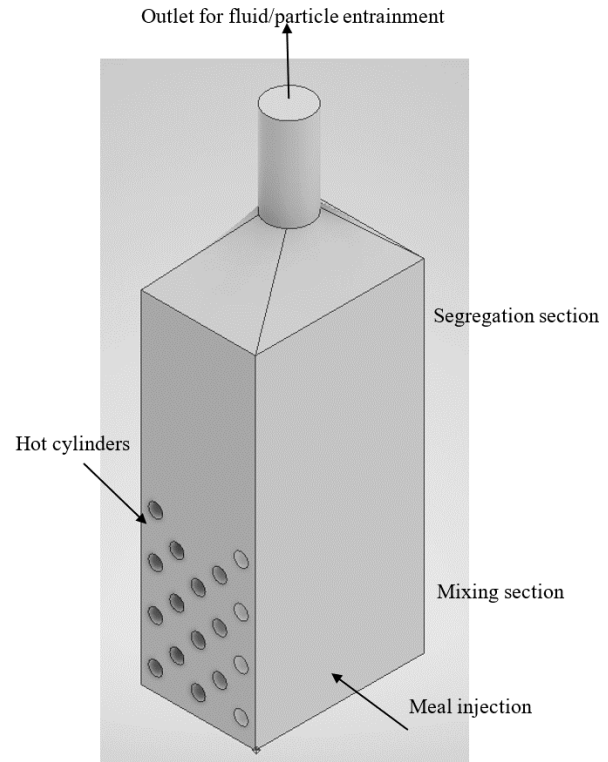


Fig. 2. The designed reactor geometry.

### 4. MODELLING AND SIMULATION

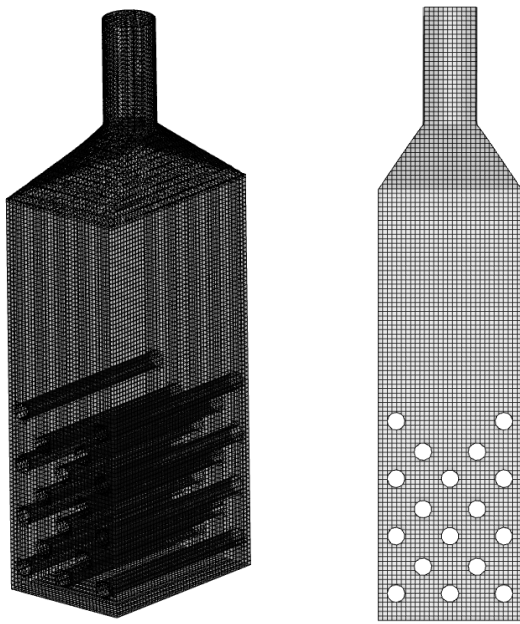
The designed reactor is modeled and investigated using the commercial software Barracuda, version 21.1.1.

The geometry of the modeled reactor is shown in Fig. 2. The simulation is done in 3D and is dynamic (time dependent). It includes flow particle interactions,  $\text{CaCO}_3$  decomposition and heat transfer between fluid and particles, between wall and particles and between fluid and wall.

A uniform grid was made in Barracuda as uniform cells will help reaching a stable and efficient simulation (Barracuda-Virtual-Reactor, 2022).

As a grid independence test, the pressure drop in the reactor was compared for four different meshes, having 54000, 130680, 192510 and 250767 cells, respectively. The last three meshes have less than 0.46 % difference in the result. The mesh with 192510 cells was finally selected as it should be fine enough for ensuring accuracy and coarse enough for not requiring too much calculation time. The final mesh is shown in Fig. 3. The design basis values are listed in Table 1. The particle size distributions of the coarse and fine particles are shown in Fig. 4.





(a) 3D mesh

(b) Mesh in the mid plane

Fig. 3. Mesh of the reactor (a) 3-dimensional, (b) The mid plane.

Table 1: Design basis values

Parameter	Description	Value	Unit
$w_{CaCO_3}$	Weight fraction of $CaCO_3$ in the raw meal	0.77	-
$\eta$	Degree of calcination	0.94	-
$T_{in,CO_2}$	Inlet temperature of fluidization gas	920	$^{\circ}C$
$T_{coarse}$	Initial temperature of coarse particle	920	$^{\circ}C$
$\dot{E}$	Electrical energy input	100	kW
$U$	Overall heat transfer coefficient	0.3	$kJ/(m^2 \cdot K)$
$D_{cyl}$	Cylinder diameter	0.055	m
$D_{top}$	Diameter at the reactor exit	0.170	m
$u_{out}$	Required entrainment velocity	1.7	m/s
$\rho_{coarse}$	Envelope density of coarse particles	1600	$kg/m^3$
$\rho_{rm}$	Envelope density of raw meal	2700	$kg/m^3$

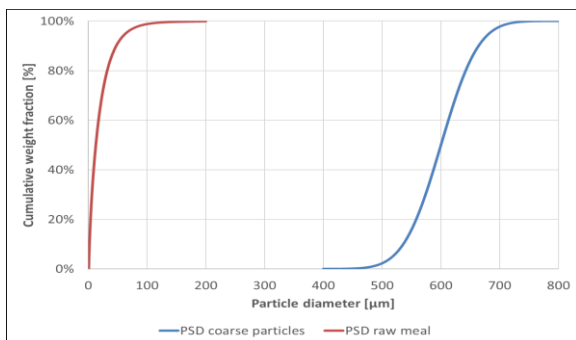


Fig. 4. Particle size PSD distribution of coarse and fine particles.

#### 4.1 Simulation setup

The initial timestep was set to 0.001 s. However, to maintain the accuracy, stability and calculation speed, the simulator checks the Courant number to stay in the preferred range ( $0.8 < CFL < 1.5$ ), and it adjusts the time step if needed (Barracuda-Virtual-Reactor, 2022). The simulation data was saved each 0.1 s.

To solve the fluid governing equations, an LES model was selected as the turbulence model. The conversion criteria were set to  $10^{-6}$  for the pressure and energy calculations,  $10^{-7}$  for the volume and velocity and  $10^{-9}$  the radiation. These values were suggested in Barracuda (Barracuda-Virtual-Reactor, 2022).

#### 4.2 Boundary conditions and case definitions

Seven different cases were defined, see Table 2. In all cases, the pressure at the outlet was set to 1 atm, a meal feed rate of 169 kg/h was applied, and the fluidization gas temperature was set to  $920^{\circ}C$ . The initial temperature was set to  $920^{\circ}C$  for both fluid and coarse particles in all cases. The initial height of the bed (before expansion) was 0.61 m.

Table 2. Case definitions

Case	$T_{in,rm}$ [ $^{\circ}C$ ]	$T_{cyl}$ [ $^{\circ}C$ ]	$u_F$ [m/s]
1	20	1050	0.3
2	20	1150	0.3
3	720	1150	0.3
4	850	1050	0.3
5	850	1100	0.3
6	850	1100	0.4
7	850	1100	0.8

#### 4.3 Calcination degree, particle entrainment and heat transfer coefficient

The calcination degree, the fine particle entrainment, the coarse particle entrainment, and the heat overall heat transfer coefficient are four important factors in the evaluation of the FB performance.

To calculate the calcination degree, a flux plane is defined at the outlet of the reactor, and the fluid mass flow rate exiting the reactor ( $\dot{m}_{CO_2,out}$ ) is found from the simulation results. This amount includes both the fluidization  $CO_2$  ( $\dot{m}_{CO_2,in}$ ) and the  $CO_2$  due to reaction. From design calculations it is known how much  $CO_2$  is injected and how much  $CO_2$  should be produced ( $\dot{m}_{CO_2,prod,design}$ ). So, the calculation of calcination degree is as follows:

$$\eta_{cal} = \frac{\dot{m}_{CO_2,out} - \dot{m}_{CO_2,in}}{\dot{m}_{CO_2,prod,design}} \times 100 \% \quad (26)$$

The flow rate of fine particles across the exit flux plane can be compared with the feed rate of particles. Sometime after



feeding is started, a pseudo-steady state should be reached where there is a balance between the inlet and outlet of fine particles (compensated for the calcination degree), cf. equations (4)-(8).

The flow rate of coarse particles across the exit flux plane should ideally be zero, but some entrainment is likely. The loss rate must be compensated by a make-up stream of coarse particles, and this make-up stream should be low compared with the meal feed rate. A small make-up may be acceptable.

The overall heat transfer coefficient can also be calculated based on the simulation results. This value can then later be verified in experiments. The calculation of this coefficient can be made reusing the energy balance in equations (1)-(3), but now utilizing that the area is known (from the design calculations) and treating the  $U$  value as an unknown.

## 5. RESULT AND DISCUSSION

All the simulations are continued until a pseudo-steady state has been reached, i.e., when the outlet mass flow rate remains steady. The results show that the pseudo-steady state is reached before 45 seconds for all the simulated cases. The avoided CO<sub>2</sub> emissions, when compared with a coal-fired cement kiln system, have been calculated based on typical operational settings for a full-scale system. There is a strong correlation to calcination degree, and the reduction varies from 44 % to 79 %, the latter being a very significant reduction. An example is shown in Fig. 5 for Case 2. Also, by injection of the raw meal in a narrow long injection area below the first row of hot cylinders a good mixing is reached in all the simulation cases. As an example, the distribution of injected raw meal particles just after 3 seconds is shown in Fig. 6. It shows that they are well distributed in the reactor. The particle volume fraction at the 45<sup>th</sup> second after reaching the pseudo steady state condition is shown in Fig. 7 which also reveals a good mixing condition.

The calcination degree in Case 2 can be calculated to 69 % using equation (23), noting that the mass flow of produced CO<sub>2</sub> based on design for cases 1-5 is 57 kg/h.

The effect of different operating conditions on the calcination degree is shown in Fig. 8. When the feed is not preheated (20 °C), a big part of the energy supply is spent on heating the raw meal (Case 1 and 2), and when the heater temperature is low (1050 °C, as in Case 1) the driving force for heat transfer is low, so, the calcination degree becomes low. The result is 50 % for Case 1 and 69 % for Case 2. Moreover, when the calcination degree is very low, such as in Case 1, the fine particles will not be properly entrained, and the system will not work.

Using a high cylinder temperature (1150 °C) and feeding raw meal preheated to 720 °C (Case 3) significantly increases the calcination degree from 69 % to 90 %, which is a very good value for stable production.

If the heater temperature is kept at 1050 °C and the meal is preheated to 850 °C (Case 4), the calcination degree reaches 81 %. However, if instead the heater temperature is increased

to 1150 °C and the meal is preheated to 720 °C (Case 5), the calcination degree reaches 90 %, i.e., the same as in Case 3. This shows the importance of being able to operate a high heater temperature in the bed.

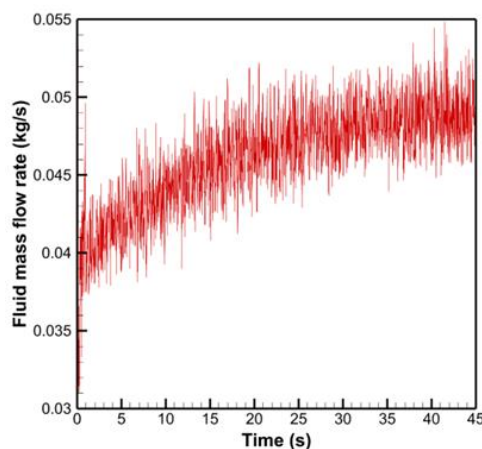


Fig. 5. Gas mass flow rate exiting the reactor (Case 2).

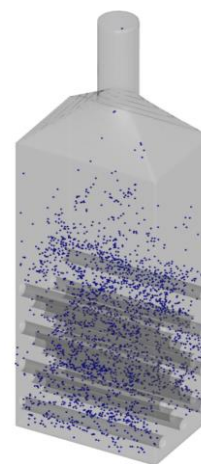


Fig. 6. The raw meal distribution after 3 seconds.

Apart from the calcination degree, particle entrainment is of great importance. The particle entrainment for two cases with low (Case 2) and high calcination degree (Case 3) is shown in Figs. 9 and 10, respectively.

In both cases, there is a very low amount of particle entrainment before 27 s. After this time, it appears that the reaction has reached steady state as the fluid flow rate exiting the reactor stopped increasing and only resonating around a specific value (Fig. 5). Also, the time integrated particle mass (Fig. 9) is increasing with a constant slope, showing that pseudo steady conditions is reached.

To check the mass balance and ensure that the pseudo steady condition is reached, the mass flow rate of raw meal injection should be equal to sum of the mass flow rate of fine particle entrainment and the produced CO<sub>2</sub> exiting the reactor. The total amount of fine particles injected into the reactor was 169 kg/h. Calculating from the 28<sup>th</sup> to the 45<sup>th</sup> second, the slope of entrainment of fine particles in Case 2 is 133 kg/h. The fluid mass flow rate exiting the reactor shows 175 kg/h, where 139 kg/h was the injected gas, so the production rate of CO<sub>2</sub> is 36

kg/h. The sum of 133 and 36 is 169, so there is a balance in the mass. Hence, the fine particles are entrained as expected.

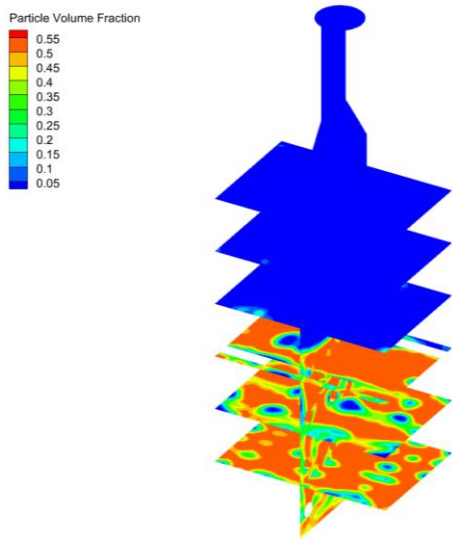


Fig. 7. Particle volume fraction in the 45<sup>th</sup> second (Case 2).

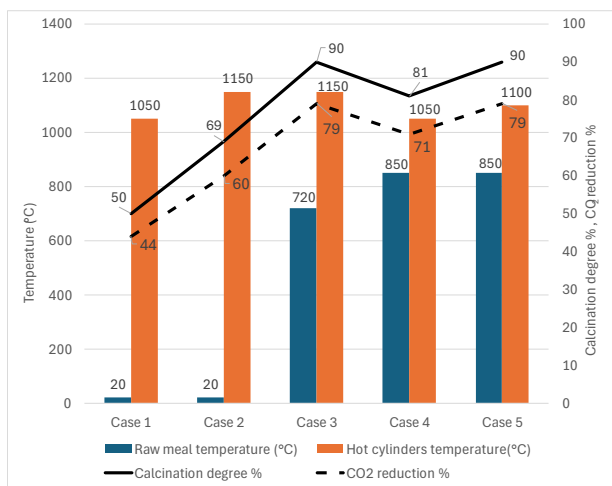


Fig. 8. Calculation degree and CO<sub>2</sub> reduction for cases 1-5.

In Case 2, a cold raw meal (20 °C) is injected, whereas in Case 3, the meal is preheated (720 °C) as it would be in an industrial system. The entrainment of coarse particles is about 0.9 % of the meal feed rate in Case 2, and about 2.4 % in Case 3. The entrainment of coarse particles is higher in the latter case due to the higher calcination degree and thereby a higher flow rate gas leaving the reactor. Anyway, both values are quite low, indicating an acceptable reactor design.

The fluidization velocity has a significant effect on the particle residence time in the bed and the entrainment of particles. Figure 11 shows that increasing the fluidization velocity from 0.3 m/s to 0.4 m/s reduces the calcination degree from 90 % to 80 %. The reason might be the residence time of the particles in the bed, which is reduced from 24.5 s to 20 s.

Increasing the velocity to 0.8 m/s gives a calcination degree of only 56 % and the residence time is only 5 s. However, this is not a stable value as the coarse particles are entrained from the

bed, as shown in Figure 12. The bed will be emptied of coarse particles in less than an hour, so operating at such a high velocity is not feasible.

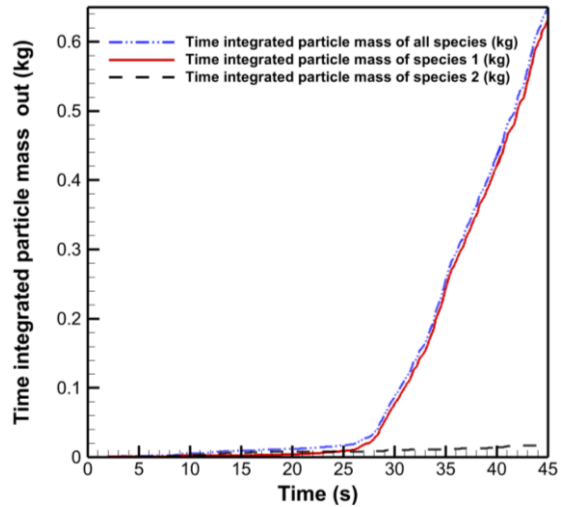


Fig. 9. Time integrated particle entrainment in Case 2. Species 1 is fine particles and species 2 is coarse particles.

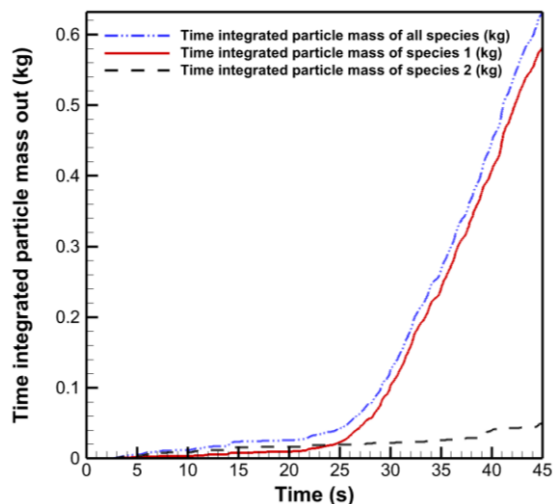


Fig. 10. Time integrated particle entrainment in Case 3. Species 1 is fine particles and species 2 is coarse particles.

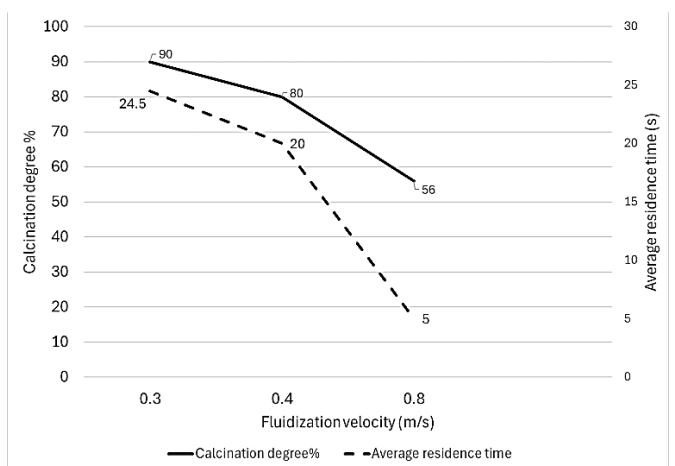


Fig. 11. The effect of fluidization velocity on calcination degree and particle residence time (Cases 5, 6 and 7).

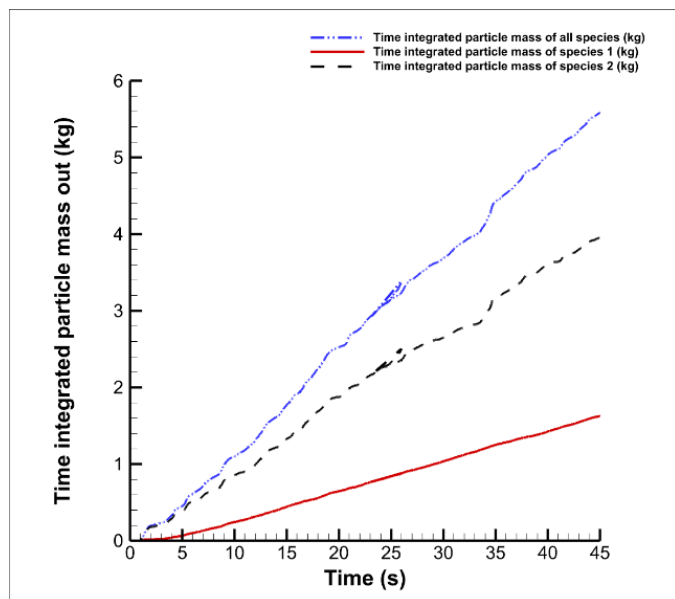


Fig. 12. Time integrated particle entrainment at a fluidization velocity of 0.8 m/s.

## 6. CONCLUSIONS

The CPFD results indicate that keeping the fluidization velocity constant but increasing the temperature of hot cylinders in the bed and or preheating the meal before injection can significantly increase the calcination degree. Moreover, keeping the temperatures constant but changing the fluidization velocity can significantly affect the calcination degree and the particle entrainment. Increasing the fluidization velocity to 0.8 m/s would give entrainment of all particles, whereas a fluidization velocity of 0.3 m/s works well for the designed reactor. The entrainment of coarse particles is in the range 1-2 % of the meal feed rate, which can be seen as an acceptable value for large-scale production. A high calcination degree of 90 % was reached when the meal was preheated, whereas no meal preheating gave a calcination degree of only 50%, which is due to more of the supplied energy being spent on heating the meal. At 90 % calcination, the resulting CO<sub>2</sub> reduction, compared to a coal-fired kiln system, is 79 %. The average residence time of the fine particles in the dense bed section after reaching the pseudo steady state condition was found to be in the range 24-26 s in five different cases. There was proper fluidization and good mixing of fine and coarse particles in all cases. All in all, calcination of raw meal in the designed electrified fluidized bed calciner appears to be feasible when the gas velocity in the reactor is kept within the specified operational limits.

## ACKNOWLEDGEMENTS

This study was done to support the project “Electrification of High Temperature and Flexible Technologies for Transforming Cement, Lime and Pulp Industry” (ELECTRA). University of South-Eastern Norway and Heidelberg Materials are two of the participants of this project, which has received funding from the European Union’s Horizon Europe research and innovation program under Grant Agreement No 101138392.

- Barracuda-Virtual-Reactor (2022). Barracuda Virtual Reactor version: 21.1.1 User Manual. Last updated on Feb 07, 2022. ed. Dupont, C. and Oberthür, S. (2015). The European Union. *Research handbook on climate governance*, 224-236.
- Geldart, D. (1973). Types of gas fluidization. *Powder Technology*, 7, 285-292.
- IEA (2020). *International Energy Agency, Paris, France, 2020*.
- IEA (2018). *Low-Carbon Transition in the Cement Industry*, IEA.
- Jacob, R. (2023). *CO<sub>2</sub> capture through electrified calcination in cement clinker production*. Ph.D dissertation, University of South-Eastern Norway.
- Jacob, R., Moldestad, B., and Tokheim, L.-A. (2022). *Fluidization of Fine Calciner Raw Meal Particles by mixing with coarser Inert Particles – Experiments and CPFD Simulations*.
- Jacob, R. and Tokheim, L.-A. (2023). CPFD simulation of an electrically heated fluidized bed calciner with binary particles. *Energy Conversion and Management: X*, 20, 100444.
- Kim, J. H., Bae, J. W., Nam, J., Kim, S. D., Choi, J.-H., and Lee, D. H. (2014). Entrainment of Geldart C particles in fluidized beds with binary particles. *Korean Journal of Chemical Engineering*, 31, 2094-2100.
- Mikulčić, H., von Berg, E., Vujanović, M., Priesching, P., Perković, L., Tatzschl, R., and Duić, N. (2012). Numerical modelling of calcination reaction mechanism for cement production. *Chemical Engineering Science*, 69, 607-615.
- Nikolakopoulos, A., Steriotis, T., Charalampopoulou, G., Karagiannakis, G., Dimitrakakis, D., Michalis, V., and Katsiotis, M. (2024). Reducing carbon emissions in cement production through solarization of the calcination process and thermochemical energy storage. *Computers & Chemical Engineering*, 180, 108506.
- Samani, N. A., Jayarathna, C. K., and Tokheim, L.-A. (2020). CPFD simulation of enhanced cement raw meal fluidization through mixing with coarse, inert particles. *SIMS 61*. Linköping.
- Snider, D. M. (2001). An Incompressible Three-Dimensional Multiphase Particle-in-Cell Model for Dense Particle Flows. *Journal of Computational Physics*, 170, 523-549.
- Snider, D. M., Clark, S. M., and O'Rourke, P. J. (2011). Eulerian-Lagrangian method for three-dimensional thermal reacting flow with application to coal gasifiers. *Chemical Engineering Science*, 66, 1285-1295.
- Tokheim, L.-A., Mathisen, A., Øi, L. E., Jayarathna, C., Eldrup, N., and Gautestad, T. (2019). Combined calcination and CO<sub>2</sub> capture in cement clinker production by use of electrical energy. *TCCS-10*. Trondheim, Norway.

# Performance Analysis of Advanced Wells in Reservoirs Using CO<sub>2</sub> Enhanced Oil Recovery

Prakash Bhattarai, Soheila Taghavi Hosnaroudi, Britt M. E. Moldestad

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway,  
Norway

E-mail corresponding author: [soheila.t.hosnaroudi@usn.no](mailto:soheila.t.hosnaroudi@usn.no)

**Abstract:** Oil and gas will remain an important source of energy for years and it is crucial to improve oil recovery with less carbon footprint to meet the future energy demands. Carbon capture utilization and storage offers a potential solution to mitigate the effects of anthropogenic CO<sub>2</sub> and to reduce the direct CO<sub>2</sub> emissions from stationary sources into the atmosphere. The captured CO<sub>2</sub> can be utilized to enhanced oil recovery (EOR) and is injected into the depleted oil fields or saline aquifers, or into the oil fields for storage and/or EOR. However, the injected CO<sub>2</sub> can be reproduced without contributing to EOR. This is due to the breakthrough of CO<sub>2</sub> into the well. Also, the corrosive mixture of CO<sub>2</sub> and water can be produced from the production well. This may cause damages to the pipeline and process equipment on the platform. Autonomous inflow control valves (AICVs) can mitigate these problems. They may reduce or stop the reproduction of CO<sub>2</sub> from the zones with CO<sub>2</sub> breakthrough and reduce the production of mixture of CO<sub>2</sub> and water. The main objective of this study is modelling and simulation of oil production in a heterogenous reservoir using CO<sub>2</sub>-EOR in combination with AICVs. The simulation models are developed using an industry standard software. The outcome of numerical simulations is analyzed to study the effect of various parameters on oil recovery. In addition, the impact of AICVs on EOR is assessed against perforated casing completion (without AICV). The results demonstrate that oil recovery factor, water cut, and cumulative gas production are better in the wells completed with AICVs than perforated casing completion. This will result into both increased oil production and a better CO<sub>2</sub> storage potential.

**Keywords:** CO<sub>2</sub>-EOR, Carbon capture utilization and storage, Autonomous inflow control valve, Advanced wells, Miscible injection, WAG

## 1. INTRODUCTION

The global energy crisis has led to a movement towards the development of clean energy technologies to ensure energy security. The oil and gas industry accounts for more than 50% of the global energy supply with oil holding approximately one-third of the global energy supply (*Supply – Key World Energy Statistics 2021 – Analysis*, n.d.) According to the World Energy Outlook 2023 (*World Energy Outlook 2023 – Analysis*, n.d.) published by the International Energy Agency (IEA), the demand for oil is estimated to reach its peak in near future. In conjunction with growing energy demand, the carbon dioxide (CO<sub>2</sub>) emission from oil production is escalating. Thus, it is essential to improve the production of oil, simultaneously reducing the carbon footprint. In this prospect, CO<sub>2</sub> Enhanced Oil Recovery (EOR) has emerged as the prospective solution to support Carbon Capture, Utilization, and Storage (CCUS) by ensuring permanent storage of CO<sub>2</sub> in geological formation while offering commercial opportunities to oil industries.

CO<sub>2</sub> EOR is the process of injecting CO<sub>2</sub> into the depleted oil reservoirs to improve oil recovery. It is a proven method and has been in commercial practice for several decades. Based on the miscibility of injected CO<sub>2</sub> with oil at reservoir conditions, the process is divided into miscible and immiscible. In miscible flooding, the CO<sub>2</sub> is injected at operating pressure

above Minimum Miscibility Pressure (MMP). MMP is the minimum pressure necessary for CO<sub>2</sub> to be miscible in the oil at reservoir conditions. The injection of CO<sub>2</sub> is carried out at operating pressure below MMP in immiscible flooding. Miscible flooding has higher efficiency than immiscible flooding due to greater sweep efficiencies. The CO<sub>2</sub> flooding process often faces the challenges of viscous fingering, channeling, and gravity override; therefore, the water is injected alternating with gas referred to as the Water Alternating Gas (WAG) process. The injection of water reduces the problem associated with gas flooding and assists in maintaining reservoir pressure above MMP (A. Khan et al., 2021). Even though the miscible CO<sub>2</sub> WAG process is very effective in recovering residual oil, the injected CO<sub>2</sub> can break through into the production well without affecting in EOR process, and the corrosive mixture of CO<sub>2</sub> and water can damage the production facility. Often these problems lead to the permanent shutting down of the reservoir.

By utilizing advanced wells in combination with Flow Control Devices (FCDs), the performance of CO<sub>2</sub> EOR processes can significantly be improved. The FCDs like Autonomous Inflow Control Valves (AICVs) have the potential to solve the issues of early breakthroughs, production of unwanted fluids, gas and water conning, and non-uniform pressure distribution in horizontal wells. Therefore, the implementation of AICVs in



CO<sub>2</sub> WAG EOR, can enhance the oil recovery, and reduce the cost associated with handling unwanted fluids.

The objective of this paper is to study the performance of AICVs integrated in the reservoir subjected to the miscible WAG EOR utilizing CO<sub>2</sub>. This aim is realized by implementing and analyzing the benefits and constraints of AICVs against the standard well perforations through numerical simulations for various WAG operating parameters. The experimental data presented by (Taghavi et al., 2022) is utilized to numerically simulate the functional behavior of AICVs in CMG STARS simulator.

## 2. CO<sub>2</sub> ENHANCED OIL RECOVERY

Enhanced oil recovery with the injection of CO<sub>2</sub> has played significant role in the tertiary recovery of residual oil since its first pilot project at Mead Strawn Field in 1964 (Holm and O'Brien, 1971). The idea was patented by Whorton Brownscombe in 1952 (Whorton et al., 1952) and the method was successfully implemented in commercial project at the Kelly-Snyder Field in 1972.

The CO<sub>2</sub> EOR is classified primarily into two types of miscible and immiscible, based on the solubility of injected CO<sub>2</sub> in the reservoir oil. The reservoir conditions, composition of injected gas, and interaction of CO<sub>2</sub> with reservoir fluid are the parameters that determine whether the CO<sub>2</sub> EOR is miscible or immiscible. The miscible and immiscible processes differ in the mechanism and subsequently, in the recovery factor (Haynes and Alston, 1990).

### 2.1 Miscible CO<sub>2</sub> EOR

The miscibility criterion is attained at the reservoir pressure higher than MMP. The MMP is determined based on the temperature and composition of the oil in the reservoir (Haynes and Alston, 1990). At supercritical condition, CO<sub>2</sub> possesses the density close to liquid phase (0.6 – 0.8 g/cm<sup>3</sup>) while viscosity remains low close to gas viscosity. The supercritical CO<sub>2</sub> dissolves into oil causing oil to swell and a reduction in the viscosity thereby improving the mobility of oil (Mansour et al., 2019).

Miscibility in the reservoir is achieved through two processes of First Contact Miscibility (FCM) and Multiple Contact Miscibility (MCM). In FCM, the injected CO<sub>2</sub> mixes with oil in reservoir in different proportions on first contact to generate a homogeneous (single-phase) solution (Clark et al., 2013). In MCM process, miscibility is attained through the vaporization of hydrocarbons into CO<sub>2</sub>, and diffusion of CO<sub>2</sub> into reservoir oil. The former MCM process is termed as vaporizing gas drive while later is called condensing gas drive (Green and Willhite, 1998). In theory, the miscible CO<sub>2</sub> EOR can have recovery factor up to 90%.

Oil swelling, viscosity reduction, mobility ratio reduction, interfacial tension reduction, vaporization of light oil, and wettability change are the mechanisms contributing towards the improvement of oil recovery in miscible CO<sub>2</sub> EOR.

### 2.2 Miscible CO<sub>2</sub> WAG EOR

Water Alternating Gas (WAG) is an EOR technique that involves injection of water and gas in a cyclic manner. The aim of WAG technique is to improve oil production utilizing microscopic displacement of the oil with injection of gas and macroscopic sweep with injection of water simultaneously (J. Wang et al., 2008).

The WAG process is affected by several factors. WAG ratio is the volumetric ratio of the injected water to the injected gas at reservoir condition (M. Y. Khan and Mandal, 2022). The WAG ratio strongly affects the oil recovery. At low WAG ratios the system works like a gas flood as the volume of injected water is low. This results in a poor vertical sweep associated with gas fingering, channeling, and early breakthrough. At high WAG ratios, the waterfront travels faster and blocks the gas from contacting the oil, consequently reducing the microscopic displacement (Belazreg et al., 2019). Designing the WAG ratio to its optimum value is important as it ensures a higher economic oil recovery by controlling the water cut, mobility ratio, and gas production (S. Chen et al., 2010). The optimum WAG ratio depends on the impacts of gravity overrides, reservoir heterogeneity, capacity of injection wells, economic constraints, etc. (Rogers and Grigg, 2000).

The WAG cycle time refers to the total duration of gas or water injection during an injection cycle in the WAG process. The cycle time directly affects the economy of EOR projects. Zhang et al., (2010) showed that shorter cycle time i.e., higher number of cycles, increases the oil production. According to (B. Chen and Reynolds, 2016) and (Abdullah and Hasan, 2021), decreasing cycle time increases the oil recovery. Araujo Cavalcante Filho et al., (2020) assumed that the shorter cycle time discourages the gravity segregation thus improving the oil recovery. The WAG process can be started at the beginning (initial WAG) and at the later phase (post WAG) of reservoir development. Initial WAG provides better incremental oil recovery than post WAG. However, the overall economics of the project will be affected. Initial WAG accelerates the oil production in both heterogeneous and homogeneous reservoirs (M. Y. Khan and Mandal, 2022).

The distance between injector and producer wells is termed as well spacing. According to (Christensen et al., 2001), well spacing directly affects the sweep efficiency and average reservoir pressure in the WAG process. The gravity segregation dominates if the well spacing is high causing a reduction in oil recovery. While, lower well spacing enhances the response time of WAG process. However, due to the short circuiting of injected gas, the oil recovery reduces.

The heterogeneity and stratification strongly affect the sweep patterns. Highly porous and permeable rocks provide better sweep efficiency resulting in improved oil recovery (Li et al., 2015).

The CO<sub>2</sub> WAG EOR involves the process of injecting CO<sub>2</sub> and water alternately. In the miscible CO<sub>2</sub> WAG EOR, CO<sub>2</sub> is

injected in the reservoir when the pressure is above MMP (Dai et al., 2014) as depicted in Fig. 1.

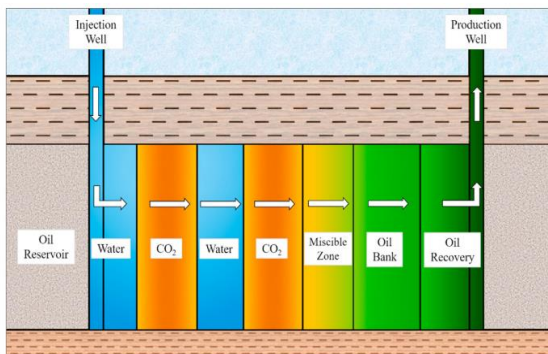


Fig. 1. Schematic of miscible CO<sub>2</sub> WAG EOR.

According to (Han and Gu, 2014), the miscibility of CO<sub>2</sub> with light oil is obtained at low MMP, at the same time, the injected water maintains the pressure above MMP, therefore CO<sub>2</sub> WAG has technical benefits. Skauge and Stensen (2003), in their review of 72 fields using WAG with hydrocarbon or non-hydrocarbon gases, reported that the miscible CO<sub>2</sub> WAG had highest average improved oil recovery of 10% of original oil in place (OOIP).

In the experiment conducted by (Yan et al., 2017), supercritical CO<sub>2</sub> played an important role and the authors found that the miscible CO<sub>2</sub> WAG injection should improve the oil production better than either CO<sub>2</sub> or water flooding. Lei et al., (2016) reported improvement in oil recovery factor between 12 -17%, and Q. Wang et al., (2020) found that the ultimate oil recovery reached to 73% from 52% due to the implementation of CO<sub>2</sub> WAG EOR process.

Miscible CO<sub>2</sub> WAG EOR is associated with problems of gravity overrides, early breakthrough, and gas channeling. In addition, Wang et al., (2020) reported a reduction in the permeability of the core due to asphaltene deposition and reaction between CO<sub>2</sub>, rock, and brine (see Fig. 2). The formation of weak acid takes place when CO<sub>2</sub> and water react with each other. The corrosive weak acid is damaging to production wells and the process equipment (Halland et al., 2013).

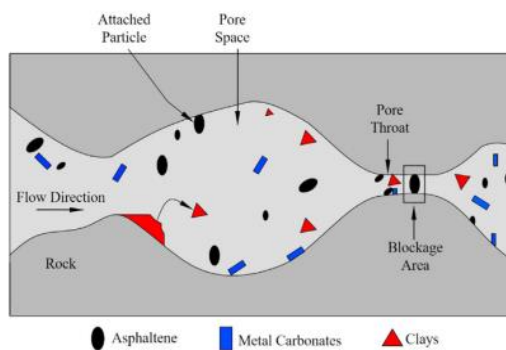


Fig. 2. Blockage of pore throat due to asphaltene deposition.

### 3. ADVANCED WELLS WITH AICV

Horizontal wells are a significant development to maximize the reservoir contact with oil in the reservoir. Increased interaction with the reservoir rock enables more effective fluid injection and drainage. The introduction of horizontal wells greatly raised the recovery factor (Behnoud et al., 2023).

Long horizontal wells allow exposure to a larger reservoir area. However, this may result in a substantial pressure difference between the toe and the heel section of the production well. This is due to the reduction in pressure caused by friction between the fluid travelling through the pipe and the inner pipe surface. As a result, there is a higher pressure drop between the wellbore and the reservoir at the heel than at the toe. Thus, the heel of the well receives more reservoir fluid flow than other regions. This phenomenon is called heel to toe effect (Mahmood et al., 2016). The difference of drawdown between the heel and the toe results in early breakthrough at the heel of the reservoir. The breakthrough is also affected by the heterogeneity of the reservoir. The considerable pressure differences along the wellbore increases the likelihood of early gas and water breakthrough, lowering the recovery efficiency. Innovative solutions, such as the Autonomous Inflow Control Valve (AICV) technology, are required to address the challenges caused by early gas and /or water breakthrough.

#### 3.1 Role of AICVs in advanced wells

AICVs are the flow control devices that work autonomously to restrict the flow of gas and water to the production wells. The valve distinguishes between fluids based on their viscosity and density. An AICV is shown in Fig. 3 (Ismail et al., 2021).



Fig. 3. View of an AICV.

AICVs are devices capable of preventing flow of unwanted fluids like gas and water, through annulus into the production wells. When gas or water flows into the production well, AICVs automatically shut off in the breakthrough zones, avoiding unwanted fluids from entering the production well. Furthermore, using AICVs eliminates the risks, costs, and logistical issues associated with removing, transporting, and handling the unwanted fluids (Aakre et al., 2014).

### 4. MODELING OF RESERVOIR AND WELLBORE

The reservoir simulation software CMG 2022.10 general release by Computer Modelling Group Ltd. is used for the modeling and simulation of CO<sub>2</sub> WAG EOR. STARS, a reservoir simulator application included in CMG software, has been selected to simulate the recovery process. The wellbore model in STARS contains a Flexible Wellbore Model (FlexWell). FlexWell allows the integration of advanced well completions including AICVs and concentric wellbores. Even



though, Flexwell is solved independently in the CMG software, it is coupled with the STARS simulator. The dataset for the STARS simulator is created in the Builder module. The data of the reservoir model including porosity, permeability, grid distribution, properties of fluid components, rock fluid properties, and well model are obtained from (Taghavi et al., 2023). The detailed descriptions of the reservoir and well models are presented in the following subsections.

4.1 Construction of reservoir model in CMG

A heterogeneous reservoir with a depth of 181 ft in k-direction (vertical) is considered for the simulation. The reservoir is divided into 9375 grid blocks with 25\*25\*15 grid blocks in i, j, and k-directions respectively. The size of each grid block in i and j-directions is 130 ft while the grid blocks along the k-direction have variable thicknesses ranging from 4 to 33 ft. The porosity variation in the reservoir is between 0.234 and 0.317. The horizontal permeability along i and j-directions, ranges from 0 to 2588 mD, while permeability in the vertical (k) direction is half of the horizontal permeability varies from 0 to 1294 mD. The distribution of permeability along the horizontal and vertical directions are presented in Figs. 4 and 5, respectively. The initial pressure and temperature of the reservoir are 4200 psi and 186 °F, respectively.

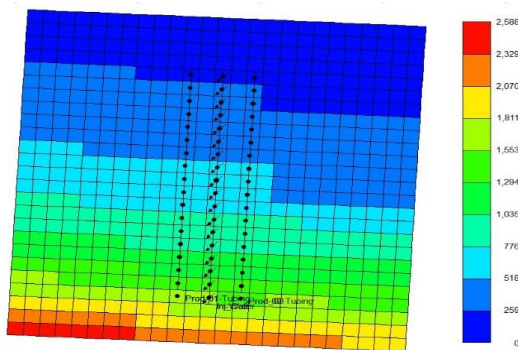


Fig. 4. Variation of horizontal permeability along the wells (i-j plane).

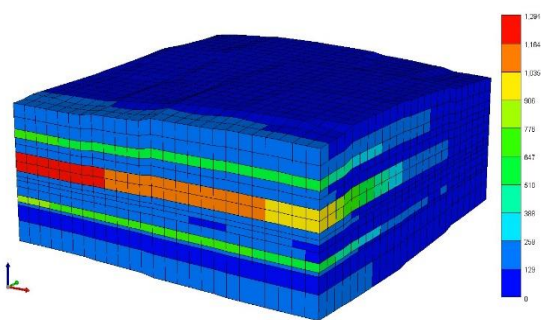


Fig. 5. 3D view of the reservoir with distribution of vertical permeability.

4.2 Description of well model

The well model consists of approximately 2210 ft long three horizontal wells placed along j-direction. Two producer wells are placed on either side of central injector well as shown in

Fig. 6. The producers and injector wells are placed at the same height in the region with higher permeabilities. The distance between the producer wells and injector well is 260 ft.

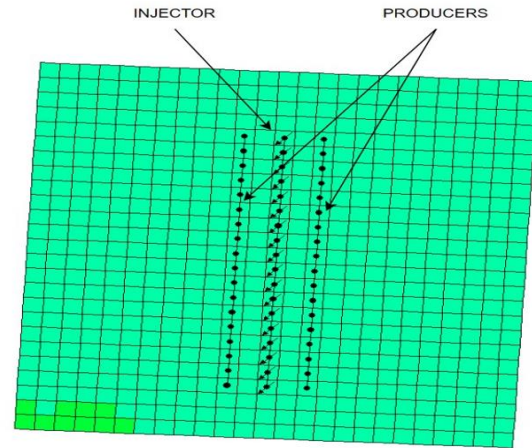


Fig. 6. Placement of wells in the reservoir.

The well constraints applied to all producer and injector wells are listed in Table 1.

Table 1. Well constraints

Wells	Function	Constraints
Inj_CO2	CO2 Injector	MAX BHP 5000 psi Variable based on WAG ratio
Inj_Water	Water Injector	MAX BHP 5000 psi Variable based on WAG ratio
Prod-01	Producer	MIN BHP 700 psi MAX STL 2500 bbl/day
Prod-01-Tubing	Producer	MIN BHP 500 psi MAX STL 2500 bbl/day
Prod-02-Tubing	Producer	MIN BHP 500 psi MAX STL 2500 bbl/day
Prod-02	Producer	MIN BHP 700PSI MAX STL 2500 bbl/day

To analyse the performance of AICVs, the producer wells are completed with two different settings. In the first completion setting, both producer wells have standard perforations i.e., without any flow control devices. In the second setting, the producer wells are completed with AICVs. In both cases, the production takes place from 18 zones and each production zone is isolated with packers. Each isolated production zone is completed with 12 AICVs in the second completion scenario. FlexWells are coupled with FCD tables developed by (Taghavi et al., 2023) to simulate the behaviour of AICVs so that, the flow of water and pure CO<sub>2</sub> gas or supercritical CO<sub>2</sub> can be restricted.

The simulations are carried out for eight years starting from 1<sup>st</sup> of January 2023 to 1<sup>st</sup> of January 2031. The well events for injection of CO<sub>2</sub> and water are setup to ensure that the injection of water and CO<sub>2</sub> takes places in the cycle of three months, while production wells are operational throughout the

simulation period. The timeline view of the CO<sub>2</sub> WAG process is presented in Fig. 7.

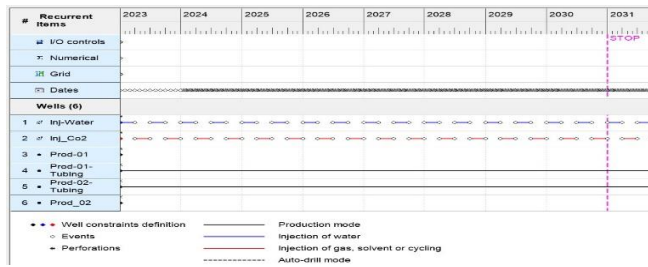


Fig. 7. Timeline view of the well events.

### 5. RESULTS AND DISCUSSION

The performance of the WAG process is influenced by various parameters. The operational parameters of WAG ratio, well spacing, and permeability are selected for performance analysis of AICVs against standard well perforations.

#### 5.1 WAG ratio

The total volume of water or gas injected in one cycle at a WAG ratio of 1:1 is approximately 0.45 hydrocarbon pore volume (HCPV). The injection rates of water and CO<sub>2</sub> at different WAG ratios at standard conditions are presented in Table 2.

Table 2. Injection rates at different WAG ratios

WAG ratio	Water injection rate (bbl/day)	CO <sub>2</sub> injection rate (ft <sup>3</sup> /day)
1:1	1000	2650000
2:1	2000	2650000
3:1	3000	2650000
4:1	4000	2650000
1:1.5	1000	3980000
1:2	1000	5310000
1:3	1000	7950000

Figure 8 to Figure 10 show the difference in oil recovery factor, cumulative GOR, and water cut at different WAG ratios between wells completed with AICVs and perforated casing. The wells completed with AICVs recovered on average 0.2% more oil than the perforated casing. The simulation case with AICVs has 25 ft<sup>3</sup>/bbl less cumulative GOR than the case with perforated casing at WAG ratio of 1:3.

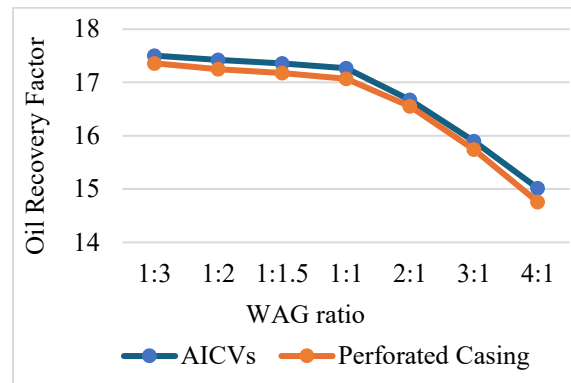


Fig. 8. Comparison of oil recovery factor between AICVs and perforated casing.

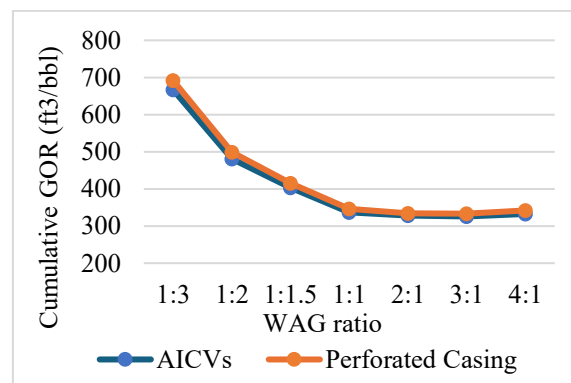


Fig. 9. Comparison of cumulative GOR between AICVs and perforated casing.

At the WAG ratio of 4:1, the water cut is 1.5% lower in the case of AICVs than perforated casing completion.

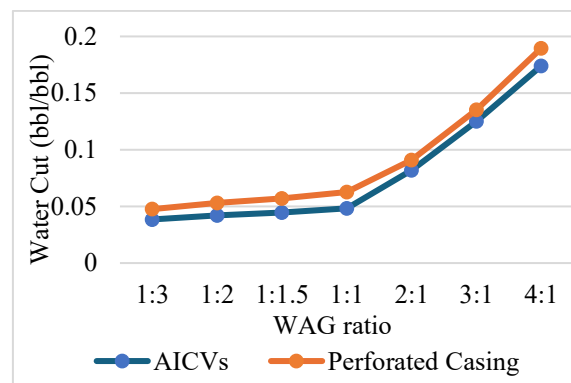


Fig. 10. Comparison of water cut between AICVs and perforated casing.

The GOR profile of AICVs and perforated casing along the well shown in Fig. 11 demonstrates that the GOR values at the heel section of the production well completed with perforated casing peaking at approximately 3000 ft<sup>3</sup>/bbl at the third production zone however, the GOR values are evenly distributed along the well completed with AICVs.

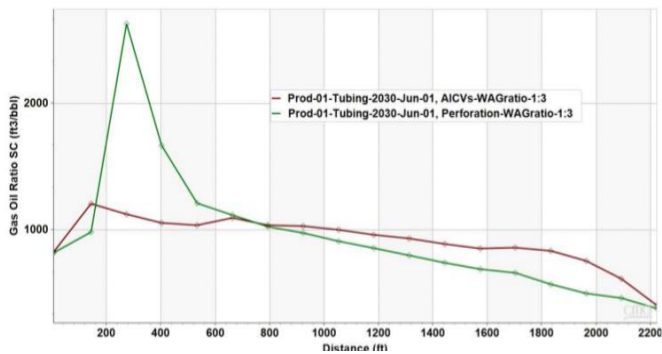


Fig. 11. Comparison of GOR profile along the well between AICVs and perforated casing

The water flow rate along the well shown in Fig. 12 illustrates that water production rate is significantly high in the heel section of the well in perforated casing scenario while the AICVs have balanced the water rate along the well.

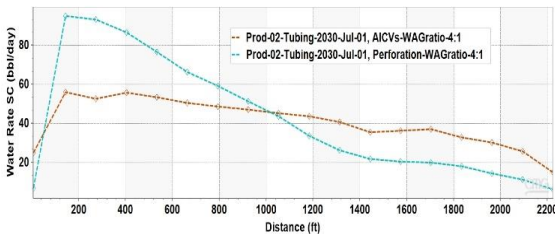


Fig. 12. Comparison of water rate along the well between AICVs and perforated casing.

### 5.2 Well spacing

Apart from the original well spacing of 260 ft between producer and injector wells, the simulations are conducted for all the WAG ratios and both completion settings at well spacing of 130 ft and 390 ft. The well distance is changed by shifting the producer wells while keeping the injector well at the same position. However, the simulation period for these cases is reduced to 5 years.

The plots displaying a comparison of the oil recovery factor and water cut of the AICVs and perforated casing completion scenarios at different WAG ratios for both well spacing of 130 ft and 390 ft are presented in Fig. 13 to Fig. 16. The differences in the results of AICVs and perforated casing in both well spacings are very marginal as in the previous case.

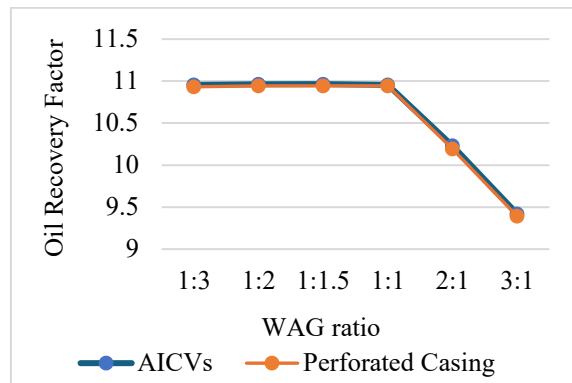


Fig. 13. Oil recovery factor at well spacing of 130 ft.

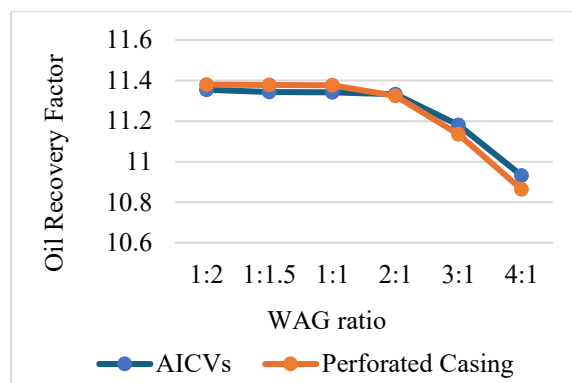


Fig. 14. Oil recovery factor at well spacing of 390 ft.

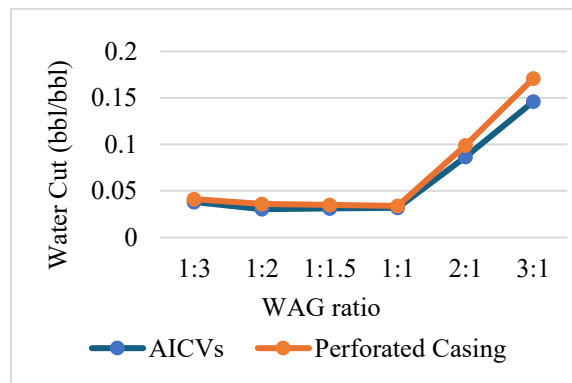


Fig. 15. Water cut at well spacing of 130 ft.

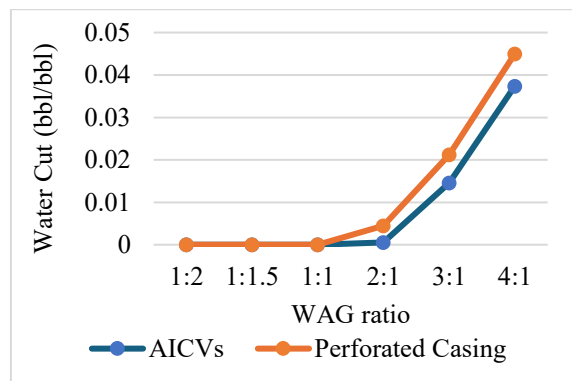


Fig. 16. Water cut at well spacing of 390 ft.

However, a significant difference for well spacing of 130 ft at a WAG ratio of 1:3 is presented in Fig. 17. The peaks in the figure indicate the breakthrough of gas in the wells completed with perforated casing while AICVs restricted the breakthrough of gas.



Fig. 17. Comparison of GOR at reservoir condition.

### 5.3 Permeability

The permeability of the reservoir is changed to twice of its original permeability and the simulation cases are developed at this changed permeability distribution for both completion settings of producer wells at all WAG ratios.

Figure 18 shows the water cut for both well completion scenarios in the reservoir with doubled permeability. The plot shows that AICVs are better at resisting inflow of water in comparison to perforated casing completion.

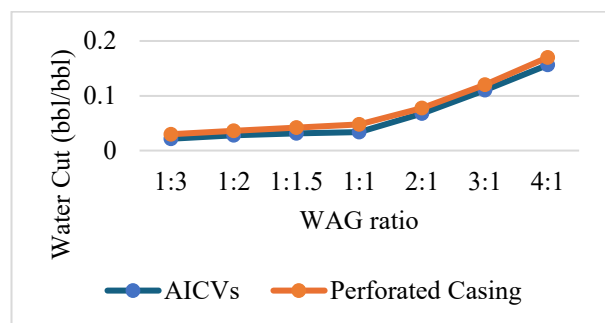


Fig. 18. Water cut in the reservoir with doubled permeability.

The profile plots of GOR and water rate as shown in Figs. 19 and 20, respectively, indicate that AICVs perform better in regulating inflow along the length of the well by mitigating the effect of reservoir heterogeneity and heel to toe effect associated with horizontal wells. The figures also show that the coning effect in the well with perforated casing completion is prevalent thus, the water rate is maximum at the heel section.

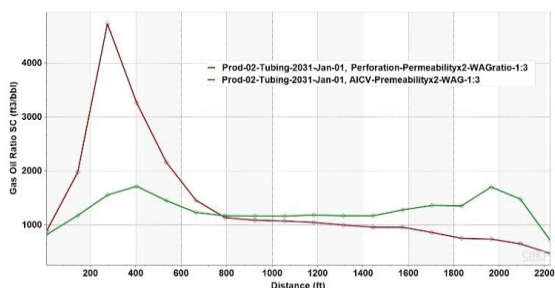


Fig. 19. GOR along the well in the reservoir with doubled permeability.

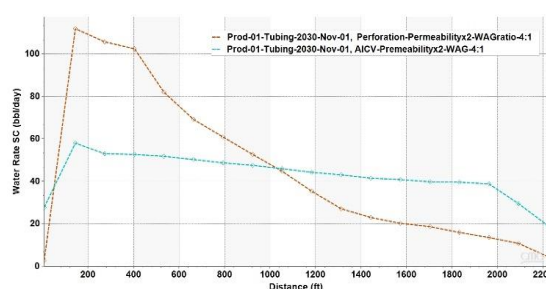


Fig. 20. Water rate along the well in the reservoir with doubled permeability.

## 6. CONCLUSIONS

The performance of AICV in a miscible CO<sub>2</sub> WAG EOR process is investigated through numerical simulations using the reservoir simulator CMG STARS. The outcome of this study demonstrates that the oil recovery of the miscible CO<sub>2</sub> WAG EOR in the horizontal well can be improved by utilizing AICVs. AICVs restricts the production of unwanted fluids. Besides, it is demonstrated that the WAG ratio, well spacing, and permeability of reservoir influence the oil production process. In overall, the miscible CO<sub>2</sub> WAG EOR method has significant potential to address the ever-growing energy demand, at the same time, resolving the problems associated with the increase of atmospheric CO<sub>2</sub>. Also, AICVs have potential to contribute substantially to improve the miscible CO<sub>2</sub> EOR processes.

## REFERENCES

Aakre, H., Halvorsen, B., Werswick, B., and Mathiesen, V. (2014, September 24). *Autonomous Inflow Control Valve for Heavy and Extra-Heavy Oil*. SPE Heavy and Extra Heavy Oil Conference: Latin America. doi:10.2118/171141-MS

Abdullah, N., and Hasan, N. (2021). The implementation of Water Alternating (WAG) injection to obtain optimum recovery in Cornea Field, Australia. *Journal of Petroleum Exploration and Production*, 11(3), 1475–1485. doi:10.1007/s13202-021-01103-7

Andrei, M., De Simoni, M., Delbianco, A., Cazzani, P., and Zanibelli, L. (2010). *Enhanced Oil Recovery with CO<sub>2</sub> Capture and Sequestration*. <https://www.osti.gov/etdeweb/biblio/21403705>

Bagci, A. S. (2007). Immiscible CO<sub>2</sub> Flooding through Horizontal Wells. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 29(1), 85–95. doi:10.1080/009083190934022

Behnoud, P., Khorsand Movaghar, M. R., and Sabooniha, E. (2023). Numerical analysis of pore-scale CO<sub>2</sub>-EOR at near-miscible flow condition to perceive the displacement mechanism. *Scientific Reports*, 13(1), Article 1. doi:10.1038/s41598-023-39706-1

Belazreg, L., Mahmood, S. M., and Aulia, A. (2019). Novel approach for predicting water alternating gas injection recovery factor. *Journal of Petroleum Exploration and Production Technology*, 9(4), 2893–2910. doi:10.1007/s13202-019-0673-2



- Chen, B., and Reynolds, A. C. (2016). Ensemble-based optimization of the water-alternating-gas-injection process. *SPE Journal*, 21(3), 786–798. Scopus. doi:10.2118/173217-pa
- Chen, S., Li, H., Yang, D., and Tontiwachwuthikul, P. (2010). Optimal parametric design for water-alternating-gas (WAG) process in a CO<sub>2</sub>-miscible flooding reservoir. *Journal of Canadian Petroleum Technology*, 49(10), 75–82. Scopus. doi:10.2118/141650-PA
- Christensen, J. R., Stenby, E. H., and Skauge, A. (2001). Review of WAG Field Experience. *SPE Reservoir Evaluation and Engineering*, 4(02), 97–106. doi:10.2118/71203-PA
- Clark, N., Shearin, H. M., Schultz, W. P., Garms, K., and Moore, J. L. (2013). Miscible Drive—Its Theory and Application. *Journal of Petroleum Technology*, 10, 11–20. doi:10.2118/1036-G
- Dai, Z., Middleton, R., Viswanathan, H., Fessenden-Rahn, J., Bauman, J., Pawar, R., Lee, S.-Y., and McPherson, B. (2014). An Integrated Framework for Optimizing CO<sub>2</sub> Sequestration and Enhanced Oil Recovery. *Environmental Science and Technology Letters*, 1(1), 49–54. doi:10.1021/ez4001033
- de Araujo Cavalcante Filho, J. S., Santos Silva, V. L., and de Sant'Anna Pizarro, J. O. (2020, August 30). *Assessment of Miscible WAG Injection Performance in a Giant Carbonate Reservoir using Analytical and Numerical Approaches*. SPE Improved Oil Recovery Conference. doi:10.2118/200333-MS
- Green, D. W., and Willhite, G. P. (1998). *Enhanced oil recovery*. Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers.
- Halland, E. K., Riis, F., Magnus, C., Johansen, W. T., Tappel, I. M., Gjeldvik, I., Solbakk, T., and Pham, V. T. H. (2013). CO<sub>2</sub> Storage Atlas of the Norwegian Part of the North Sea. *Energy Procedia*, 37, 4919–4926. doi:10.1016/j.egypro.2013.06.403
- Han, L., and Gu, Y. (2014). Optimization of Miscible CO<sub>2</sub> Water-Alternating-Gas Injection in the Bakken Formation. *Energy and Fuels*, 28(11), 6811–6819. doi:10.1021/ef501547x
- Haynes, S., and Alston, R. B. (1990, April 22). *Study of the Mechanisms of Carbon Dioxide Flooding and Applications to More Efficient EOR Projects*. SPE/DOE Enhanced Oil Recovery Symposium. doi:10.2118/20190-MS
- Holm, L. W., and O'Brien, L. J. (1971). Carbon Dioxide Test at the Mead-Strawn Field. *Journal of Petroleum Technology*, 23(04), 431–442. doi:10.2118/3103-PA
- Ismail, I., Mathiesen, V., Abraham, A., Ranjbar, E., Zarei, F., and Walter, J. (2021, January 22). *An Innovative Modelling Approach in Characterization of Autonomous Inflow Control Valve Performance to Maximizing Oil Recovery in Heavy Oil-SAGD Application*. doi:10.2118/203859-MS
- Khan, A., Saxena, S., Baloni, S., Sharma, M., and Kodavaty, J. (2021). Overview and methods in Enhanced Oil Recovery. *Journal of Physics: Conference Series*, 2070(1), 012061. doi:10.1088/1742-6596/2070/1/012061
- Khan, M. Y., and Mandal, A. (2022). Analytical model of incremental oil recovery as a function of WAG ratio and tapered WAG ratio benefits over uniform WAG ratio for heterogeneous reservoir. *Journal of Petroleum Science and Engineering*, 209, 109955. doi:10.1016/j.petrol.2021.109955
- Lei, H., Yang, S., Zu, L., Wang, Z., and Li, Y. (2016). Oil Recovery Performance and CO<sub>2</sub> Storage Potential of CO<sub>2</sub> Water-Alternating-Gas Injection after Continuous CO<sub>2</sub> Injection in a Multilayer Formation. *Energy and Fuels*, 30(11), 8922–8931. doi:10.1021/acs.energyfuels.6b01307
- Li, L., Khorsandi, S., Johns, R., and Dillmore, R. (2015). CO<sub>2</sub> enhanced oil recovery and storage using a gravity-enhanced process. *International Journal of Greenhouse Gas Control*, 42, 502–515. doi:10.1016/j.ijggc.2015.09.006
- Mahmood, M. N., Sultan, Z., and Yousuf, N. (2016, December 31). *A Review on Smart Well Completion System: Route to the Smartest Recovery*.
- Mansour, E. M., Al-Sabagh, A. M., Desouky, S. M., Zawawy, F. M., and Ramzi, M. (2019). A laboratory investigation of carbon dioxide-enhanced oil recovery by focusing on CO<sub>2</sub>-oil physical properties. *Egyptian Journal of Petroleum*, 28(1), 21–26. doi:10.1016/j.ejpe.2018.10.004
- Rogers, J. D., and Grigg, R. B. (2000, April 3). *A Literature Analysis of the WAG Injectivity Abnormalities in the CO<sub>2</sub> Process*. SPE/DOE Improved Oil Recovery Symposium. doi:10.2118/59329-MS
- Skauge, A., and Stensen, J. (2003, May 19). *Review of WAG Field Experience*.
- Supply – Key World Energy Statistics 2021 – Analysis. (n.d.). IEA. Retrieved January 18, 2024, from <https://www.iea.org/reports/key-world-energy-statistics-2021/supply>
- Taghavi, S., Aakre, H., and Moldestad, B. M. E. (2022, March 11). *Performance Analysis of Autonomous Inflow Control Valve in a SAGD Late Life Process with Non-Condensable Gases*. SPE Canadian Energy Technology Conference. doi:10.2118/208915-MS
- Taghavi, S., Tahami, S. A., Aakre, H., Furuvi, N. C. I., and Moldestad, B. M. E. (2023, October 9). *Performance Analysis of Autonomous Inflow Control Valve in a Heterogeneous Reservoir Using CO<sub>2</sub> Enhanced Oil Recovery*. SPE Annual Technical Conference and Exhibition. doi:10.2118/215153-MS
- Wang, J., Abiazie, J., McVay, D., and Ayers, W. (2008). Evaluation of Reservoir Connectivity and Development Recovery Strategies in Monument Butte Field, Utah. *Proceedings - SPE Annual Technical Conference and Exhibition*, 7. doi:10.2118/116695-MS
- Wang, Q., Yang, S., Lorinczi, P., Glover, P. W. J., and Lei, H. (2020). Experimental Investigation of Oil Recovery Performance and Permeability Damage in Multilayer Reservoirs after CO<sub>2</sub> and Water-Alternating-CO<sub>2</sub> (CO<sub>2</sub>-WAG) Flooding at Miscible Pressures. *Energy and Fuels*, 34(1), Article 1.
- Whorton, L. P., Brownscombe, E. R., and Dyes, A. B. (1952). *Method for producing oil by means of carbon dioxide* (Patent 2623596). <https://www.freepatentsonline.com/2623596.html>
- World Energy Outlook 2023 – Analysis. (n.d.). IEA. Retrieved January 18, 2024, from <https://www.iea.org/reports/world-energy-outlook-2023>
- Yan, Y., Li, C., Dong, Z., Fang, T., Sun, B., and Zhang, J. (2017). Enhanced oil recovery mechanism of CO<sub>2</sub> water-alternating-gas injection in silica nanochannel. *Fuel*, 190, 253–259. doi:10.1016/j.fuel.2016.11.019



# CO<sub>2</sub> Enhanced Oil Recovery in Reservoirs with Advanced Wells: Simulations and Sensitivity Analysis

Isu Uchechukwu Agha, Nora C.I. Furuvik, Soheila Taghavi \*

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway.*

*\*Corresponding author: [soheila.t.hosnaroudi@usn.no](mailto:soheila.t.hosnaroudi@usn.no)*

**Abstract:** Injection of CO<sub>2</sub> for enhanced oil recovery (CO<sub>2</sub>-EOR) is used in fields with high amount of residual oil. CO<sub>2</sub>-EOR refers to a technology where supercritical CO<sub>2</sub> is injected into an oil reservoir to increase the oil production. CO<sub>2</sub>-EOR in combination with CO<sub>2</sub>-storage is an attractive method to increase the oil production from mature oilfields, and at the same time reduce the carbon footprint from industrial sources. Utilizing autonomous inflow control valves (AICVs) in CO<sub>2</sub>-EOR projects contributes to a better distribution of CO<sub>2</sub> in the reservoir, reduction in production of water and CO<sub>2</sub> mixture, and thereby increased storage capacity of CO<sub>2</sub>. The main objective of this study is modelling and simulation of oil production from an oil reservoir using CO<sub>2</sub> water alternating gas (CO<sub>2</sub> WAG) injection in combination with advanced wells that are completed with AICVs. Furthermore, performance evaluation of the AICV technology and sensitivity analysis of parameters affecting the WAG process are completed. The results from the simulations indicate that well completion with AICV can maintain good oil production while the production of water is decreased from 3e+06 m<sup>3</sup> to 9.8e+04 m<sup>3</sup> which corresponds to 97% reduction in water production. The sensitivity analysis of the simulation results affirms that permeability, well placement, and well spacing have impact on productivity in terms of both oil recovery and water production in the WAG EOR method. The results indicate that permeability increase has a slight increment effect on oil recovery. The well spacing analysis shows that increasing the distance between the wells will increase the oil recovery and delay the water breakthrough. Lastly the well placement analysis shows that vertical injection of miscible CO<sub>2</sub> produces more oil than horizontal injection of miscible CO<sub>2</sub>. AICVs restrict the production of mixture of CO<sub>2</sub> and water, and thereby cause a better distribution of CO<sub>2</sub> in the reservoir.

**Keywords:** Miscible CO<sub>2</sub>, Enhanced Oil Recovery, Autonomous Inflow Control Valve, Water Alternating Gas, Computer Modelling Group, Minimum Miscible Pressure

## 1. INTRODUCTION

The oil and gas industry has played a pivotal role for the world energy production for decades. The oil and gas will remain important sources of energy in the future. Hence, improving oil recovery with reduced carbon footprint is necessary to meet the future energy demands. The CO<sub>2</sub> water alternating gas enhanced oil recovery (WAG EOR) is one of the methods used in the tertiary stage of oil production. WAG is a process of injecting CO<sub>2</sub> in alternating sequence with water into the oil field formation (Bahagio, 2013).

Studies suggest that the injection of CO<sub>2</sub> into the oil field reservoirs is beneficial for both the oil recovery and the greenhouse gas emissions (Safi et al., 2020). One example of the application of WAG EOR, is the commercial project at Lula offshore oil field, Brazil. Compared with CO<sub>2</sub>-EOR, the CO<sub>2</sub>-WAG EOR gives improved oil displacement and sweep efficiencies (Bahagio, 2013). Norway has technical potential for CO<sub>2</sub>-WAG EOR on the North Sea oil fields. However, one problem is that the CO<sub>2</sub> injected can be recirculated into the producer well leading to poor distribution of CO<sub>2</sub> in the reservoir and thereby damage the process equipment due to the corrosive mixture of CO<sub>2</sub> and water (E. K. Halland et al., 2019). Advanced wells or smart wells are used to avoid the problems with recirculation of CO<sub>2</sub>, thus forcing CO<sub>2</sub> to distribute over a larger area in the reservoir. Examples of advanced well completion technologies are the autonomous inflow control valve (AICV) developed by InflowControl AS

and the passive inflow control device (ICD) (Aakre et al., 2018). Restricting CO<sub>2</sub> recirculated using AICV may potentially lead to higher drawdown in high-oil saturation zones. There is also a broader contact between CO<sub>2</sub> and the residual oil in the reservoir, all of which will boost oil production and recovery. CO<sub>2</sub>-WAG can be either miscible or immiscible depending on the minimum miscibility pressure, however this study will solely investigate the miscible process. The producer and injector wells can either be vertical or horizontal. The CO<sub>2</sub>-WAG performance depends on well spacing, well placing, CO<sub>2</sub> and water injection rates, permeability, and porosity differences in the reservoir (Taghavi et al., 2023).

This study aims at modelling and simulation of enhanced oil recovery for miscible CO<sub>2</sub> injection with advanced wells completed with AICV. Further performance evaluation of the AICV technology and sensitivity analysis of parameters affecting the WAG process are completed. The Miscible CO<sub>2</sub>-WAG with advanced wells model was developed using the commercial software Computer Modelling Group (CMG). In this study, different available modules such as Builder, FlexWell, and STARS are used to achieve the modelling and simulations. The collected data from different simulation cases are used to perform sensitivity analysis on parameters that impact the EOR process.

## 2. CO<sub>2</sub> EOR

### 2.1 EOR method of CO<sub>2</sub> water alternating gas (WAG)

CO<sub>2</sub>WAG is an improvement of the gas injection methods. CO<sub>2</sub>, when dissolved in oil, reduces the oil viscosity which helps to increase the mobility of the oil hence improving the oil recovery. The CO<sub>2</sub> injection alone often results in low sweep efficiency because of unstable displacement due to gravity segregation and viscous fingering caused by early gas breakthrough. (Cherian et al., 2012)

Figure 1 illustrates the principle of the CO<sub>2</sub>-WAG EOR process and shows how the miscibility between CO<sub>2</sub> and oil happens in the miscible zones after flooding.

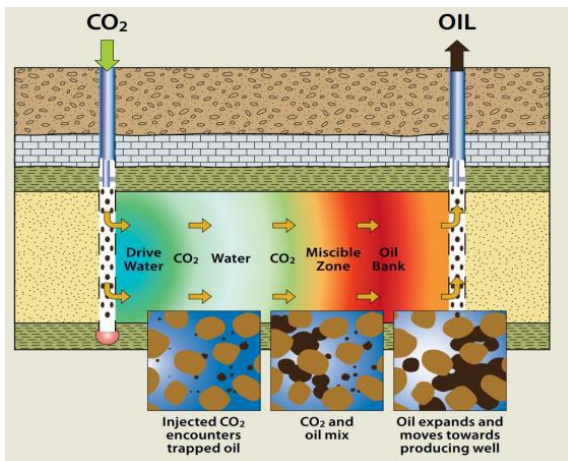


Fig. 1. The diagram of CO<sub>2</sub> WAG process (E. Halland et al., 2012).

When CO<sub>2</sub> EOR takes place at a pressure equal to or higher than the minimum miscibility pressure (MMP) it is called miscible CO<sub>2</sub> EOR, while CO<sub>2</sub> EOR at pressures lower than MMP is called immiscible CO<sub>2</sub> EOR. The advantage of the miscible CO<sub>2</sub> EOR process is that the oil volume is increased, and the oil viscosity is lowered causing oil to travel easier towards the producing wells (Chathurangani and Halvorsen, 2015). The MMP is the reservoir pressure above which CO<sub>2</sub> and oil can combine into a single-phase fluid.

CO<sub>2</sub>-WAG can help to control the mobility of the gas because the water will limit fractional flow of gas which will lead to improved sweep efficiency as well as displacement efficiency. The parameters which can affect the result of CO<sub>2</sub>-WAG are injection rates and WAG cycle length for each injection phases (Bahagio, 2013).

A problem with CO<sub>2</sub>-WAG is that CO<sub>2</sub> dissolved in water can form corrosive acid with calcite component presence in the rock (Oomole and Osoba, 1983):



This phenomenon can lead to economic challenges after breakthrough if this corrosive mixture reaches the producer wells.

### 2.2 Advanced wells and their impact on increased EOR

Advanced well completion might be necessary in maximizing the efficiency of the EOR process in order to avoid the common challenge of early CO<sub>2</sub> and water breakthrough. Presently in the oil and gas industry advanced wells can be achieved with flow control devices, annular flow isolation, and sand control screens (Moradi et al., 2022).

The ICD (Fig. 2) is an example of a passive flow control device with no moving parts inside. ICD was innovated to solve the phenomena of the heel-to-toe effect along the well because it can provide additional pressure drop, and by that balance the pressure variation from the toe to the heel along the well. The installation of ICD in the wells can delay gas and water breakthrough in an EOR process, but it cannot restrict the flow of unwanted effluents once a breakthrough of these fluids occurs (Kais et al., 2016).



Fig. 2. The picture of the nozzle type ICD technology (Kais et al., 2016).

The AICV (Fig. 3) is an example of a reactive flow control device. The AICV responds with a contrary course of action without direct human control when present in the well. It is a modern technology with a movable piston which acts after water breakthrough in EOR. The operating procedure of the AICV device is governed by viscosity and density differences which determines the pressure drop for different reservoir fluids (Aakre et al., 2018).



Fig. 3. The picture of the modern AICV technology (InflowControl, 2024).

If high viscous fluids like oil is around the valve, the piston acts downwards which opens the valve. If low viscous fluids like CO<sub>2</sub> or water is around the valve, the piston acts upwards which closes the valve.

Taghavi et al. (2023) compared the ICD with the AICV performance. With both devices having the same oil flow rate at a specific differential pressure, the results from the study showed that there is a significant gas and water reduction by using AICV under the same conditions, see Fig. 4.

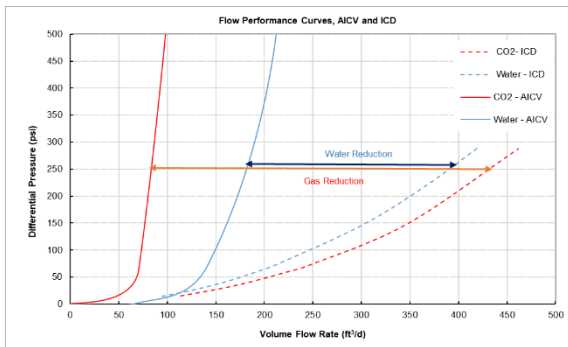


Fig. 4. The performance curves of pressure and volumetric flow rate for AICV and ICD (Taghavi et al., 2023).

### 3. MODEL DEVELOPMENT ON CMG

#### 3.1 Reservoir fluid components and characterization

The WinProp package is capable of fluid characterization, matching experimental data, and constructing phase diagrams. WinProp uses equation of states such as Peng-Robinson combined with data obtained from laboratory analysis of reservoir samples. However, the aim of the created fluid model in this work is to calculate the MMP required to achieve miscibility between oil and CO<sub>2</sub> injected. It was determined to be 15284 kPa at reservoir temperature of 85.5°C. Figure 5 shows the pressure-temperature phase envelope of CO<sub>2</sub> generated by WinProp. The two-phase boundary is the green curve, and the critical temperature and pressure are approximately 6500kPa and 425°C.

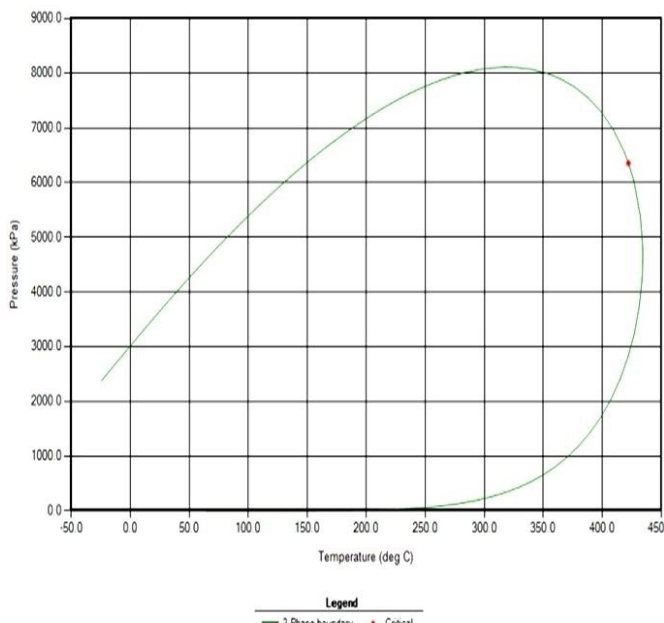


Fig. 5. The P-T phase envelope of CO<sub>2</sub> created in WinProp.

#### 3.2 The reservoir

One homogeneous reservoir and one heterogeneous reservoir were built in builder suite package with cartesian plane. For both reservoirs, there are ten grids in the I-J direction, and fifteen in the K direction. The length, width and height dimensions of the reservoir are 300 m, 500 m, and 150 m, respectively. The top of the reservoir is at a depth of 1000 m and the bottom of the reservoir is at depth 1150 m. Most

properties of the reservoirs were left in the original preset initial values specified by CMG, however both reservoirs porosity was modified to 0.35. The initial reservoir temperature is constant at 85.5°C. The reference pressure is 20684.3 kPa, which is much higher than the MMP, to ensure the process remains a miscible CO<sub>2</sub> process. The surface pressure condition was 101 kPa and the surface temperature condition was 16.85°C.

Figure 6 shows the pictorial view of the homogeneous reservoir (left-hand-side) and the heterogeneous reservoir (right-hand-side). The homogeneous reservoir permeability is constant all through the layers at 2500 mD. The heterogeneous reservoir permeability varies from 2500 mD (blue color) to 10000 mD (red color). The highest permeability region for the heterogeneous reservoir was placed at the heel section of the producer wells.

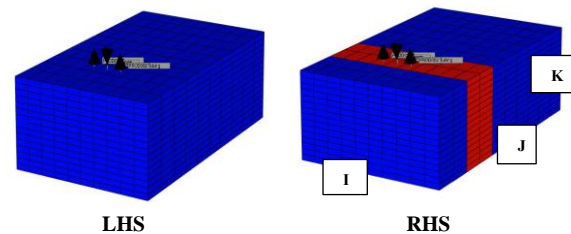


Fig. 6. The 3-D view of the homogeneous (LHS) and heterogeneous reservoir (RHS).

The wettability state of the rock is water wet. The relative permeability curves datasets were calculated based on the Stone II model for two-phase. The oil is immobile below 0.25 saturation, and the water maximum saturation is 0.78.

#### 3.3 The simulation cases

The developed simulation cases were based on the homogeneous reservoir labelled Case-A and the heterogeneous reservoir labelled Case-B. The simulation cases were investigated with different injection methods (water-EOR and WAG-EOR), different wellbore placement (horizontal and vertical wells), and different wellbore completion (with AICV and without AICV). Additional simulations were performed in order to investigate the effect of parameters such as well spacing and permeability.

The timeline of the simulated cases was for 10 years from the period of 2024-01-01 to 2034-01-01. The base case is defined as the water injection mode with two producer wells open for continuous production all year, and one injector well perforated in the middle between the two producer wells. The injector well is open all year during these periods to inject water into the reservoirs. The WAG-EOR case involves the same wells and the same perforation location as the base case, but the injection cycle period was modified to injection period for both water and CO<sub>2</sub>. Figure 7 shows an illustration of the timeline for the water injected alternately with CO<sub>2</sub> with all year continuous production. The annulus of the producer well is shut, but the tubing is open.

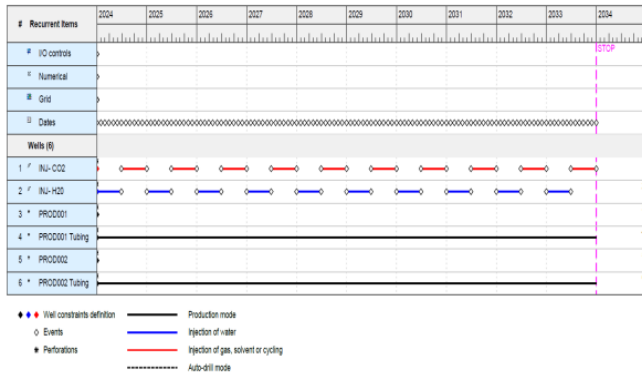


Fig. 7. The timeline of the WAG cycle periods.

The wall inner and outer diameters were modified to 0.3 m and 0.35 m respectively. Table 1 shows the type of constraint and the specified values for the simulations. Both the injector well, the producer well-1 and the producer well-2 which are installed in the horizontal and vertical perforations have the same constraints values specified, except STL surface liquid rate.

**Table 1. The constraint specification for both the injector well and producer wells**

Constraint Type	Limit	Value
BHP bottom hole pressure	MAX	22000 kPa
BHP bottom hole pressure	MIN	15000 kPa
STG surface gas rate	MAX	50000 m3/day
STW surface water rate	MAX	10000 m3/day
STL surface liquid rate	MAX	840 m3/day

Figure 8 shows the picture of the horizontally placement producer well-1 from the J-k direction view.

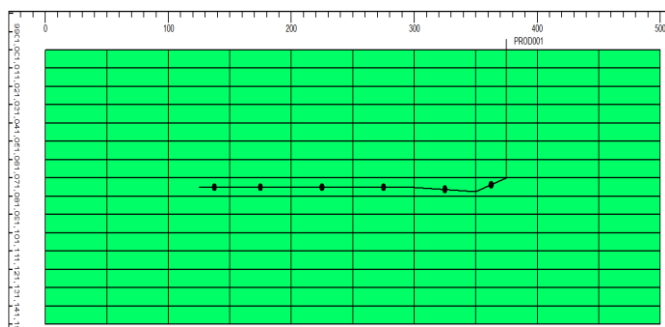


Fig. 8. The J-K direction view of the horizontal producer well-1.

Figure 9 shows the picture of the vertically placement wells from the I-k direction view of the producer well-1 to the left, the injector well in the middle, and the producer well-2 to the right.

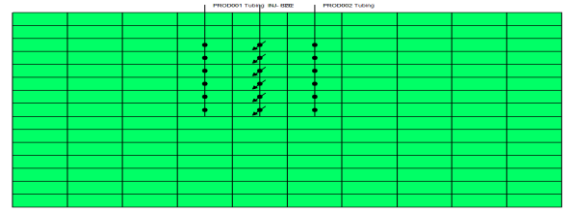


Fig. 9. The I-K direction view of the vertical placement of the injector and producer wells.

## 4. RESULTS AND DISCUSSION

### 4.1 Comparison of WAG and water injection

Figure 10 shows the field oil rate of the two producer wells at standard condition. The thick green line represents the oil rate for water injection, and the dash green line represents Case-A-1 which is the WAG.

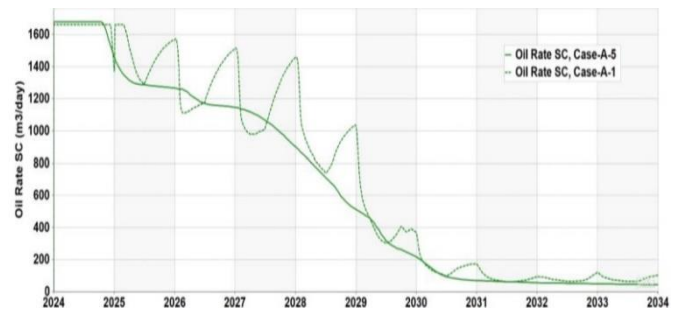


Fig. 10. The field oil rate of the two producer wells at standard condition for case-A-1 and case-A-5.

The field oil rate in the figure illustrates that both WAG and water injection can promote oil productivity. The case-A-1 has an oscillating curve because the highest peaks in the oil production appear during the CO<sub>2</sub> injection period. This is because if CO<sub>2</sub> is well circulated around the reservoir region of high oil saturation, the mobility of the oil toward the producer wells increases.

In year 2034, the cumulative oil production for the WAG is 2.7e+06 m<sup>3</sup>, represented with the thin green line in Fig. 11. This is approximately 12.5 % more oil than the cumulative oil production for water injection (Thick green line in Fig. 11), of which the oil cumulative production is 2.4e+06 m<sup>3</sup>.

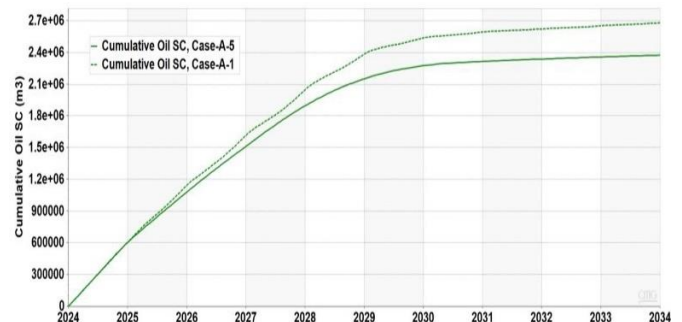


Fig. 11. The field cumulative oil of the two producer wells at standard condition for case-A-1 and case-A-5.



Oil saturation is an important parameter to observe when comparing WAG and water injection. Figure 12 indicates that WAG produces more oil than water injection. In Fig. 12 the green color zones in the reservoir are where oil has been produced and replaced with water. The scaling shows that the red color zone is the high oil saturation zone and has a mole fraction of 1 which means oil is the only component present. The orange and yellow color indicates two-phase zone of oil and water.

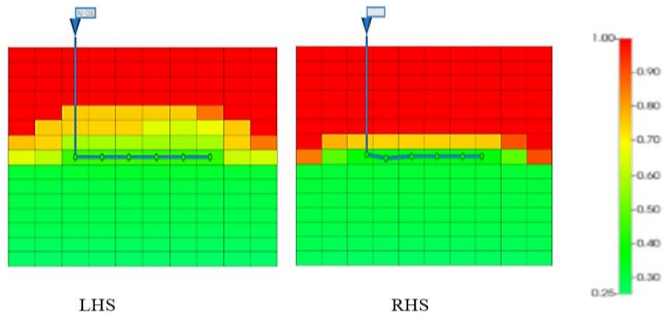


Fig. 12. The reservoir oil saturation for the WAG (LHS) and the water injection (RHS).

4.2 The performance of AICV

Figure 13 shows the graph of the field cumulative oil of the two producer wells at standard condition. The solid green line represents Case-A-1 i.e. with AICV completion, and the dash green line represents Case-A-2 which is without AICV completion. From the figure it is seen that the cumulative oil production without AICV corresponds to around  $3.3e+06 \text{ m}^3$  for the entire production period, while the cumulative oil production with AICV corresponds to  $2.7e+06 \text{ m}^3$ .

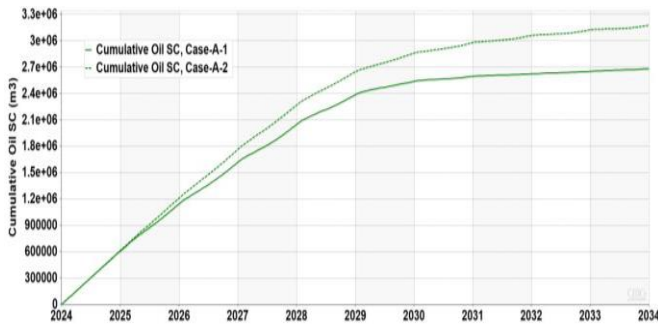


Fig. 13. The field cumulative oil of the two producer wells at standard condition for case-A-1 and case-A-2.

As water can form corrosive mixture with  $\text{CO}_2$  (Oomole and Osoba, 1983), the goal is to produce as little water as possible. This is, among other things, to prevent the corrosive mixture from entering the top side facilities where it can cause major damage. Figure 14 compares the cumulative water production with and without AICV completion in the wells. The figure shows that by installing AICVs in the well, the cumulative water production can be reduced from  $3e+06 \text{ m}^3$  to  $9.8e+04 \text{ m}^3$  during the production period, which corresponds to approximately 97% less water production with AICV.

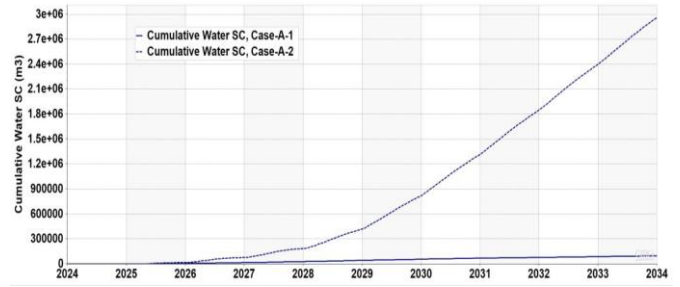


Fig. 14. The field cumulative water of the two producer wells at standard condition for case-A-1 and case-A-2.

4.2 Comparison of homogeneous and heterogeneous reservoir

Figure 15 and Figure 16 illustrate that permeability of the reservoir plays an important role in oil and water production. In Fig. 15 the heterogeneous reservoir is represented by the thin green line, while the homogenous reservoir is represented by the thick green line. The heterogenous reservoir, which has a higher permeability (10000 mD) around the heel section of the well has a slightly higher cumulative oil production in the year 2034 compared to the homogenous reservoir, of which the oil production has increased from  $2.65e+06 \text{ m}^3$  to  $2.7e+06 \text{ m}^3$  respectively.

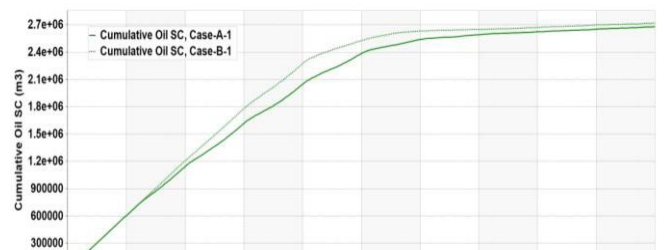


Fig. 15. The field cumulative oil at standard condition for case-A-1 (homogeneous) and case-B-1(heterogeneous).

However, the heterogeneous reservoir produces considerably less water than the homogenous reservoir. During the entire production period the heterogenous reservoir produces  $5.2e+04 \text{ m}^3$ , which is around 50% reduction compared to the homogeneous reservoir that produces  $9.8e+04 \text{ m}^3$ .

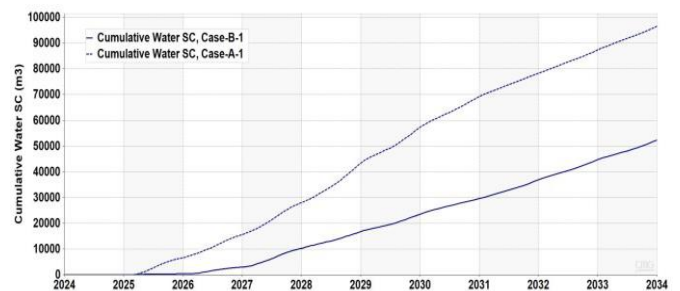


Fig. 16. The field cumulative water at standard condition for case A-1 (homogeneous) and case-B-1(heterogeneous).



4.2 The impact of well spacing and position on production

To investigate the impact of well spacing and position on the oil and water production, the producer wells were modified to have a shorter distance from the injector well, as shown in Fig. 17.

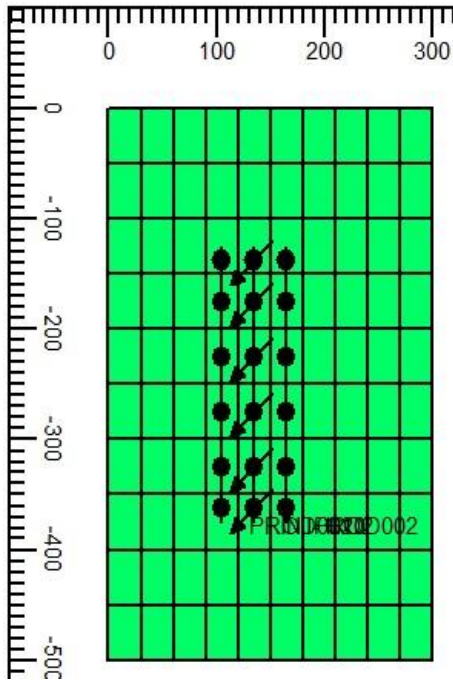


Fig. 17. The modified case-A-2 with less well spacing distances.

It is important to avoid that the producer wells are too close to the injector well, this is because an early water breakthrough at the start of the WAG is observed. Possibly because the injected fluid (water) at the start of every year is produced directly in the producer well instead of being distributed in the reservoir. This effect reduces the oil rate, as shown in Fig. 18 where the production drops towards zero at the start of the years which corresponds to the periods when water was injected.

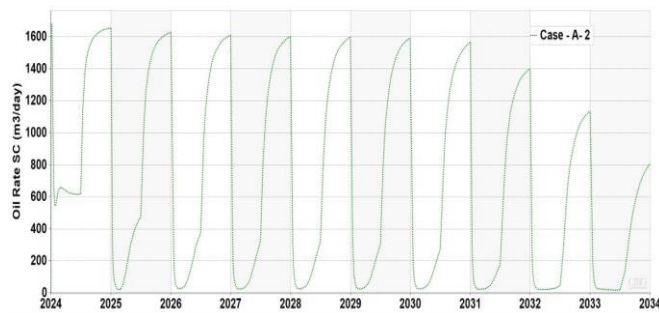


Fig. 18. The field oil rate for the modified case-A-2 with less well spacing distance.

At year 2034, the cumulative water production has increased from  $3e+06$  m<sup>3</sup> (Fig. 14) to  $3.5e+06$  m<sup>3</sup>, this is because of early breakthrough at the beginning of year 2024, see Fig. 19.

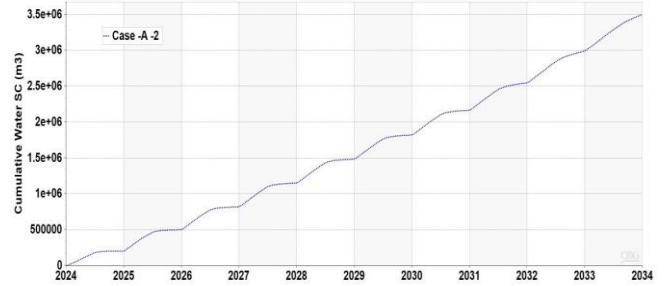


Fig. 19. The field cumulative water volume for the modified case-A-2 with less well spacing distance.

4.5 Comparison of horizontal and vertical wells

The impact of well placement on miscible CO<sub>2</sub> injection was investigated by comparing a vertical injector well with vertical producer wells, to a horizontal injector well with horizontal producer wells. This comparison was done for cases with AICV and cases without AICV.

The vertical injection (thick green line) of miscible CO<sub>2</sub> injection gives a higher cumulative oil production than the horizontal injection (thin green line), producing  $3.3e+06$  m<sup>3</sup> and  $3e+06$  m<sup>3</sup> respectively, see Fig. 20.

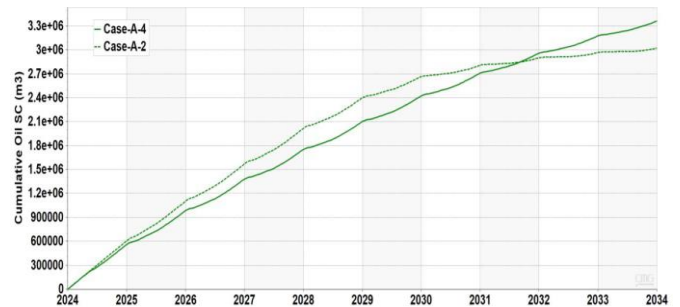


Fig. 20. Cumulative oil production without AICV for the horizontal case (A-2) and the vertical case (A-4).

The vertical injection (thick blue line) of miscible CO<sub>2</sub> injection gives less cumulative water production than the horizontal injection (thin blue line), producing around  $2.7e+06$  m<sup>3</sup> and  $3e+06$  m<sup>3</sup> respectively, see Fig. 21.

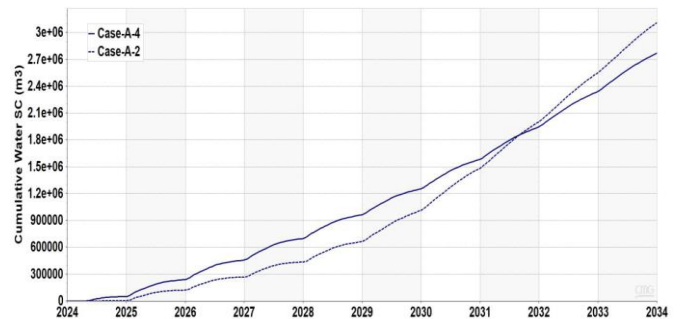


Fig. 21. Cumulative water production without AICV for the horizontal case (A-2) and the vertical case (A-4).

Another important observation is the oil saturation from the injector for both vertical and horizontal well with time. Figure

22 illustrates the sweeping of oil because of the miscible CO<sub>2</sub> injection at the end of the production period (year 2034). The figure shows that at the end of the production period at year 2034, the vertical miscible CO<sub>2</sub> injection has better sweep of the oil resulting in less oil saturation (light green) compared to horizontal miscible CO<sub>2</sub> injection.

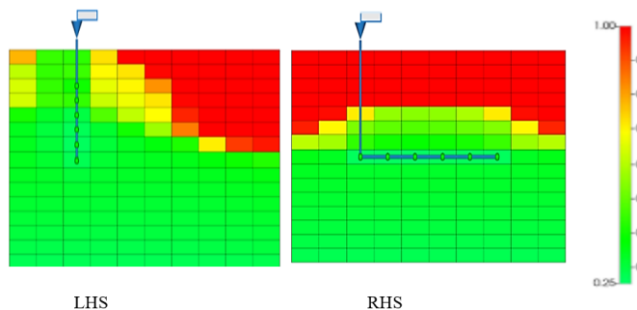


Fig. 22. The oil saturation at year 2034 of case-A-4 and case-A-1 for the vertical (LHS) and horizontal (RHS) CO<sub>2</sub> injection.

## 6. CONCLUSIONS

The objective of this study was modelling of the miscible CO<sub>2</sub> injection for WAG process and evaluation of the performance of the AICV including sensitivity analysis of the parameters affecting the EOR process.

The results show that the production wells completed with AICVs maintain good oil production while the production of water is decreased from  $3e+06$  m<sup>3</sup> to  $9.8e+04$  m<sup>3</sup> which corresponds to 97% reduction in water production.

The sensitivity analysis of the simulation results affirms that permeability, well placement, and well spacing have impact on productivity in terms of both oil recovery and water production in the WAG EOR method. The results indicate that permeability increase has a slight increment effect on oil recovery and 50% decrease in water production. The well spacing analysis shows that increasing the distance between the wells will increase the oil recovery and delay the water breakthrough. Also, if the wells are too close, recirculation of injected water and CO<sub>2</sub> in the producer wells occurs at the start date. Lastly the well placement analysis shows that vertical injection of miscible CO<sub>2</sub> produces more oil than horizontal injection of miscible CO<sub>2</sub>.

As future study, it is recommended to investigate the optimum perforation location and distance for the CO<sub>2</sub> injector from the producer well which favors maximum oil recovery, reduced operational cost and economic challenges.

## REFERENCES

Aakre, H., Mathiesen, V., and Moldestad, B. (2018). Performance of CO<sub>2</sub> flooding in a heterogeneous oil reservoir using autonomous inflow control. *Journal of Petroleum Science and Engineering*, 167, 654–663. [doi.org/10.1016/j.petrol.2018.04.008](https://doi.org/10.1016/j.petrol.2018.04.008)

Bahagio, D. N. T. (2013). *AES/PE/13-12 Ensemble Optimization of CO<sub>2</sub> WAG EOR*.

Chathurangani, L. B. J., and Halvorsen, B. M. (2015). *Near Well Simulation of CO<sub>2</sub> Injection for Enhanced Oil Recovery (EOR)*. 309–318. [doi:10.3384/ecp15119309](https://doi.org/10.3384/ecp15119309)

Cherian, B. V., Stacey, E. S., Bressler, S., Iwere, F. O., Heim, R. N., and Higgins-Borchardt, S. (2012). *Evaluating Horizontal Well Completion Effectiveness in a Field Development Program*. SPE Hydraulic Fracturing Technology Conference. [doi:10.2118/152177-MS](https://doi.org/10.2118/152177-MS)

Halland, E. K., Pham, V., Riis, F., and Hansen, A.-H. (2019). CO<sub>2</sub> for EOR Combined with Storage in the Norwegian North Sea. *SSRN Electronic Journal*. [doi:10.2139/ssrn.3365602](https://doi.org/10.2139/ssrn.3365602)

Halland, E., Riis, F., Johansen, W. T., Gjeldvik, I. T., and Magnus, C. (2012). *CO<sub>2</sub> Storage Atlas of the Norwegian Part of the North Sea*. cp. [doi:10.3997/2214-4609.20148261](https://doi.org/10.3997/2214-4609.20148261)

Kais, R., Mathiesen, V., Aakre, H., Woiceshyn, G., Elarabi, A., and Hernandez, R. (2016). *First Autonomous Inflow Control Valve AICV Well Completion Deployed in a Field Under an EOR Water & CO<sub>2</sub> Injection Scheme*. SPE Annual Technical Conference and Exhibition. [doi:10.2118/181552-MS](https://doi.org/10.2118/181552-MS)

Moradi, A., Samani, N. A., Kumara, A. S., and Moldestad, B. M. E. (2022). Evaluating the performance of advanced wells in heavy oil reservoirs under uncertainty in permeability parameters. *Energy Reports*, 8, 8605–8617. [doi:10.1016/j.egy.2022.06.077](https://doi.org/10.1016/j.egy.2022.06.077)

Oomole, O., and Osoba, J. S. (1983, May 9). *Carbon Dioxide—Dolomite Rock Interaction During CO<sub>2</sub> Flooding Process*. Annual Technical Meeting. [doi:10.2118/83-34-17](https://doi.org/10.2118/83-34-17)

*Planning and implementing an EOR project for the pre-salt Lula field*. (2012, September 8).

<https://www.worldoil.com/magazine/2012/august-2012/supplement/planning-and-implementing-an-eor-project-for-the-pre-salt-lula-field>

*Product Range | InflowControl*. (n.d.). Retrieved 4 May 2024, from <https://www.inflowcontrol.no/reservoir-solutions/products->

Safi, A., Chen, Y., Wahab, S., Ali, S., Yi, X., and Imran, M. (2020). Financial Instability and Consumption-based Carbon Emission in E-7 Countries: The Role of

Trade and Economic Growth. *Sustainable Production and Consumption*, 27, 383–391.  
[doi:10.1016/j.spc.2020.10.034](https://doi.org/10.1016/j.spc.2020.10.034)

Taghavi, S., Tahami, S. A., Aakre, H., Furuvik, N. C. I., and Moldestad, B. M. E. (2023). Performance Analysis of Autonomous Inflow Control Valve in a Heterogenous Reservoir Using CO2 Enhanced Oil Recovery. *Day 3 Wed, October 18, 2023*, D031S045R002. [doi:10.2118/215153-MS](https://doi.org/10.2118/215153-MS)

## CO<sub>2</sub> storage and evaluation of important parameters affecting the CO<sub>2</sub> plume distribution: simulation and sensitivity analysis

Mohammad Rakibul Hasan Chowdhury and Soheila Taghavi

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway.*

E-mail corresponding author: [soheila.t.hosnaroudi@usn.no](mailto:soheila.t.hosnaroudi@usn.no)

**Abstract:** Carbon capture utilization and storage (CCUS) offers a potential solution to mitigate the effects of anthropogenic CO<sub>2</sub> and to reduce the direct CO<sub>2</sub> emissions from stationary sources into the atmosphere. The captured CO<sub>2</sub> is injected into deep saline-water saturated formations or in depleted oil and gas fields, or into the oil fields for storage and/or enhanced oil recovery (EOR). The primary objective of this study is to identify and analyze the critical parameters affecting CO<sub>2</sub> plume development in the reservoir. Understanding the subsurface dynamics of carbon sequestration will facilitate to plan the subsurface process better. The simulation models are developed using the commercial software Computer Modelling Group, CMG. The plume dynamics that include plume volume and plume geometry over 30 years of injection and 170 years of post-injection period is investigated. Additionally, the contribution of different trapping mechanisms over the time horizon in the storage process is assessed. Moreover, a sensitivity analysis is done for evaluating the impact of variables including porosity, permeability, injection rate, and injector bottom hole pressure. The simulation results show that CO<sub>2</sub> plume propagates at an increased rate during the injection period and continues to disperse at a comparatively reduced rate after the injection ends. The horizontal spread of plume is significantly greater than the vertical propagation when the horizontal permeability is larger than the vertical. Additionally, the plume volume shows a linear relationship with the injected CO<sub>2</sub> amount. In terms of storage efficiency, the most prevalent CO<sub>2</sub> is free phase super critical CO<sub>2</sub> that contributes around 80% of the stored CO<sub>2</sub> whereas the rest are structurally or residually trapped and dissolved CO<sub>2</sub>. From the sensitivity analysis in a homogenous reservoir, it can be concluded that the horizontal permeability is impacting the most (42%) for structural and residual trapping of CO<sub>2</sub> whereas porosity impacts the most (38%) for dissolution of CO<sub>2</sub> contributing to solubility trapping mechanism.

*Keywords:* CCUS, Plume dynamics, CO<sub>2</sub> Storage, CO<sub>2</sub> trapping mechanism, Sensitivity

### 1. INTRODUCTION

Carbon dioxide (CO<sub>2</sub>) reduction from the atmosphere has become a global attention from last decade which resulted in imposing several laws against emitting CO<sub>2</sub> in the atmosphere. However, CO<sub>2</sub> production is inevitable in lot of processes which compelled the industries and researchers to draw more attention in capturing the produced CO<sub>2</sub> and storing them to a safe place. In the storage purpose, depleted oil reservoirs or aquifers has a good potential to be used for storage of CO<sub>2</sub>. According to a study, there is approximately 139 giga tones of CO<sub>2</sub> storage potential in worldwide oil reservoirs (Godec et al., 2011). Geological sequestration of CO<sub>2</sub> is presently the most viable, and probably the sole, short-to-medium term strategy for substantially increasing CO<sub>2</sub> sinks and thereby decreasing overall carbon emissions into the atmosphere (Bachu, 2008). So, CO<sub>2</sub> storage in oil reservoirs, coupled with enhanced oil recovery (EOR), has gained attention as a strategy for both mitigating climate change and improving oil recovery. Currently, CO<sub>2</sub> storage in geological formations such as oil and gas reservoirs and deep saline aquifers is not a new technology. The extensive history of natural gas storage in North America and Europe and CO<sub>2</sub>-EOR practices primarily in the U.S (Moritis, 2006) provide some evidences of this fact.

Additionally, numerous commercial-scale projects worldwide engage in CO<sub>2</sub> injection for various purposes, ultimately contributing to the mitigation of greenhouse gas emissions, such as, in Canada, the Weyburn CO<sub>2</sub>-EOR project uses CO<sub>2</sub> obtained from coal gasification (Whittaker et al., 2004). In the North Sea, the Sleipner project injects CO<sub>2</sub> stripped from natural gas into the Utsira formation (Torp and Gale, 2003). Over in Algeria, the In Salah project pumps CO<sub>2</sub> back into an aquifer underneath the gas field it came from (Riddiford et al., 2003) and some examples in Western Canada, show that CO<sub>2</sub> storage is often implemented for economic or regulatory reasons, not just for addressing climate change (Bachu, 2008) but resulted in the mitigation of diverse impact on climate change.

However, the success of CO<sub>2</sub> storage is highly dependent on the behavior of the CO<sub>2</sub> plume within the reservoir, which is influenced by reservoir characteristics (Birkholzer et al., 2015). The study by Luo et. al. (2022) suggests that studies on structural sequestration should take into account all relevant factors, and that the capacity of structural sequestration should be assessed in light of the characteristics of the caprock, the rate of CO<sub>2</sub> injection, and the saline aquifer actual geological

conditions. Current research in the field of CO<sub>2</sub> plume evolution and the impact of reservoir parameters on storage efficiency emphasizes the complexity and variability of geological storage environments. Al-Khdheawi et al. (2018) discusses the effects of heterogeneity, reservoir temperature, and wettability on CO<sub>2</sub> migration and trapping mechanisms. Myshakin et al. (2023) explores the impact of reservoir heterogeneity on fluid displacement and volumetric efficiency. But these studies underscore the critical role of physical reservoir characteristics in determining CO<sub>2</sub> storage efficiency. Moreover, Zapata et al. (2020) and Luo et al. (2022) contribute to this topic by investigating CO<sub>2</sub> plume dynamics over long-term injection periods and the effects of key parameters on gas recovery and storage efficiency, respectively. However, there remains a significant gap in comprehensive analyses integrating multiple reservoir parameters under varied operational conditions, particularly understanding the simultaneous effects of the reservoir parameters and dynamic injection scenarios on CO<sub>2</sub> storage efficiency.

The primary objective of this work is to observe plume evolution and analyze the reservoir parameters affecting CO<sub>2</sub> plume, and storage efficiency in the reservoir. These parameters include, reservoir permeability, porosity, pressure and temperature conditions, and fluid properties as well as the injection rate and injection bottom hole pressure. CO<sub>2</sub> storage model is developed in the commercial software Computer Modelling Group, CMG to investigate the CO<sub>2</sub> migration after a 30-year injection period and 170-year post-injection period. Sensitivity analysis is done among corresponding variables to understand the impact of different parameters on storage capacity. The simulation models are developed using CMG, by utilizing its five modules.

## 2. CO<sub>2</sub> STORAGE IN GEOLOGICAL MEDIA

Under normal atmospheric conditions, CO<sub>2</sub> exists as gas. However, when subjected to pressures exceeding 7.39 Mpa and temperatures above 31.1°C, reaching what is termed its critical point, it transitions into a supercritical fluid state (Zhi et al., 2019). In this form, CO<sub>2</sub> exhibits properties of both gases and liquids, making it ideal for underground storage. It becomes as dense as a liquid, which allows it to hold more CO<sub>2</sub> in the pore spaces of rock formations. At the same time, it maintains a gas-like viscosity, facilitating its movement through the rock layers. Reservoirs suitable for CO<sub>2</sub> storage are typically found deeper than 1 km, have a thickness of around 10500 m, and may extend for hundreds of kilometers across (Szulczewski, 2013). At these depths, CO<sub>2</sub> is kept in its supercritical condition, where it is somewhat lighter (~700 kg/m<sup>3</sup>) than the brine, oil, or any other existing fluids, leading it to migrate upwards due to buoyant forces (Verma et al., 2021). The upward movement of CO<sub>2</sub> ceases when it meets the caprock, effectively sealing it within the subsurface.

CO<sub>2</sub> has an influence on the relative permeability curve. CO<sub>2</sub> injection influences the wettability of the reservoir, causing the rock to become somewhat more water-wet, which promotes better oil displacement efficiency and optimizes the reservoir capacity to trap CO<sub>2</sub> effectively (Kułynycz, 2015). This process alters the endpoint relative permeabilities as well as

modifying the shape of the corresponding relative permeability curves (Taghavi et al., 2023).

This indicates that CO<sub>2</sub> injection reduces the mobility ratio,  $M$ , between oil and water. Mobility,  $\lambda$ , is characterized as the ratio of the endpoint relative permeability,  $K_r$ , to dynamic viscosity,  $\mu$ .

$$\lambda = \frac{K_r}{\mu} \quad (1)$$

As defined by Ahmed (2010), the mobility ratio is the ratio of the mobility of the fluid causing displacement, such as  $\lambda_w$  to the mobility of the fluid being displaced, such as  $\lambda_o$  where the subscripts denote the water for w and oil for o.

$$M = \frac{\lambda_w}{\lambda_o} = \frac{k_{rw}}{k_{ro}} \cdot \frac{\mu_o}{\mu_w} \quad (2)$$

Enhancing the water-wetness of the rock decreases the residual oil saturation while increasing the irreducible water saturation (Taghavi et al., 2023). Consequently, the oil relative permeability is increased. Moreover, there is a decrease in oil viscosity, collectively leading to a lower mobility ratio (Aakre et al., 2018).

### 2.1 Geological media

Potential CO<sub>2</sub> storage options include deep saline aquifers, operational or depleted oil and gas fields, unmineable deep coal beds, and mined salt caverns. According to Saeedi (2012), deep saline aquifers have the advantage of extensive capacity and wide availability but face the disadvantage of unproven storage reliability. Active or depleted oil and gas reservoirs offer demonstrated storage security, established infrastructure, and enhanced hydrocarbon recovery, yet they are not available in all regions and may not be available for immediate use (Saeedi, 2012). Unminable coal seams can potentially enhance methane recovery but are limited by uncertain storage capacity and regional availability. Basalt formations provide a permanent trapping mechanism for CO<sub>2</sub>, although they are constrained by slow reaction rates and limited field experiences (Saeedi, 2012). CO<sub>2</sub> is stored safely without risking the contamination of underground resources. Historically, they have effectively contained oil and gas under high pressure and temperature, thereby reducing the likelihood of CO<sub>2</sub> leakage over extended periods. Furthermore, these abandoned hydrocarbon storage sites can maintain the necessary temperature and pressure for CO<sub>2</sub> to reach supercritical condition (Van Der Meer, 2005).

### 2.2 CO<sub>2</sub> Trapping Mechanism

At least six mechanisms exist that can secure CO<sub>2</sub> within a storage complex over extended duration. Among the most recognized are structural trapping, capillary trapping, solubility trapping, and mineral trapping. Table 1 presents a comparative analysis of the trapping mechanisms.



**Table 1. Comparison of different trapping mechanism (Ketzer et al., 2012; Raza et al., 2015)**

Trapping Mechanism	Description	Advantages	Considerations
Structural and Stratigraphic Trapping	CO <sub>2</sub> is trapped beneath an impermeable cap rock, similar to how oil and gas are trapped in petroleum fields.	-Direct and immediate trapping - Utilizes existing geological structures	- Dependent on the integrity of the cap rock - Limited by the structure's capacity and closure.
Capillary Trapping	CO <sub>2</sub> becomes immobilized as a residual phase within the pore spaces of the storage medium.	- Rapid and efficient entrapment of CO <sub>2</sub> - Enhanced storage security through immobilization. - Higher capillary force than buoyant force, leading to pore-scale CO <sub>2</sub> bubbles	- Efficiency varies with rock properties and fluid characteristics - Requires sophisticated understanding of pore space interactions
Solubility Trapping	CO <sub>2</sub> dissolves in brine, and the denser CO <sub>2</sub> -saturated brine sinks within the storage medium.	- Contributes to long-term storage stability - Reduces risk of leakage by dissolving CO <sub>2</sub>	- Dependent on diffusion rates and storage medium properties - Slower process compared to capillary trapping
Mineral Trapping	Dissolved CO <sub>2</sub> reacts with minerals in the storage medium to form stable carbonate minerals.	- Provides the most permanent form of CO <sub>2</sub> storage. - Enhances storage security by chemically binding CO <sub>2</sub>	- Slowest trapping mechanism. - Dependent on geochemical conditions and mineral availability

3. METHODOLOGY AND SIMULATION SETUP

CMG (*Computer Modelling Group Ltd., 2023*) is used for this study. Among 13 of its products, GEM, Builder, cEdit, CMOST, and Results have been used to simulation setup, solving the system and analysis of result.

3.1 Structural and Petrophysical Modeling of Reservoir

Initially two reservoirs were modeled for the simulation. One is homogenous and the other is heterogenous. Both reservoirs are defined within a 3D Cartesian grid, delineating its structural and petrophysical attributes essential for CO<sub>2</sub> storage simulation. The model is structured into a 20×20×24 grid, translating into 9600 cells, with a uniform lateral cell dimension of 150 meters across both the X- and Y-axes which

resulted in 3000 m × 3000 m reservoir dimensions. Vertically, each cell exhibits a consistent thickness of 8.8 m, summing up to a reservoir thickness of approximately 211 m. The reservoir top boundary is placed at a depth of 1200 m below the surface, establishing the initial conditions for simulation purposes. The other properties and initial conditions of the homogenous reservoir are shown in Table 2.

**Table 2. Data used in the reservoir model**

Property	Values
Porosity	12%
Permeability (Layers 1-3)	0 millidarcies (mD)
Permeability (Horizontal)	1000 millidarcies (mD)
Permeability (Vertical)	100 millidarcies (mD)
Initial Reservoir Pressure	11800 kpa
Initial Reservoir Temperature	70°C
Compressibility Factor	5.5E-7 1/psi
Water saturation	25%
Reference depth	1200 m
Water-oil contact depth	1300 m

Porosity and permeability data of the grid cells in the heterogenous reservoir are presented in Fig. 1 which is the 3D grid view of a heterogenous reservoir.

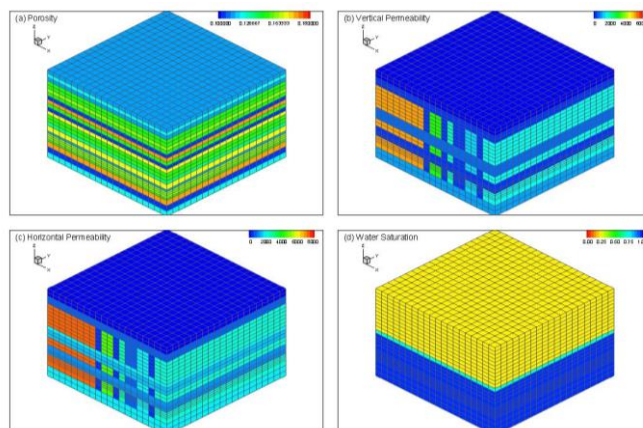


Fig. 1. Heterogenous reservoir showing a) porosity b) vertical permeability c) horizontal permeability d) water saturation.

The relative permeability data used for the study is presented in Fig. 2.

3.2 Well (CO<sub>2</sub> Injector) Modeling

A well is characterized as an injection well, with the purpose of injecting a solvent composed entirely of pure CO<sub>2</sub>. The operational parameters are governed by constraints implemented through a continuous repeat command to maintain stability and control over the injection process. These constraints are defined as follows: a maximum surface gas injection rate (STG) of 50,000 m<sup>3</sup>/day, a maximum allowable bottom hole pressure (BHP) of 30,000 kPa, an injection period set to continue for 30 years, and a total simulation period of 200 years. Perforations have been done through three cells with coordinates (1,1,18), (1,1,19), and (1,1,20).

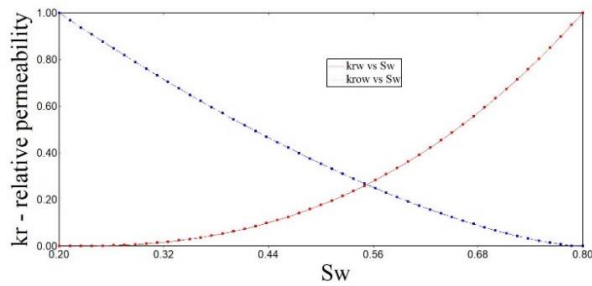


Fig. 2. Water-oil relative permeability curve.

### 3.3 Simulation Cases

Two types of reservoirs, homogenous and heterogenous, were studied for achieving the objective of this study. Table 3 presents the parameters used in all case studies done in the homogenous reservoir. The other properties will follow the base case described until this chapter.

Table 3. Simulation cases for homogenous reservoir

Case No.	Injection rate m <sup>3</sup> /d	Permeability X axis (mD)	Permeability Y axis (mD)	Permeability Z axis (mD)
5	50000	1000	1000	100
6	60000	1000	1000	100
7	70000	1000	1000	100
8	50000	1500	1500	50
9	50000	600	600	600
10	50000	100	100	100

In contrast, heterogenous reservoir have less freedom to change the parameters as the reservoir was considered as a predetermined property as it naturally is. This study, however, conducted some simulations by changing injection rate which are shown in Table 4.

Table 4. Simulation cases for heterogenous reservoir

Case No.	Injection rate (m <sup>3</sup> /d)
1	50000
2	60000
3	70000
4	80000

Sensitivity analysis was done with CMOST which performs effect estimation by using Design of Experiments (DoE) to systematically vary input parameters and run multiple simulations. It constructs a response surface to model the relationship between inputs and outputs, and conducts sensitivity analysis to quantify the impact of each parameter on the results (Wang et al., 2023). The range of parameter values used in the CMOST sensitivity analysis for the homogenous reservoir are shown in Table 5. In total 36 experiments were selected by CMOST AI within these parameter ranges. The objective function taken in the CMOST study are as follows:

- 1) Trapped CO<sub>2</sub> (due to structural and residual trapping).
- 2) Dissolved CO<sub>2</sub> mol (solubility trapping).

Table 5. Range of parameter values for sensitivity analysis

Parameter Name	Lower Limit	Upper Limit
Injection Rate	50000 m <sup>3</sup> /day	100000 m <sup>3</sup> /day
Bottom Hole Pressure	25000 KPa	37500 KPa
Horizontal Permeability	300 mD	1500 mD
Vertical Permeability	100 mD	1000 mD
Porosity	9%	15%

## 4. RESULTS AND DISCUSSION

The primary objective of the study is to analyze the dynamics of CO<sub>2</sub> plume and to assess the roles played by various trapping mechanisms, along with investigating the influence of various reservoir characteristics on the plume development and storage efficiency. The primary objective is achieved through the following objectives: investigating the size of the CO<sub>2</sub> plume under different conditions, examining how the plume develops over time with continuous extended CO<sub>2</sub> injection, evaluating the effect of different CO<sub>2</sub> trapping methods in terms of storage, and conducting a sensitivity analysis to determine the impact of various reservoir parameters on the stored CO<sub>2</sub>.

### 4.1 Plume Dynamics

An important element of CO<sub>2</sub> storage in an aquifer involves identifying the area of the aquifer surrounding the injection well that is affected by CO<sub>2</sub> injection. This affected region is referred to as the CO<sub>2</sub> plume. In this work, the criterion used to define the plume is based on the molality of CO<sub>2</sub> (Zapata et al., 2020). Because among all other criterion that generally used to define plume, CO<sub>2</sub> in aqueous phase spread across the aquifer region most, resulting in the maximum possible plume volume (Zapata et al., 2020). Cells exhibiting a CO<sub>2</sub> molality greater than the threshold of 0.4 are considered active within the plume.

In this study, plume volume is defined based on the pore volume of affected cells which are satisfying the threshold values of molality. So, the plume volume is not referring to pure CO<sub>2</sub> volume. Figure 3 is a plot of plume volume with respect to the time for both heterogenous and homogenous reservoirs. In the heterogeneous reservoir, the plume volume initially increases rapidly, indicative of varied pathways that facilitate quicker CO<sub>2</sub> spread through regions of higher permeability. In contrast, the plume in the homogeneous reservoir expands more gradually and uniformly, reflecting the consistent geological properties that regulate a steadier CO<sub>2</sub> migration. Despite these initial differences in growth rates, both scenarios eventually start to stabilize after the injection period ends. Additionally, Figure 4 is a graphical representation of the development of plume in a consistent CO<sub>2</sub> injection where the plume is defined based on the aqueous phase CO<sub>2</sub> as a function of total CO<sub>2</sub> injected with an injection rate of 50000 m<sup>3</sup>/day in standard condition. In both cases it can be observed that the plume develops almost in a linear trend with the injected CO<sub>2</sub> and the development continues even after the injection stops. This continuation generally occurred by the buoyancy force or molecular diffusion. But these stabilized quickly after the injection period ended.

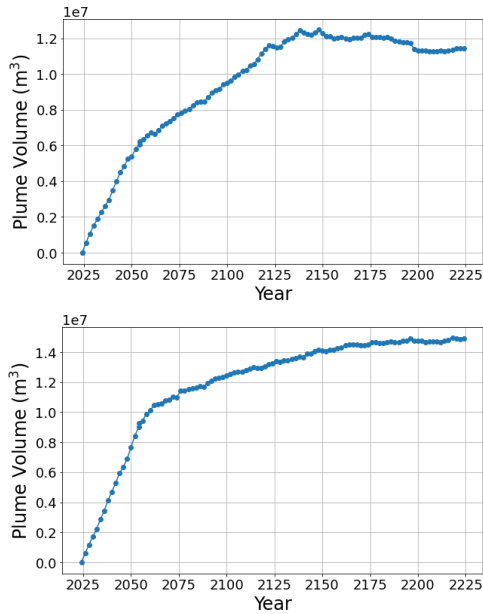


Fig. 3. Plume volume with time (years) in (a) homogenous reservoir (b) heterogeneous reservoir.

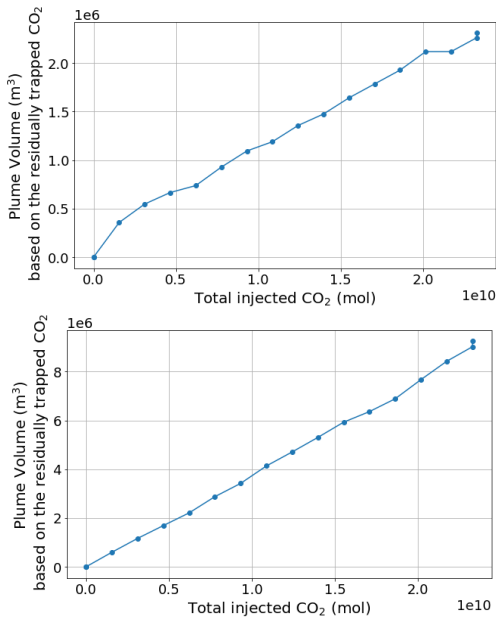


Fig. 4. Volume of CO<sub>2</sub> plume with total CO<sub>2</sub> injected in (a) homogenous (b) heterogeneous reservoir where plume is defined based on CO<sub>2</sub> in aqueous phase.

Studying plume geometry and its evolution is another objective in this study. Figure 5 illustrates how the plume shape changes over time in different scenarios cases in a homogenous reservoir. The development of the plume is influenced by the structural and layered composition of the formation. The irregular plume shape of a heterogeneous reservoir with various injection rates is presented in Fig. 6. It can be stated from the figures that in case 7 with high injection rate (70000 m<sup>3</sup>/day), the plume spread cross vertical direction is larger than the other case. Additionally, in case 10, the case of low permeability (100 mD in each direction), shows almost a very small but dense CO<sub>2</sub> plume as gas cannot move freely due to very low permeability.

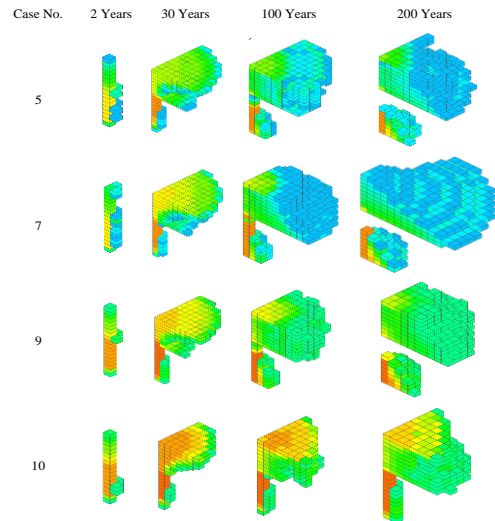


Fig. 5. Plume evolution after different time for several cases in a homogenous reservoir.

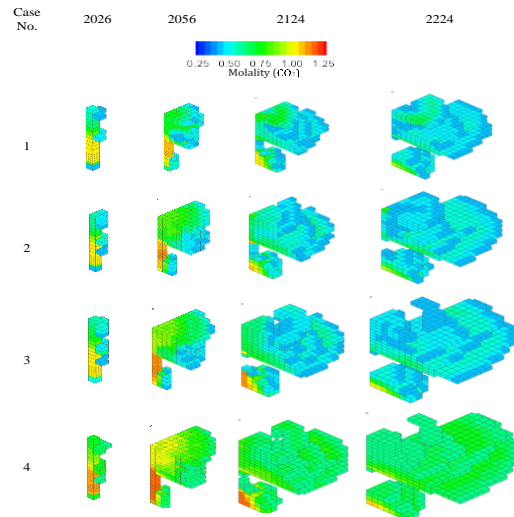


Fig. 6. Plume evolution after different time for different cases in a heterogeneous reservoir.

It can be implied that CO<sub>2</sub> plume is developing quickly in both vertical and horizontal direction during the injection period, then horizontal expansion rate becomes larger than the vertical. The permeability of the reservoir plays a critical role for the behavior of the plume evolution. At the very beginning of the injection period the injected CO<sub>2</sub> spreads vertically due to buoyancy-driven flow and initial pressure gradient with reservoir depth. Eventually, the horizontal propagation rate starts to increase as the horizontal permeability is 10 times larger than the vertical permeability. However, the propagation rate varies with the injection rate and permeability values. For example, plume evolution in case 10, with permeability value of 100 mD in each direction, has almost equal rate of propagation in both horizontal and vertical direction and the molality variation across the plume cells tends to be very low.

The heterogeneous reservoir also shows a similar trend of plume propagation. As the reservoir has larger permeability

value in horizontal direction, the plume propagates mostly horizontally. So, the radial expansion will prevail over the vertical expansion. As a result, a larger plume volume is generated after 200 years where the average CO<sub>2</sub> molality is reduced but the spreader plume volume results in a larger surface area of the CO<sub>2</sub>-brine and interaction between CO<sub>2</sub> and rock.

#### 4.2 Storage Efficiency

This chapter is focused on the impact of various trapping mechanisms in the geological storage of CO<sub>2</sub> in the aquifers. CO<sub>2</sub> is stored in two forms of super critical CO<sub>2</sub> and aqueous phase CO<sub>2</sub> due to solubility trapping mechanism. However, super critical CO<sub>2</sub> can be found in two different conditions of mobile free phase CO<sub>2</sub> and immobilized CO<sub>2</sub> by structural and residual trapping mechanism.

To compare the contributions of various trapping mechanisms over a period of 200 years, a metric known as the "storage ratio" has been established. The storage ratio is defined as the proportion of stored CO<sub>2</sub> (in moles) to the total injected CO<sub>2</sub> (in moles) expressed by: (Zapata et al., 2020).

$$\text{Storage ratio} = \frac{\text{CO}_2 \text{ stored due to that specific trapping mechanism (moles)}}{\text{Injected CO}_2 \text{ (mol)}} \quad (3)$$

The storage ratios for each mechanism across both type of reservoirs (case 1 and case 5) is presented in Fig. 7. At the initial stage, structural, residual, and solubility trapping shows a huge storage ratio which drastically falls as the injection of CO<sub>2</sub> continues and the system allows the CO<sub>2</sub> to move. On the other hand, mobile free phase CO<sub>2</sub> drastically increased and stabilized at the storage ratio around 0.9.

The scenario is elaborated in Fig. 8. It is visible that the amount of structurally and residually trapped CO<sub>2</sub> (red color) increases rapidly until the injection period ends, then it gets eclipsed by the solubility trapping (orange color). Because more CO<sub>2</sub> can move through the reservoir freely and get chance to be dissolved more in the aquifer. Moreover, some mobile free phase super critical CO<sub>2</sub> is also being dissolved over time which makes a negative slope in mobile free phase CO<sub>2</sub> curve too. As a result, even if the slope of solubility trapping curve is not too steep at the beginning but that positive slope holds with a close value throughout the long-term period.

The contribution of different trapping mechanisms in both homogenous and heterogenous reservoirs (case 1 and case 5) is presented in Fig. 9 as a stacked area diagram. It can be observed that the largest portion of CO<sub>2</sub> is retained as mobile CO<sub>2</sub> in the free phase. The ratio of mobile CO<sub>2</sub> rises rapidly throughout the injection phase; nonetheless, upon discontinuation of injection, the significance of alternative trapping mechanisms is amplified, leading to a decline in the mobile CO<sub>2</sub> fraction.

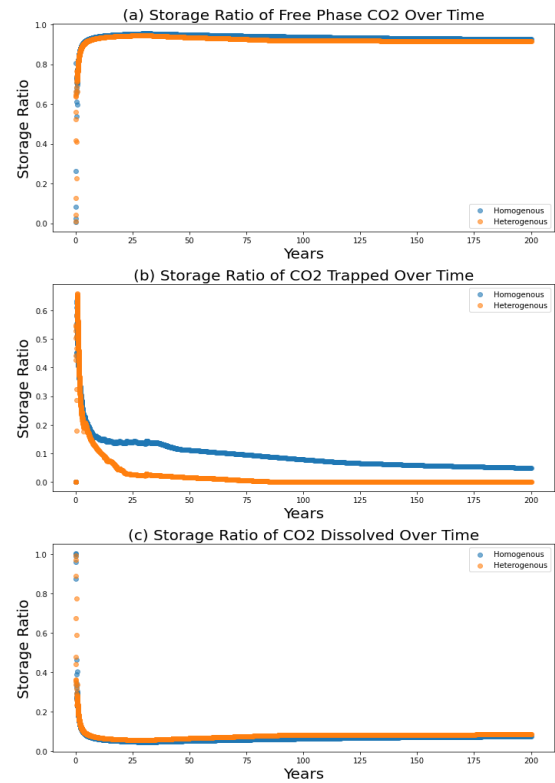


Fig. 7. Contribution of (a) mobile free phase CO<sub>2</sub> (b) residual and structural trapping and (c) solubility trapping in a homogenous and heterogenous reservoir in terms of storage ratio with time.

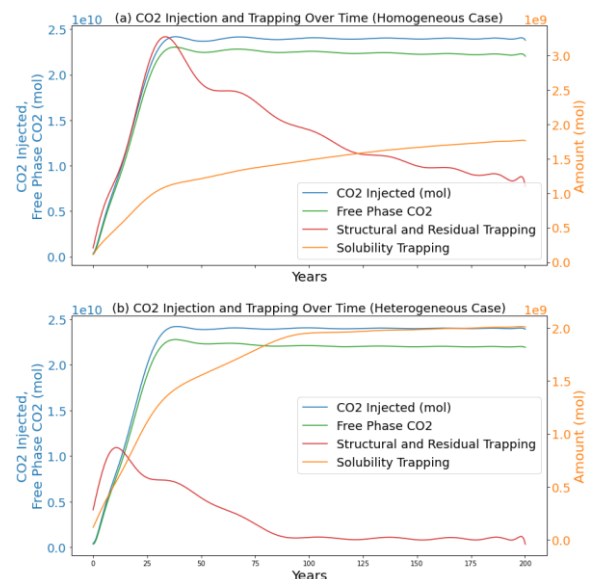


Fig. 8. Injected and stored CO<sub>2</sub> amount (mol) in (a) homogenous and (b) heterogenous case.

#### 4.3 Sensitivity Analysis

CMOST AI, an updated module was used in the sensitivity analysis. Response Surface Methodology (RSM) was employed to evaluate the effects of various operational parameters on the amount of trapped and dissolved CO<sub>2</sub>. RSM comprises statistical and mathematical techniques for exploratory experiments aimed at developing, analyzing, and optimizing various processes (Bauer Jr. et al., 1999). RSM is particularly useful and efficient in performing sensitivity



analyses for decision-making problems, offering a way to notably shorten the time required to conduct these analyses (Bauer Jr. et al., 1999). In this study, the parameters considered include bottom hole pressure (kPa), injection rate (m<sup>3</sup>/day), horizontal permeability (mD), vertical permeability (mD), and porosity. The objective functions considered are trapped CO<sub>2</sub> and dissolved CO<sub>2</sub>.

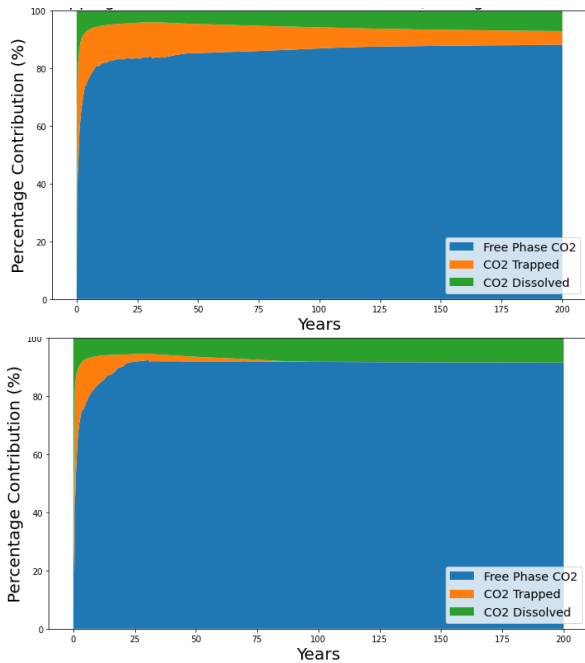


Fig. 9. Contribution of different trapping mechanism in (a) homogenous reservoir (b) heterogenous reservoir.

The charts displayed in Figs. 10 and 11 show the normalized impact of each parameter against the maximum effect value observed in our models.

From the analysis, it is evident that certain parameters are particularly influential, which are described in the following. The sensitivity analysis for trapped CO<sub>2</sub> reveals several key insights:

- 1) The horizontal permeability shows the most substantial positive quadratic effect, indicating that higher horizontal permeability increases CO<sub>2</sub> trapping significantly.
- 2) Although injection rate is positively impacting the trapped CO<sub>2</sub> amount, it has a negative quadratic impact which shows that after a certain level, the trapped CO<sub>2</sub> will decrease with the increase in injection rate.
- 3) The squared term of bottom hole pressure also positively affects CO<sub>2</sub> trapping.
- 4) Porosity and its squared value moderately influence CO<sub>2</sub> trapping, indicating that more porous formations tend to trap more CO<sub>2</sub>.

The analysis of dissolved CO<sub>2</sub> presented in Fig. 11 indicates that porosity exhibits the strongest positive effect on the dissolution of CO<sub>2</sub> into the reservoir fluids. Higher porosity levels enhance the capacity for CO<sub>2</sub> dissolution, due to the increased fluid interactions within porous media. Then both injection rate and horizontal permeability positively impact CO<sub>2</sub> dissolution.

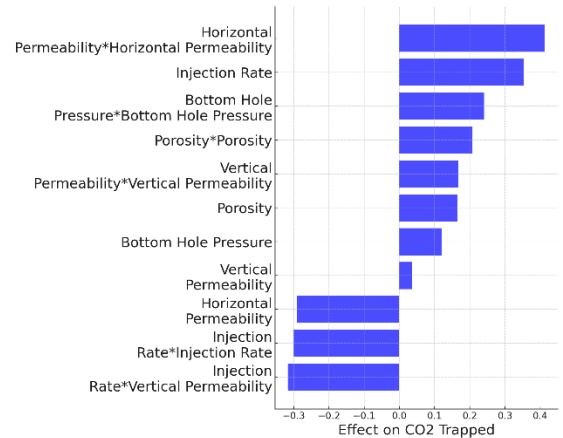


Fig. 10. Relative impact of different parameters on structural and residually trapped CO<sub>2</sub>.

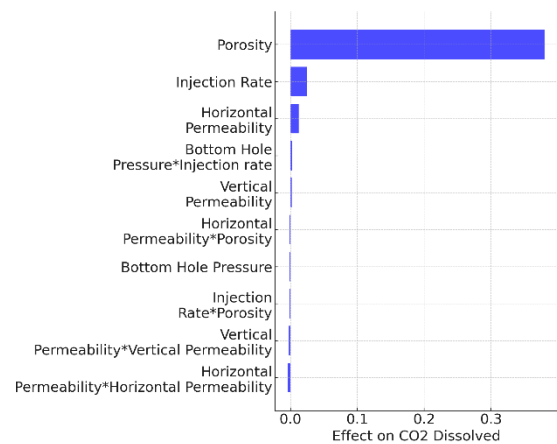


Fig. 11. Relative impact of different parameters on dissolved CO<sub>2</sub> amount (moles).

## 5. CONCLUSIONS

CO<sub>2</sub> plume dynamics, storage capacity, and impact of different reservoir properties and parameters were explored in this study. The results show that horizontal plume spread exceeds vertical due to higher horizontal permeability. Case studies showed that permeability and injection rate significantly influence plume volume, with higher rates and permeabilities resulting in larger plumes. Additionally, the plume volume shows a linear relationship with the injected CO<sub>2</sub> amount. In terms of storage efficiency, the most prevalent CO<sub>2</sub> is free phase super critical CO<sub>2</sub> that contributes around 80% of the stored CO<sub>2</sub> whereas the rest are structurally or residually trapped and dissolved CO<sub>2</sub>. Initially, trapped CO<sub>2</sub> contributed almost 15%. Over time, some of the trapped CO<sub>2</sub> dissolved into the reservoir or aquifer fluid. This led to a reduction in the percentage contribution to structural and residual trapping mechanisms, decreasing to 5% in homogeneous reservoirs and 0% in heterogeneous reservoirs. At the same time, the percentage contributed to solubility trapping increased to 15% in homogeneous reservoirs and 20% in heterogeneous reservoirs. Sensitivity analyses revealed that horizontal permeability and injection rate significantly affect trapped CO<sub>2</sub>, while porosity impacts CO<sub>2</sub> dissolution. The Future research should incorporate more realistic reservoir models, explore mineral trapping, and conduct further sensitivity analyses.



## REFERENCES

- Aakre, H., Mathiesen, V., and Moldestad, B. (2018). Performance of CO<sub>2</sub> flooding in a heterogeneous oil reservoir using autonomous inflow control. *Journal of Petroleum Science and Engineering*, 167, 654–663. doi:10.1016/j.petrol.2018.04.008
- Ahmed, T. (2010). Chapter 14—Principles of Waterflooding. In T. Ahmed (Ed.), *Reservoir Engineering Handbook (Fourth Edition)* (pp. 909–1095). Gulf Professional Publishing. doi:10.1016/B978-1-85617-803-7.50022-5
- Al-Khdheawi, E. A., Vialle, S., Barifcani, A., Sarmadivaleh, M., and Iglauer, S. (2018). Effect of wettability heterogeneity and reservoir temperature on CO<sub>2</sub> storage efficiency in deep saline aquifers. *International Journal of Greenhouse Gas Control*, 68, 216–229. doi:10.1016/j.ijggc.2017.11.016
- Bachu, S. (2008). CO<sub>2</sub> storage in geological media: Role, means, status and barriers to deployment. *Progress in Energy and Combustion Science*, 34(2), 254–273.
- Bauer Jr., K. W., Parnell, G. S., and Meyers, D. A. (1999). Response surface methodology as a sensitivity analysis tool in decision analysis. *Journal of Multi-Criteria Decision Analysis*, 8(3), 162–180. doi:10.1002/(SICI)1099-1360(199905)8:3<162::AID-MCDA241>3.0.CO;2-X
- Birkholzer, J. T., Oldenburg, C. M., and Zhou, Q. (2015). CO<sub>2</sub> migration and pressure evolution in deep saline aquifers. *International Journal of Greenhouse Gas Control*, 40, 203–220. doi:10.1016/j.ijggc.2015.03.022
- Computer Modelling Group Ltd. (2023). [Computer software]. CMG. <https://www.cmg.ca/>
- Godec, M., Kuuskraa, V., Van Leeuwen, T., Stephen Melzer, L., and Wildgust, N. (2011). CO<sub>2</sub> storage in depleted oil fields: The worldwide potential for carbon dioxide enhanced oil recovery. *Energy Procedia*, 4, 2162–2169. doi:10.1016/j.egypro.2011.02.102
- Ketzer, J. M., Iglesias, R. S., and Einloft, S. (2012). Reducing Greenhouse Gas Emissions with CO<sub>2</sub> Capture and Geological Storage. In W.-Y. Chen, J. Seiner, T. Suzuki, and M. Lackner (Eds.), *Handbook of Climate Change Mitigation* (pp. 1405–1440). Springer US. doi:10.1007/978-1-4419-7991-9\_37
- Kulynycz, V. (2015). The influence of wettability on oil recovery. *AGH Drilling, Oil, Gas*, 32, 493. doi:10.7494/drill.2015.32.3.493
- Luo, A., Li, Y., Chen, X., Zhu, Z., and Peng, Y. (2022). Review of CO<sub>2</sub> sequestration mechanism in saline aquifers. *Natural Gas Industry B*, 9(4), 383–393. doi:10.1016/j.ngib.2022.07.002
- Moritis, G. (2006). CO<sub>2</sub> injection gains momentum. *Oil and Gas Journal*, 104, 37–41.
- Myshakin, E. M., Haeri, F., Moore, J., Crandall, D., and Goodman, A. L. (2023). Numerical Simulations of Carbon Dioxide Storage Efficiency in Heterogeneous Reservoir Models. *Geofluids*, 2023, 5089508. doi:10.1155/2023/5089508
- Raza, A., Rezace, R., Gholami, R., Rasouli, V., Bing, C. H., Nagarajan, R., and Hamid, M. A. (2015). Injectivity and quantification of capillary trapping for CO<sub>2</sub> storage: A review of influencing parameters. *Journal of Natural Gas Science and Engineering*, 26, 510–517. doi:10.1016/j.jngse.2015.06.046
- Riddiford, F., Tourqui, A., Bishop, C., Taylor, B., and Smith, M. (2003). *A cleaner development: The In Salah Gas project, Algeria*. 595–600.
- Saedi, A. (2012). *Experimental study of multiphase flow in porous media during CO<sub>2</sub> Geo-Sequestration processes*. Springer Science & Business Media.
- Szulczewski, M. L. (2013). *The Subsurface Fluid Mechanics of Geologic Carbon Dioxide Storage* by.
- Taghavi, S., Tahami, S. A., Aakre, H., Furuvik, N. C. I., and Moldestad, B. M. E. (2023). Performance Analysis of Autonomous Inflow Control Valve in a Heterogeneous Reservoir Using CO<sub>2</sub> Enhanced Oil Recovery. *Day 3 Wed, October 18, 2023*, D031S045R002. doi:10.2118/215153-MS
- Torp, T. A., and Gale, J. (2003). *Demonstrating storage of CO<sub>2</sub> in geological reservoirs: The Sleipner and SACS projects*. 311–316.
- Van Der Meer, B. (2005). Carbon Dioxide Storage in Natural Gas Reservoir. *Oil & Gas Science and Technology*, 60(3), 527–536. doi.org/10.2516/ogst:2005035
- Verma, Y., Vishal, V., and Ranjith, P. G. (2021). Sensitivity Analysis of Geomechanical Constraints in CO<sub>2</sub> Storage to Screen Potential Sites in Deep Saline Aquifers. *Frontiers in Climate*, 3, 720959. doi.org/10.3389/fclim.2021.720959
- Wang, Y., Luo, J., and Shang, Z. (2023). Sensitivity Analysis of Oil and Gas Production in the In Situ Pyrolysis of Oil Shale. *Processes*, 11(7). doi.org/10.3390/pr11071948
- Whittaker, S., Wilson, M., and Monea, M. (2004). *IEA GHG Weyburn CO<sub>2</sub> monitoring & storage project summary report 2000-2004: From the proceedings of the 7th International Conference on Greenhouse Gas Control Technologies : September 5-9, Vancouver, Canada : Volume III*. Petroleum Technology Research Centre.
- Zapata, Y., Kristensen, M. R., Huerta, N., Brown, C., Kabir, C. S., and Reza, Z. (2020). CO<sub>2</sub> geological storage: Critical insights on plume dynamics and storage efficiency during long-term injection and post-injection periods. *Journal of Natural Gas Science and Engineering*, 83, 103542. doi:10.1016/j.jngse.2020.103542
- Zhi, S., Elsworth, D., and Liu, L. (2019). W-shaped permeability evolution of coal with supercritical CO<sub>2</sub> phase transition. *International Journal of Coal Geology*, 211, 103221. doi:10.1016/j.coal.2019.103221

# Equilibrium analysis for methanation focusing on CO<sub>2</sub> derived substitute natural gas

Rakhi and Fabian Mauss

*BTU Cottbus - Senftenberg, Siemens-Halske-Ring 14, 03046 Cottbus,  
Germany (e-mail: rakhi.rakhi@b-tu.de).*

**Abstract:** In this study the methanation of synthesis gas (syngas) is investigated with a focus on achieving maximum methane and minimum CO by full methanation of CO<sub>2</sub>. For this study, we have considered a comprehensive thermodynamics analysis of CO<sub>2</sub> hydrogenation. This will help us to understand the thermodynamic behaviour of the reactions involved in the methanation process. We have discussed the behavior of the species, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, and H<sub>2</sub>O at the equilibrium with temperature, pressure, and fuel ratio variation in order to get the desired output. The preliminary study will focus on selecting the optimum conditions (temperature, pressure, and H<sub>2</sub>/CO<sub>2</sub> ratio) for performing the experiments and for catalyst development.

*Keywords:* Equilibrium calculations, Thermodynamic analysis, CO<sub>2</sub> methanation, Gibbs free energy minimization

## 1. INTRODUCTION

It is assumed that the main element for the increase in atmospheric temperature is the increase of the CO<sub>2</sub> concentration. The greenhouse gas CO<sub>2</sub>, therefore, needs to be reduced in the scope of the energy transition from fossil to renewable sources. This needs new concepts for the sustainable energy supply and also for energy storage. Hence, it is essential to reintegrate the secondary products like H<sub>2</sub> and CO<sub>2</sub> into the energy supply in order to reduce the direct CO<sub>2</sub> emission discussed by Kopyscinski et al. (2010); Seemann et al. (2010).

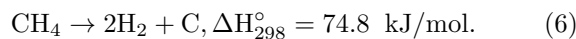
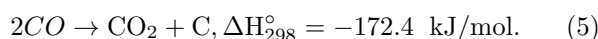
The conversion of CO<sub>2</sub> into methane is a promising approach for a CO<sub>2</sub> neutral production circle and this process is commonly known as methanation. The excess energy produced in methanation by renewable energies is converted into chemical energy. There is a possibility to feed the produced CH<sub>4</sub> into the existing network of natural gas referred as ‘power-to-gas’ approach (PtG) (Müller et al. (2013, 2019)).

The sufficient supply of H<sub>2</sub> required for the hydrogenation of CO<sub>2</sub> is generated by some kind of renewable energy. This ensures a CO<sub>2</sub> neutral process (Ursua et al. (2012); Razaq. et al. (2013); Müller et al. (2019)). The exhaust CO<sub>2</sub> can be used as a CO<sub>2</sub> source to fully convert the hydrogen/CO<sub>2</sub> to methane by the Sabatier reaction:



This reaction is thermodynamically favoured and catalysts can be useful to achieve acceptable conversion from CO<sub>2</sub> into CH<sub>4</sub> (Du et al. (2007); Ma et al. (2009); Müller et al. (2017); Rachow (2017)). In this study, we want to comprehend the thermodynamics involved in the methanation processes.

Other than the Sabatier reaction, the following main competitive reactions depending on the fuel composition also needs to be considered:



The above listed reactions are some of the important reactions expected to happen in a methanation process. However, in this study our focus is only to consider the thermodynamic point of view by using the thermochemistry of all the species which are used in the system. Some of the thermodynamic investigations of methanation reactions are discussed by Greyson et al. (1955); Anderson (1986); Gao et al. (2012); Jia et al. (2016) and catalytic studies are conducted by Beuls et al. (2012); Ocampo et al. (2009); Hu et al. (2012).

Although there are some thermodynamic investigations available in literature, we still need a further comprehension of the complex methanation reactions. Hence, in this work, we discuss the thermodynamic impact of CO<sub>2</sub> methanation on formation of products. The main species considered for the study are CH<sub>4</sub>, CO<sub>2</sub>, H<sub>2</sub>, CO, and H<sub>2</sub>O with a focus to maximize the methane formation and

minimize the CO formation for the considered conditions for investigation.

## 2. METHODOLOGY

Apart from the LOGEcat model for catalyst investigations used by Rakhi et al. (2022b,a, 2023); Rakhi and Mauss (2024) from the LOGEsoft software suite (LOGEsoft (2008)), there is also an Equilibrium Reactor model in the software package. We have used the equilibrium reactor model to perform the simulations discussed in this paper. We only need the thermodynamic data for all the species involved in the methanation process in each phase for the equilibrium reactor model and this allows to determine the chemical state of a mixture under equilibrium conditions including any number of gas-phase or bulk species for the thermodynamic equilibrium calculations.

The equilibrium composition of a reactive system is calculated using the Gibbs free energy minimization method which is based on the principle that the total Gibbs energy of the system has its minimum value at chemical equilibrium without considering the individual equilibrium constants (Adhikari et al. (2007); Ozkara-Aydinoglu (2010); Rossi et al. (2009)). The distribution of the products under a minimum free energy is achieved by utilizing a general mathematical technique which does not require the knowledge of the chemistry of the reactions. However, all the species in a reaction system including reactants as well as the products needs to be given. Then the Gibbs energy is calculated using the Gibbs energy under standard conditions and the law of the mass action:

$$\Delta_r G = \Delta_r G^\theta + RT \ln \prod_i a_i^{\nu_i} \quad (7)$$

The meaning of the symbols,  $\Delta_r G$ ,  $\Delta_r G^\theta$ ,  $R$ ,  $T$ ,  $\prod$ ,  $a$ , and  $\nu$  are the change in the Gibbs free energy, standard Gibbs free energy for the reaction  $r$ , universal gas constant, temperature, product across all  $i$ -indexed variables, activity coefficient, and stoichiometric coefficient of species  $i$ , respectively. Since the change in the free energy at equilibrium is zero, the equilibrium constant ( $K_{eq}$ ) is given as:

$$K_{eq} = \exp\left(\frac{-\Delta_r G^\theta}{RT}\right) \quad (8)$$

This equations determines the direction of the reaction. The change in the standard free energy,  $\Delta_r G^\theta$  can be calculated using the Gibbs-Helmholtz equation given as:

$$\Delta_r G^\theta = \Delta H^\theta - T \Delta S^\theta \quad (9)$$

The symbols,  $\Delta H^\theta$  and  $\Delta S^\theta$  represent the change in the reaction enthalpy and the change in the reaction entropy, respectively. The polynomial functions of temperature are used to determine the state functions and heat capacity in the equilibrium reactor model. The polynomial coefficients for all the species available in the methanation process can then be provided in the state function input file using a standard format for NASA coefficients. These polynomials can be used to drive all the other

thermodynamic properties needed for the thermodynamic equilibrium calculations. The Gibbs free energy for the Sabatier reaction (Equation 1) can be calculated using the equations ( $t=T[K]/1000$ ):

$$\Delta H^\theta(t) = \Delta H_{298.15K}^0 + dH_{CH_4} + 2 * dH_{H_2O} - dH_{CO_2} - 4 * dH_{H_2} \quad (10)$$

$$\Delta S^\theta(t) = S_{CH_4}^0 + 2 * S_{H_2O}^0 - S_{CO_2}^0 - 4 * S_{H_2}^0 \quad (11)$$

$$\Delta_r G^\theta(T) = \Delta H^\theta(t) - T * \frac{\Delta S^\theta(t)}{1000} \quad (12)$$

The equilibrium constant,  $K_{eq}$ , is expected to be reduced with increasing temperature for the Sabatier reaction (Equation 1) due to its exothermic nature. While solving all the above equations, a system of linear simultaneous equations is achieved containing many unknowns and it yields a new composition representing new approximation of the composition giving minimum free energy. The information achieved with these thermodynamic equilibrium calculations can be useful to develop a catalyst for the methanation process by providing a background in selecting the most favourable conditions for the experiments and catalytic simulations.

A detailed summary of the possible reactions involved in the methanation of carbon oxides is given by Mills and Steffgen (1974); Nahar and Madhani (2010). We have used the gaseous compounds, CO, H<sub>2</sub>O, CO<sub>2</sub>, H<sub>2</sub>, O<sub>2</sub>, and CH<sub>4</sub> for the equilibrium calculations. The high hydrocarbons, solid carbon, and oxygen-containing compounds (methanol, methanoic acid, acetic acid etc.) are ignored because they are available in very small amount in the equilibrium gas mixture.

## 3. VALIDATION

The validation of the equilibrium solver is done against the literature results from Gao et al. (2012). The validation is done for one inlet condition, i.e., H<sub>2</sub>/CO<sub>2</sub>=4 at 1 atm in a temperature range of 200-800°C. Note that the equilibrium calculations provided in literature are also simulations and these simulations are performed using the CHEMCAD solver. Therefore, we have taken the reference data simulated with CHEMCAD and compared with the calculations from our equilibrium reactor model from LOGEsoft software package.

For the considered inlet condition for validation, we have compared the product fraction of CO<sub>2</sub> methanation at equilibrium for all the species, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>O, and CO before exploring the new conditions. The equilibrium calculations with our solver matched with the reference very well (not shown here). After a successful validation, the solver is further used for equilibrium calculations at various unexplored conditions to select the optimum parameters for the methanation process and the the inlet conditions considered for the simulations are given in Table 1.

Table 1. Summary of the inlet fuel composition on which the investigation is performed.

Simulation Run	H <sub>2</sub> /CO <sub>2</sub> (vol. %)	H <sub>2</sub> (vol. %)	CO <sub>2</sub>
R1	2	66.6	33.4
R2	4	80.0	20.0
R3	6	85.7	14.3

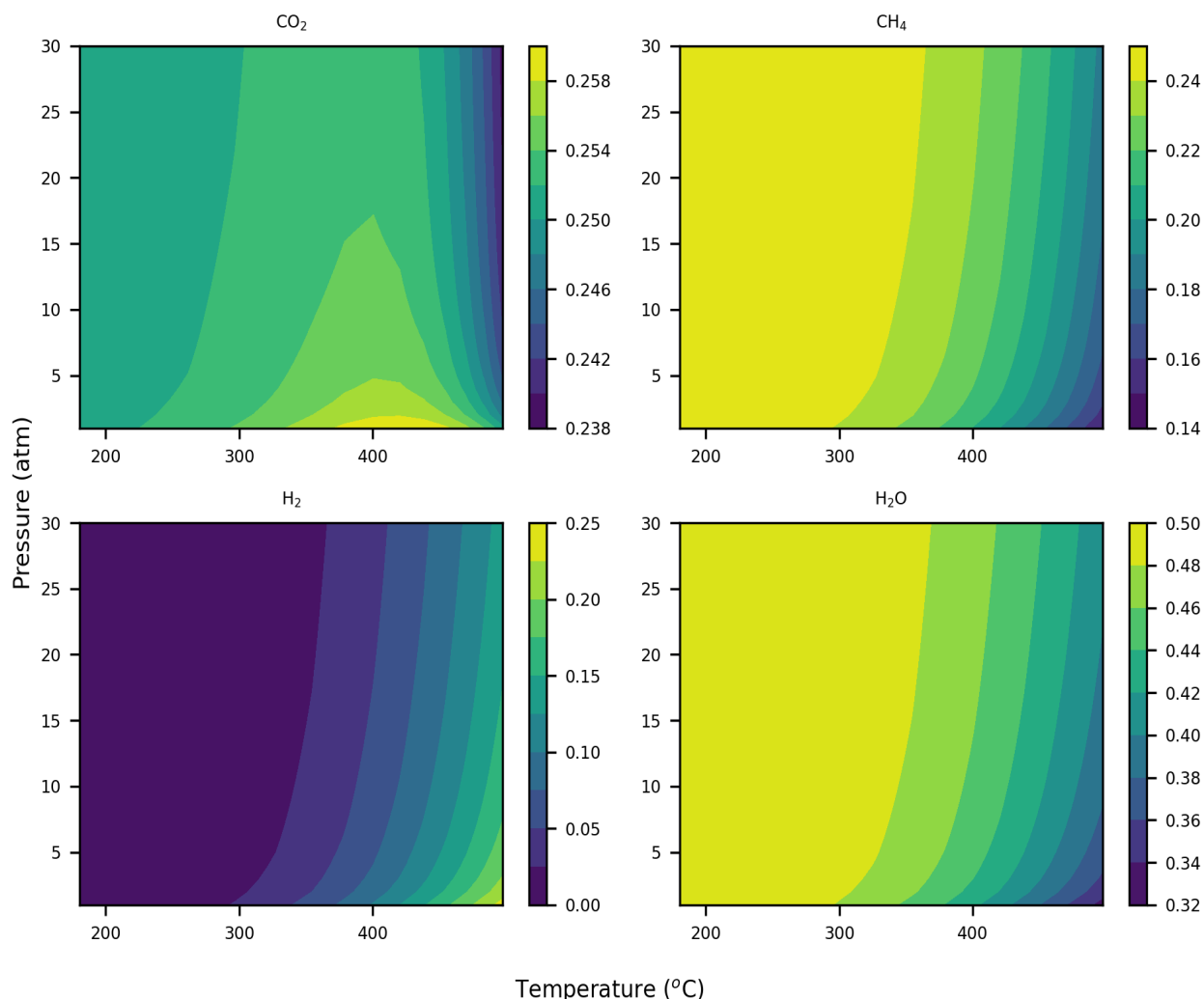


Fig. 1. Products fractions at the reactor outlet for species, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, and H<sub>2</sub>O varying with the pressure and temperature for H<sub>2</sub>/CO<sub>2</sub>=2.

#### 4. RESULTS

After validating the equilibrium reactor model using the methodology explained in the previous sections, the model is further applied to investigate the effect of CO<sub>2</sub> methanation. The simulations performed with the inlet conditions given in Table 1 are discussed in this section to select the H<sub>2</sub>/CO<sub>2</sub> fuel composition for catalytic simulations. Note that the temperature used to perform the simulations is varied in the range 200-500°C and pressure from 1 atm to 30 atm for all the inlet conditions given in the table.

Starting with the low fuel ratio, i.e., H<sub>2</sub>/CO<sub>2</sub>=2, where the amount of hydrogen is double the amount of CO<sub>2</sub> at the inlet, the measurement of the species, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, and H<sub>2</sub>O varying with the pressure and temperature is presented in Fig. 1. With this inlet composition, the CO<sub>2</sub>

is mostly unused in the entire temperature and pressure range considered for the study.

However, hydrogen is fully used as indicated in the lower panel of Fig. 1. The H<sub>2</sub> mole fraction at the reactor outlet is zero or close to zero up to 400 °C in the entire pressure range, i.e., up to 30 atm shown in the figure with blue colour. The figure also depicts that with increasing temperature and pressure, more unreacted hydrogen is expected at the reactor outlet for the methanation process.

Hence, low temperatures and low pressures are recommended for better conversion of hydrogen for this fuel composition. However, for H<sub>2</sub>/CO<sub>2</sub>=2, the low temperatures and low pressures are not favourable for CO<sub>2</sub> conversion. The CO<sub>2</sub> conversion is expected to be good for high temperatures.

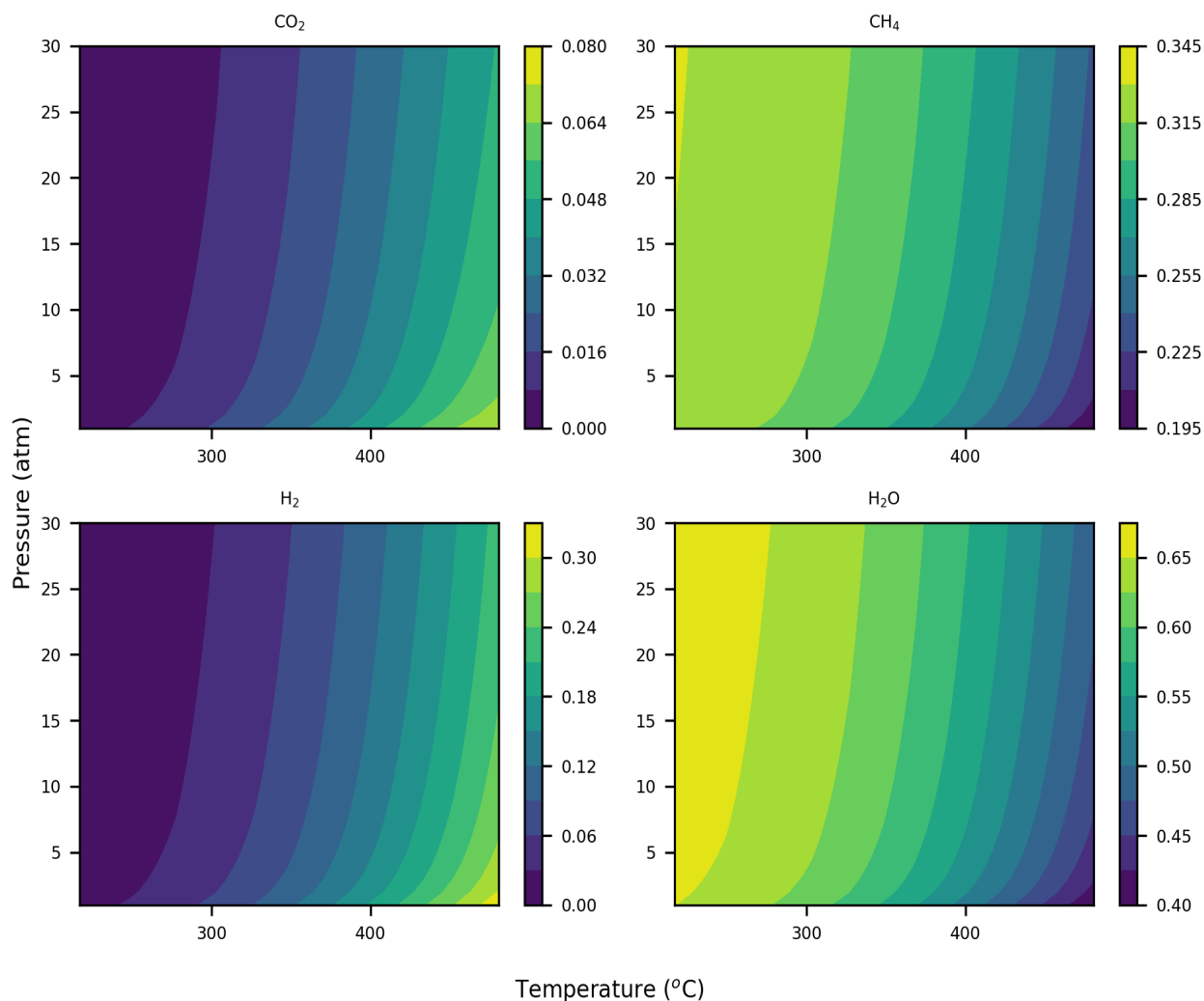


Fig. 2. Products fractions at the reactor outlet for species,  $\text{CO}_2$ ,  $\text{H}_2$ ,  $\text{CH}_4$ , and  $\text{H}_2\text{O}$  varying with the pressure and temperature for  $\text{H}_2/\text{CO}_2=4$ .

$\text{CH}_4$  and  $\text{H}_2\text{O}$  product fractions shown in the right side of Fig. 1 indicates that methane as well as water formation at low temperature in the entire pressure range considered for the investigation is very high shown with the yellow colour in the figure. So, in order to maximise the methane formation, it may be useful to operate the reactor at the condition with fuel ratio,  $\text{H}_2/\text{CO}_2=2$  at low temperature. However, this fuel ratio is not favourable considering the conversion of  $\text{CO}_2$ . Therefore, we considered also the other fuel compositions for the investigation.

Next, we computed the thermodynamic equilibrium by increasing the hydrogen mole fraction and reducing the  $\text{CO}_2$  mole fraction at the inlet condition, i.e.,  $\text{H}_2/\text{CO}_2=4$  and the product fraction of all the species at the considered temperature and pressure range is shown in Fig. 2. The species  $\text{CO}_2$  and  $\text{H}_2$  mole fractions are on the left side and for  $\text{CH}_4$  and  $\text{H}_2\text{O}$  on the right side of the figure.

Note that by reducing the  $\text{CO}_2$  mole fraction in the fuel composition leads to almost 100% conversion of  $\text{CO}_2$ . This is indicated in the figure by the blue colour showing zero unreacted  $\text{CO}_2$  at low temperatures. For operation at high temperatures, high pressures are recommended. From the

$\text{H}_2$  mole fraction scale, we note that the conversion for this species is very good even after increasing hydrogen for this fuel composition, i.e.,  $\text{H}_2/\text{CO}_2=4$ .

From right side of the figure, we note the the formation of methane and water is good for this fuel ratio as well, however, both the species mole fraction at the reactor outlet in equilibrium for  $\text{H}_2/\text{CO}_2=4$  are slightly reduced.

The hydrogen is further increased and  $\text{CO}_2$  is reduced to analyse the influence for another fuel ratio, i.e.,  $\text{H}_2/\text{CO}_2=6$  on all the species at thermodynamic equilibrium. The results for this fuel ratio are shown in Fig. 3 for  $\text{CO}_2$  and  $\text{H}_2$  mole fractions on the left side and for  $\text{CH}_4$  and  $\text{H}_2\text{O}$  on the right side of the figure.

By comparing the scales for  $\text{CO}_2$  in Fig. 2 and 3, the  $\text{CO}_2$  consumption is improved more by further reducing the  $\text{CO}_2$  mole fraction in the initial fuel composition. This behaviour is seen in the complete range of temperature and pressure considered for the calculation. However, with the increase in the hydrogen for  $\text{H}_2/\text{CO}_2=6$ , a lot of unreacted hydrogen is noted at the equilibrium. For improving the hydrogen conversion for this fuel composition,



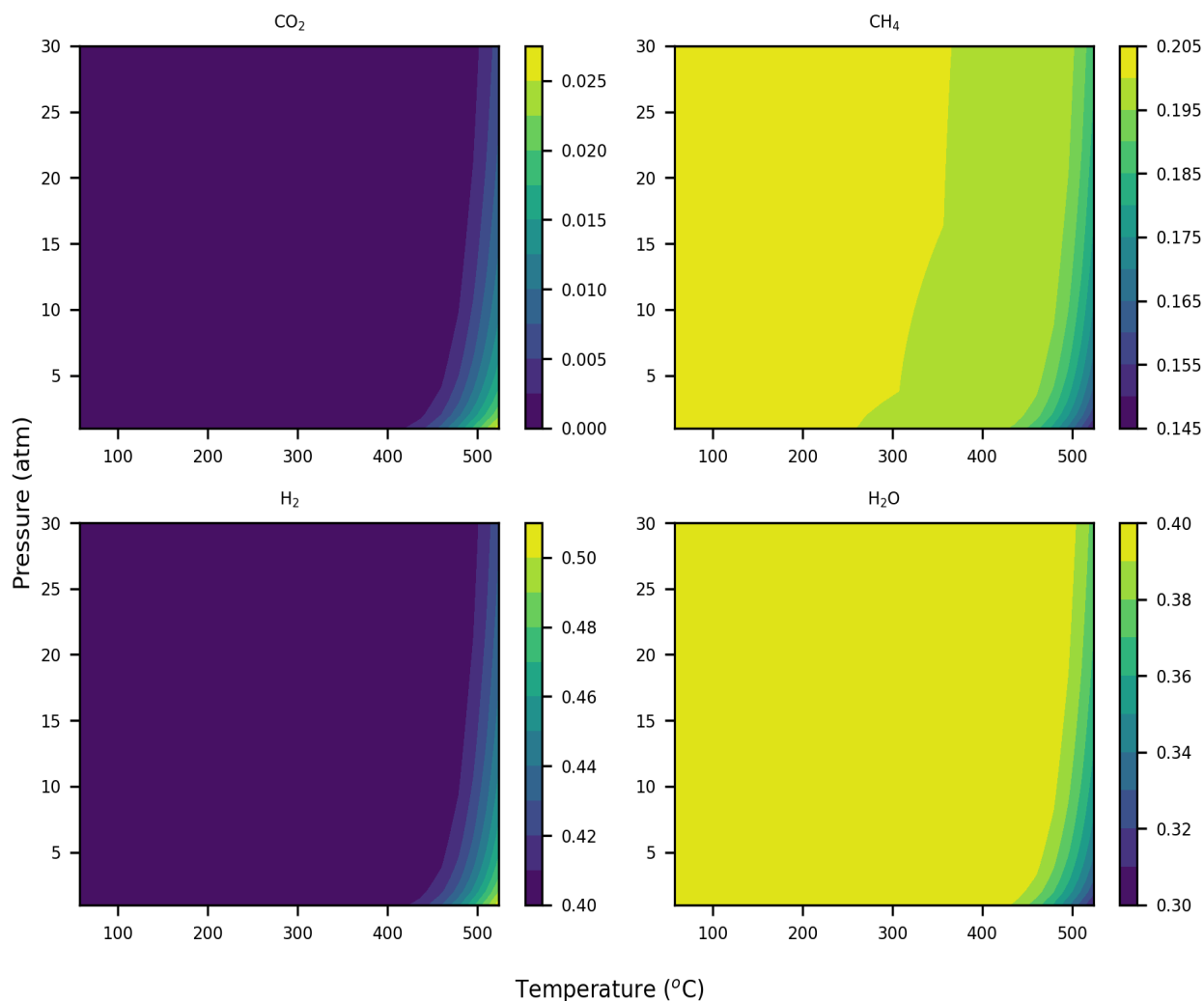


Fig. 3. Products fractions at the reactor outlet for species,  $\text{CO}_2$ ,  $\text{H}_2$ ,  $\text{CH}_4$ , and  $\text{H}_2\text{O}$  varying with the pressure and temperature for  $\text{H}_2/\text{CO}_2=6$ .

high temperatures and pressures can be investigated. This indicated that this fuel composition is not favourable for hydrogen conversion.

Not only hydrogen is unreacted but also the formation of methane at equilibrium for this fuel composition is reduced. This can be seen by comparing the scales for methane from Fig. 2 and 3. Similar to methane, water formation is also reduced for this fuel composition.

Considering the conversion of  $\text{CO}_2$ ,  $\text{H}_2/\text{CO}_2=2$  is not favourable and for conversion of  $\text{H}_2$ ,  $\text{H}_2/\text{CO}_2=6$  is not favourable.  $\text{H}_2/\text{CO}_2=4$  is the most favourable fuel composition for conversion of  $\text{H}_2$  as well as  $\text{CO}_2$  along with maximum formation of methane for this ratio.

## 5. CONCLUSIONS

The thermodynamic equilibrium composition for  $\text{CO}_2$  methanation of the species,  $\text{CO}_2$ ,  $\text{H}_2$ ,  $\text{CH}_4$ , and  $\text{H}_2\text{O}$  are shown at the reactor outlet for the Sabatier temperature range, i.e., approximately 250-450°C in the pressure range from 1 atm to 30 atm.

For the equilibrium calculations for the methanation process, fuel composition plays very important role in the entire pressure and temperature range considered for the investigation. For the low range of temperature,  $\text{H}_2/\text{CO}_2$  should be more than 2 to achieve good conversion of  $\text{CO}_2$ , however, this ratio should be less than 6 for good conversion of  $\text{H}_2$ . For maximum methane formation,  $\text{H}_2/\text{CO}_2=4$  is the most favourable fuel composition.

## ACKNOWLEDGEMENTS

Financial support by the federal ministry of education and research (Bundesministerium für Bildung und Forschung, BMBF) under the Grant Number 03SF0693A of the collaborative research project “Energie-Innovationszentrum” is gratefully acknowledged. The authors also thank the Graduate Research School (GRS) of the BTU Cottbus-Senftenberg for the partial financial support.

## REFERENCES

Adhikari, S., Fernando, S., and Haryanto, A. (2007). A comparative thermodynamic and experimental analysis on

- hydrogen production by steam reforming of glycerin. *Energy Fuels*, 21, 2306–2310.
- Anderson, R.B. (1986). Thermodynamics of the hydrogenation of oxides of carbon. *Journal of Physical Chemistry*, 90 (20), 4806–4810.
- Beuls, A., Swalus, C., Jacquemin, M., Heyen, G., Karelavic, A., and Ruiz, P. (2012). Methanation of CO<sub>2</sub>: Further insight into the mechanism over Rh/γ-Al<sub>2</sub>O<sub>3</sub> catalyst. *Applied Catalysis B*, 113–114, 2–10. doi:10.1016/j.apcatb.2011.02.033.
- Du, G.A., Lim, S., Yang, Y.H., Wang, C., Pfefferle, L., and Haller, G.L. (2007). Methanation of carbon dioxide on ni-incorporated MCM-41 catalysts: The influence of catalyst pretreatment and study of steady-state reaction. *Journal of Catalysis*, 249 (2), 370–379.
- Gao, J., Wang, Y., Ping, Y., Hu, D., Xu, G., Gu, F., and Su, F. (2012). A thermodynamic analysis of methanation reactions of carbon oxides for the production of synthetic natural gas. *RSC Advances*, 2, 2358–2368. doi:10.1039/c2ra00632d.
- Greyson, M., Demeter, J.J., Schlesinger, M.D., Johnson, G.E., Jonakin, J., and Myers, J.W. (1955). Synthesis of Methane. D. O. Interior Report 5137, Bureau of Mines, U.S.
- Hu, D., Gao, J., Ping, Y., Jia, L., Gunawan, P., Zhong, Z., Xu, G., Gu, F., and Su, F. (2012). Enhanced Investigation of CO Methanation over Ni/Al<sub>2</sub>O<sub>3</sub> Catalysts for Synthetic Natural Gas Production. *Industrial % Engineering Chemistry Research*, 51, 4875–4886. doi:10.1021/ie300049f.
- Jia, C., Gao, J., Dai, Y., Zhang, J., and Yang, Y. (2016). The thermodynamic analysis and experimental validation for complicated systems in CO<sub>2</sub> hydrogenation process. *Journal of Energy Chemistry*, 25, 1027–1037. doi:10.1016/j.jechem.2016.10.003.
- Kopyscinski, J., Schildhauser, T.J., and Biollaz, S.M.A. (2010). Production of synthetic natural gas (SNG) from coal and dry biomass – A technology review from 1950 to 2009. *Fuel*, 89, 1763–1783. LOGEsoft (2008). V1.10 www.logesoft.com.
- Ma, J., Sun, N., Zhang, X., Zhao, N., Xiao, F., Wie, W., and Sun, Y. (2009). A short review of catalysis for CO<sub>2</sub> conversion. *Catalysis Today*, 148, 221–231.
- Mills, G.A. and Steffgen, F.W. (1974). Catalytic methanation. *Catalysis Reviews*, 8, 159–210. doi:10.1080/01614947408071860.
- Müller, K., Rachow, F., Günther, V., and Schmeisser, D. (2019). Methanation of Coke Oven Gas with Nickel-based Catalysts. *International Journal of Environmental Science*, 4, 73–79.
- Müller, K., Rachow, F., Israel, J., Charlafti, E., Schwiertz, C., and Schmeisser, D. (2017). Direct Methanation of Flue Gas at a Lignite Power Plant. *International Journal of Environmental Science*, 2, ISSN: 2367–8941.
- Müller, K., Stˆadter, M., Rachow, F., Hoffmannbeck, D., and Schmeisser, D. (2013). Sabatier-based CO<sub>2</sub>-methanation by catalytic conversion. *Environmental Earth Science*, 70, 3771–3778.
- Nahar, G.A. and Madhani, S.S. (2010). Thermodynamics of hydrogen production by the steam reforming of butanol: Analysis of inorganic gases and light hydrocarbons. *International Journal of Hydrogen Energy*, 35, 98–109. doi:10.1016/j.ijhydene.2009.10.013.
- Ocampo, F., Louis, B., and Roger, A.C. (2009). Methanation of carbon dioxide over nickel-based Ce<sub>0.72</sub>Zr<sub>0.28</sub>O<sub>2</sub> mixed oxide catalysts prepared by solgel method. *Applied Catalysis A*, 369, 90–96. doi:10.1016/j.apcata.2009.09.005.
- Ozkara-Aydinoglu, S. (2010). Thermodynamic equilibrium analysis of combined carbon dioxide reforming with steam reforming of methane to synthesis gas. *International Journal of Hydrogen Energy*, 35, 12821–12828.
- Rachow, F. (2017). *Prozessoptimierung für die Methanisierung von CO<sub>2</sub> – from a laboratory to an industrial scale*. urn:nbn:de:kobv:co1-opus4-42625, Doctoral thesis.
- Rakhi, Günther, V., and Mauss, F. (2023). Insights into dry reforming of methane over nickel catalyst using a thermodynamic model. *Reaction Kinetics, Mechanisms and Catalysis*, 136, 1197–1210. doi:10.1007/s11144-023-02426-8.
- Rakhi, Günther, V., Richter, J., and Mauss, F. (2022a). Steam reforming of methane over nickel catalyst using a one-dimensional model. *International Journal of Environmental Science*, 5(1), 1–32.
- Rakhi and Mauss, F. (2024). Thermodynamic model: Steam and oxidative reforming of methane over nickel catalyst. *Reaction Kinetics, Mechanisms and Catalysis*, 137, 791–812. doi:10.1007/s11144-024-02571-8.
- Rakhi, Shrestha, K.P., Günther, V., and Mauss, F. (2022b). Kinetically consistent detailed surface reaction mechanism for steam reforming of methane over nickel catalyst. *Reaction Kinetics, Mechanisms and Catalysis*, 135, 3059–3083. doi:10.1007/s11144-022-02314-7.
- Razzaq., Li, C., and Zhang, S. (2013). Coke oven gas: Availability, properties, purification, and utilization in China. *Fuel*, 113, 287–299.
- Rossi, C.C.R.S., Alonso, C.G., Antunes, O.A.C., Guirardello, R., and Cardozo-Filho, L. (2009). Thermodynamic analysis of steam reforming of ethanol and glycerine for hydrogen production. *International Journal of Hydrogen Energy*, 34, 323–332.
- Seemann, M.C., Schildhauser, T.J., and Biollaz, S.M.A. (2010). Fluidized bed methanation of wood-derived producer gas for the production of synthetic natural gas. *Industrial % Engineering Chemistry Research*, 49, 7034–7038.
- Ursua, A., Gandia, L.M., and Sanchis, P. (2012). Hydrogen Production From Water Electrolysis: Current Status and Future Trends. *Proceedings of the IEEE*, 100 (2), 410.

## Modelling and Simulation of CO<sub>2</sub> Capture through Aqueous Indirect Mineralization using CaO-containing By-products

Amirhossein Ghazi\*, Lars-Andre Tokheim\*\*

\* Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
(Tel: +47 96 72 84 65; e-mail: [Amirhossein.ghazi11@gmail.com](mailto:Amirhossein.ghazi11@gmail.com))

\*\* Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
(Tel: +47 35 57 51 71; e-mail: [Lars.A.Tokheim@usn.no](mailto:Lars.A.Tokheim@usn.no))

**Abstract:** The amount of CO<sub>2</sub> in the atmosphere is continuously increasing, resulting in climate change and global warming. Industrial processes contribute a substantial share in the amount of CO<sub>2</sub> released to the atmosphere. On the other hand, different types of wastes and by-products are being produced by different industries which are deemed pollutants and require energy and capital to be safely managed through a circular economy perspective. A solution to simultaneously tackle both the CO<sub>2</sub> emission and waste pollution problems would be of high value. CO<sub>2</sub> sequestration by mineralization of CaO-rich industrial wastes is one potential solution. In such a process, CO<sub>2</sub> reacts with the CaO in the waste and CaCO<sub>3</sub> is produced. This product is thermodynamically stable and has multiple uses. Many studies in the literature have reported use of various CaO-rich wastes to capture CO<sub>2</sub>, but they are mostly based on lab-scale experiments, and mostly the focus is on the chemistry of the suggested processes. Hence, there is a need to study the technical and economic feasibility of up-scaled industrial versions of such processes. In this study, four different aqueous indirect mineralization processes applying different chemicals, all with a relatively high performance documented from laboratory experiments, are scaled up to industrial size with a CO<sub>2</sub> capturing capacity of 400 t/y using an in-house-made process simulation tool. Furthermore, an economic analysis and environmental assessment are conducted for all processes, and the results are compared. Finally, parameters impacting the techno-economic feasibility of each process are evaluated through a sensitivity study. The results indicate that the potential of capturing CO<sub>2</sub> and producing CaCO<sub>3</sub> can be as high as 530 kg and 1200 kg per ton of the waste while the yearly energy consumption can be as low as 0.7 kWh per kilogram of captured CO<sub>2</sub>. The aqueous indirect mineralization of CO<sub>2</sub> can be profitable and the emitted CO<sub>2</sub> by the process can be so low as 6% of the captured amount.

**Keywords:** CO<sub>2</sub> sequestration, CO<sub>2</sub> capture, Mineralization, Aqueous indirect mineralization, Industrial wastes, Process simulation, Techno-economic analysis, Mass and energy balance, Environmental assessment

### 1. INTRODUCTION

Mineralization of silicates (Seifritz, 1990) is a direct process for carbonation and mineralization of natural alkaline minerals, such as olivine, serpentine and basalt, or industrial alkaline wastes like ashes and slags. This, due to the slow kinetics of mineral carbonation, needs a large energy demand to accelerate, or a long reaction time under ambient conditions. In addition, directly carbonating alkaline minerals usually produces low-quality products (Zhang et al., 2019).

On the other hand, there is the indirect carbonation process, which makes it possible to produce higher-value goods such as pure precipitated calcium carbonate (PCC) (Zhang et al., 2019). Indirect carbonation can be separated into aqueous indirect carbonation and stepwise gas-solid reactions. Prior to starting the mineral carbonation reaction, an indirect reaction must be used to extract alkaline earth metal ions from silicates

using the appropriate organic or inorganic acids or salts. Because the reactions usually occur in ambient conditions, this method may save energy in the mineral carbonation step. However, the extraction step (also known as the enrichment or separation step) may be uneconomical due to expensive reagents, reagent recovery, and energy consuming equipment (Zhang et al., 2019).

To determine which processes that have the lowest energy and cost intensity and the greatest amount of CO<sub>2</sub> captured, reaction modelling, process simulation, environmental impact assessment, energy analysis, and economic evaluation should be conducted. This will help determine the most promising options for scale-up (Zhang et al., 2020). Figure 1 shows a schematic overview of an indirect mineralization process.

Compared to solid phase mineralization methods, aqueous phase mineralization has demonstrated advantages in terms of

process operational parameters which are not under harsh conditions. Promising benefits of aqueous phase mineralization include the potential to speed up the process and the viability of large-scale implementation (Zhang et al., 2020).

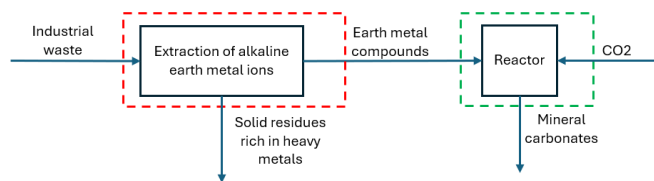


Fig. 1. A schematic overview of indirect mineralization process

The aqueous indirect mineralization of industrial wastes usually consists of two main steps. The first one is leaching, where the earth metal ions are separated from the waste. The second one is mineralization, where  $\text{CO}_2$  is introduced to the earth metal ions to be mineralized. Different wastes and methods based on literature have been used to perform the mentioned two steps. The waste can originate from various sources and the extracted earth metal could be different.

To extract the earth metal ions in the leaching process, an aqueous solution of acids or salts (reagents) with a specific concentration ( $C$ ) is prepared. Then, the solution is mixed with solids at a certain solid-to-liquid ratio ( $S:L$ ) at a specific leaching temperature ( $LT$ ) for a certain leaching time ( $TL$ ). The extraction efficiency ( $EF$ ) is defined as the part of the primary content of the earth metal ions which is extracted.

After the extraction, the  $\text{CO}_2$  is introduced to the solution, which is now rich in earth metal ions, at a specific mineralization temperature ( $MT$ ) for a certain mineralization time ( $TM$ ). In this way,  $\text{CO}_2$  is sequestered, and carbonates (mostly  $\text{CaCO}_3$ ) are produced. Appendix A gives details from the literature review done in this work.

In this study, the viability of aqueous indirect mineralization of  $\text{CO}_2$  by utilizing  $\text{CaO}$ -rich wastes for a plant in Norway with a  $\text{CO}_2$  capturing capacity of 400 t/y is investigated. The first step is, based on a literature study, to choose four processes with comparably good results. The second step is to model and simulate the chosen processes using an in-house-customized simulation tool (using MS Excel®) to solve mass and energy balances and do economic calculations. The third and final step is to define key performance indicators (KPI) and use these to compare the four processes in a sensitivity analysis. The work is based on a master's thesis work at USN (Ghazi, 2024).

## 2. PROCESS SELECTION

From the reviewed literature, the process using converter slag (CS) and  $\text{NH}_4\text{Cl}$  (Kodama et al., 2008) will be referred to as process 1, the process using recycled concrete fines (RCF) and  $\text{NH}_4\text{Cl}$  (Mehdizadeh et al., 2023) will be referred to as process 2, the process using blast furnace slag (BFS) and  $\text{HCl}$  (Liu et al., 2023) will be referred to as process 3, and the process using BFS and  $\text{CH}_3\text{COOH}$  (Teir et al., 2007) will be referred to as

process 4. These processes are chosen due to their comparably high  $\text{CaO}$  content in the wastes and high leaching efficiencies and because they use different reagents and apply different leaching and mineralization temperatures. For process 4, the  $MT$  and  $TM$  are not mentioned in the literature. Hence, the same values for  $LT$  and  $TL$  are assumed for these.

## 3. MODELLING AND SIMULATION

To be able to model and simulate the processes, first the processes are designed, and process flow diagrams (PFD) are prepared based on the details of the laboratory work in the papers. Then, using the PFDs, the mass and energy balance calculations can be conducted. Finally, based on mass and energy calculations, the economic and environmental assessment can be performed.

### 3.1 Process flow diagrams and descriptions

Figure 2 shows the PFD for process 1 and 4, and Table 1 provides the definitions of the streams in the PFD.

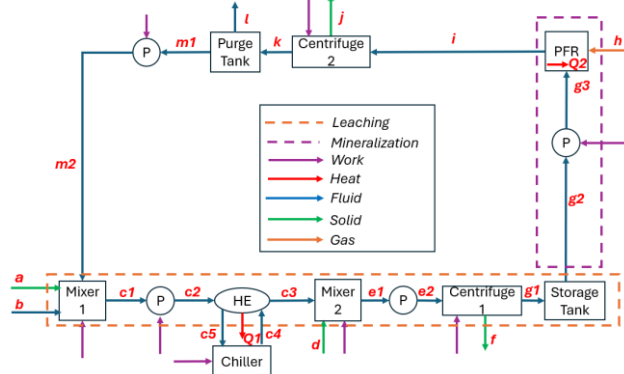


Fig. 2. Suggested PFD for process 1 and 4.

Table 1. Definition of streams in Fig. 2.

Stream	Definition
a	Reagent make-up
b	Water make-up
c1 and c2	Leaching solution
c4 and c5	Cooling water from chiller
c3	Cooled leaching solution
d	Industrial waste
e1 and e2	Leachate solution
f	Solid residues
g1, g2, and g3	Filtrate solution
h	$\text{CO}_2$
i	Mineralization solution
j	Precipitated $\text{CaCO}_3$
k	Recovered leaching solution
l	Purge
m1 and m2	Recovered leaching solution
Q1	Heat extracted from the process
Q2	Heat generated in the reactor

The water (b) and reagent (a and m2) are mixed in mixer 1, and the leaching solution (c1 and c2) is then pumped (P) through the heat exchanger (HE), where the stream is cooled down (Q1) to the leaching and mineralization temperature

using cooling water (c4 and c5) from a chiller. The leaching and mineralization temperatures are the same in process 1 and 4. After that, the industrial waste (d) is added to the leaching solution (c3) in mixer 2, where the leaching process happens in the leaching time. After leaching, the leachate (e1 and e2) is pumped (P) into centrifuge 1, where the solid residues (f) are separated, and the filtrate solution (rich in Ca) accumulates in the storage tank. Then the filtrate solution (g1, g2 and g3) is pumped (P) to the plug flow reactor (PFR), where CO<sub>2</sub> (h) is introduced to be mineralized, generating heat (Q2) due to exothermic nature of the reaction. After the PFR, the mineralization solution (i), containing CaCO<sub>3</sub>, passes through centrifuge 2, where the precipitated CaCO<sub>3</sub> (j) is separated from the stream. The recovered reagent (k) then accumulates in the purge tank, and a part of it (l) is purged out of the process to prevent accumulation of heavy metals and undesired materials. Finally, the recovered reagent (m1) is pumped (P) to mixer 1 to repeat the cycle.

Figure 3 depicts the suggested PFD for processes 2 and 3, and Table 2 provides the definition of streams that are different from those in Fig. 2. Processes 2 and 3 are quite similar to processes 1 and 4, but the operational temperatures are different. The leaching temperature is much higher than the mineralization temperature in processes 2 and 3. This requires the heat to be added to the stream before leaching to increase the temperature to the leaching temperature. Meanwhile, the heat must be removed from the stream before mineralization to reach the mineralization temperature. To reduce the heating and cooling demands, a heat recovery line using heat exchangers 1 (HE1) and 3 (HE3) is established to recover a part of the heat. The rest of the heat will be added to and will be removed from the stream using heat exchangers 2 (HE2) and 4 (HE4), respectively.

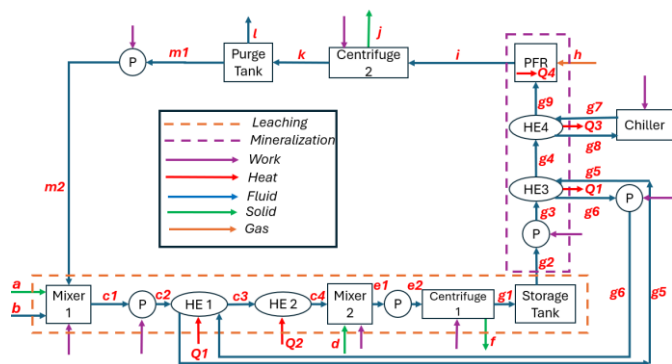


Fig. 3. Suggested PFD for process 2 and 3.

Table 2. Definition of streams in Fig. 3 that are new or different from those in Fig. 2.

Stream	Definition
c3 and c4	Cooled leaching solution
g5 and g6	Heat recovery line
g4 and g9	Cooled filtrate solution
g7 and g8	Cooling water from chiller
Q1	Recovered heat in the process
Q2	Heat added to the process
Q3	Heat extracted from the process
Q4	Heat generated in the reactor

### 3.2 Mass balance calculations

In the mineralization process, 1 mol of CO<sub>2</sub> reacts with 1 mol of CaO to produce 1 mol of CaCO<sub>3</sub>. Laboratory data from the literature, such as calcium content in the waste, reagent concentration, and solid-to-liquid ratio, are used to calculate flow rates, extraction efficiency and plant capacity for all the processes.

### 3.3 Energy balance calculations

The energy consumption of the four chosen processes is the sum of mechanical work (in pumps, agitators, mixers and centrifuges) and thermal energy related to heating and cooling demands. By knowing the mass flow rates, the power of agitators and centrifuges, the enthalpy of mineralization, and the specific heat capacity of the streams, the yearly energy consumption is calculated. Since the heat transfer area of the mixers, tanks, and pipes is not specified at this level, the heat loss from these equipment units is not calculated. Instead, the system is treated as well insulated, and heat loss is neglected.

### 3.4 Economic calculations

The economic calculations in this study are limited to material (reagent and process water) costs, energy cost, and revenue from sales of CaCO<sub>3</sub>. The mass flow rates of the required materials, the energy consumption, the energy price in Norway, and the reagent price (assuming three different origins; East Asia, the European Union, and the US) are used as inputs, see Appendix B for details. Then the revenue from sales of CaCO<sub>3</sub>, the yearly cost and the revenue of the processes are calculated. The economic assessment in terms of checking the profitability is then conducted.

### 3.5 CO<sub>2</sub> footprint

Although the CO<sub>2</sub> footprint of a process depends on numerous factors and aspects, in this study, the CO<sub>2</sub> footprint calculated based on the electrical energy consumption, assumed to be only electricity from grid. The CO<sub>2</sub> emission factors of electricity in the country where the plants are located are given in Appendix B.

### 3.6 Simulation tool and simulation settings

Laboratory data from the literature are used as inputs to the mass and energy balance equations. Assumptions are made where sufficient data are not available. Then a simulation tool is implemented in MS Excel<sup>®</sup> to calculate and scale up the unknowns, such as mass flow rates, energy consumption, etc. This tool is also used to perform the sensitivity analysis. The simulation settings are shown in appendix B.

## 4. KEY PERFORMANCE INDICATORS

To illustrate and compare the performance of the processes, eight KPIs are defined in Table 3 based on the results of mass, energy, economic, and environmental calculations.

KPIs 1 to 3 are mass-based and can be used to compare how much CO<sub>2</sub> that can be captured, how much waste that can be handled and how much make-up reagents that are required. KPI1 and KPI2 should be high, whereas KPI3 should be low.



Table 3. KPIs and their definitions.

KPI	Definition
KPI1	Mass of captured CO <sub>2</sub> per mass of waste
KPI2	Mass of produced CaCO <sub>3</sub> per mass of waste
KPI3	Mass of make-up reagent per mass of captured CO <sub>2</sub>
KPI4	Energy consumption per mass of captured CO <sub>2</sub>
KPI5	Mass of produced CO <sub>2</sub> per mass of captured CO <sub>2</sub>
KPI6	Yearly profit per mass of captured CO <sub>2</sub> if the reagent is supplied from East Asia
KPI7	Yearly profit per mass of captured CO <sub>2</sub> if the reagent is supplied from the EU
KPI8	Yearly profit per mass of captured CO <sub>2</sub> if the reagent is supplied from the US

KPI4 is energy-based and can be used to compare how much energy that must be supplied to the different processes. This value should be low.

KPI5 is also mass-based, but the main aim of this parameter is to show the environmental impact of the processes; the lower the better. KPIs 6 to 8 are used to compare the economic performance of the processes, the higher values the better.

## 5. RESULTS AND DISCUSSION

### 5.1 Mass balance results

The calculated mass flow results for the streams in the four processes are shown in Table 4.

Although the CO<sub>2</sub> and CaCO<sub>3</sub> streams (h and j) are the same in all four processes, other streams are different. Taking g1 as for comparison, process 4 has the highest flow rate of 159 kg/min, which is the double of the flow rate in process 1 with 81 kg/min. The lowest flow rate is for process 2 with 24 kg/min. This will result in a higher energy consumption of the pumps and centrifuges. Considering stream a, process 4 has the highest consumption of reagent followed by processes 3, 1, and 2. The higher the reagent consumption, the higher the process costs.

### 5.2 Energy balance

The energy consumption results for the four processes are shown in Table 5. Compared to other rotary equipment, the pump energy consumption is negligible in all four processes. For process 4, the energy consumption of the centrifuges is almost twice as high as in process 1. This is due to a higher flow rate in process 4. Heating and cooling, with more than 70 % of the whole energy consumption, are the main role-players in the energy consumption of processes 2 and 3.

Table 4. Mass flow rates (kg/min) (– = Not applicable)

Stream	Process			
	1	2	3	4
a	0.6	0.4	0.8	2.4
b	8	2.3	3.7	16
c1 and c2	80	23	37	158
c3	80	23	37	158
c4	22	23	37	67
c5	22	–	–	67
d	5.0	2.3	3.0	2.7
e1 and e2	85	25	40	161
f	4.3	1.5	2.2	1.9
g1, g2, and g3	81	24	38	159
g5 and g6	–	25	39	–
g4 and g9	–	24	38	–
g7 and g8	–	123	151	–
h	0.8	0.8	0.8	0.8
i	82	25	39	160
j	1.7	1.7	1.7	1.7
k	80	23	37	158
l	8	2.3	3.7	16
m1 and m2	72	21	33	142

Table 5. Energy streams (kW) (– = Not applicable)

Utility	Process			
	1	2	3	4
Agitation	15	15	15	15
Pumps	0.06	0.02	0.03	0.12
Centrifuges	14	4	6	26
Heating	–	31	49	–
Cooling (Chiller)	10	17	20	10
Total energy	39	67	91	51

### 5.3 KPI results

The mass and energy balance results are used, along with cost and environmental calculations, in the calculation of KPIs. Figure 4 shows KPIs 1 to 3 and 5.

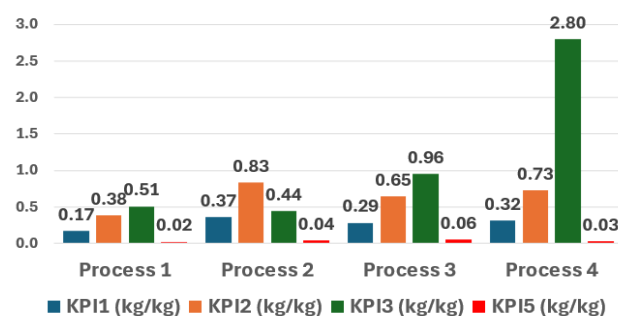


Fig. 4. KPIs 1, 2, 3, and 5

As shown in Fig. 4, process 2 has the highest capacity of capturing CO<sub>2</sub> per mass of used waste (KPI1) with 0.37 kg/kg. Process 4 and 3 follow with values of 0.32 kg/kg and 0.29 kg/kg, respectively. Process 1 has the lowest potential of capturing CO<sub>2</sub> with only 0.17 kg/kg.

The trend for KPI2 is similar to that of KPI1; process 2 has the highest amount of produced CaCO<sub>3</sub> per mass of waste (0.83 kg/kg) followed by processes 4, 3, and 1 (0.73, 0.65, and 0.38 kg/kg, respectively).

The reasons for the higher capacity for CO<sub>2</sub> capture and CaCO<sub>3</sub> production of process 2 is the higher content of CaO in the waste, a higher solid-to-liquid ratio, and a higher extraction efficiency. Although the waste CaO content of process 1 is almost the same as in processes 3 and 4, the lower extraction efficiency results in a lower capture capacity.

When it comes to KPI3, the lower mass of make-up reagent per mass of captured CO<sub>2</sub>, the better in terms of economic assessment. Hence, process 2 is still the better process, with a reagent consumption of 0.44 kg/kg, followed by process 1 with 0.51 kg/kg and process 3 with 0.96 kg/kg. Process 4 has a staggeringly high reagent consumption of 2.8 kg/kg. Although the concentration of reagent for this process is not high compared to the other three processes, the comparably lower solid-to-liquid ratio results in a higher volume of leaching solution which increases the reagent consumption.

KPI5 is an indicator of the emitted mass of CO<sub>2</sub> per mass of captured CO<sub>2</sub>. This KPI is of high importance and should be calculated in the early stages of a scale-up project since it can be a showstopper. If KPI5 shows 1 kg/kg or more, it means that the plant is emitting more CO<sub>2</sub> than it captures. This depends on the CO<sub>2</sub> emission factor of electricity generation. Figure 4 shows KPI5 for power production in Norway, which has a very low CO<sub>2</sub> emission factor. Accordingly, all four processes emit negligibly low amounts of CO<sub>2</sub> per unit of captured CO<sub>2</sub> with 0.02 kg/kg, 0.03 kg/kg, 0.04 kg/kg, and 0.06 kg/kg for processes 1, 4, 2, and 3, respectively.

Based on Fig. 5, processes 1 and 4 are the most efficient ones in terms of energy consumption per mass of captured CO<sub>2</sub> with 0.86 kWh/kg and 1.13 kWh/kg followed by process 2 (1.47 kWh/kg) and process 3 (1.98 kWh/kg).

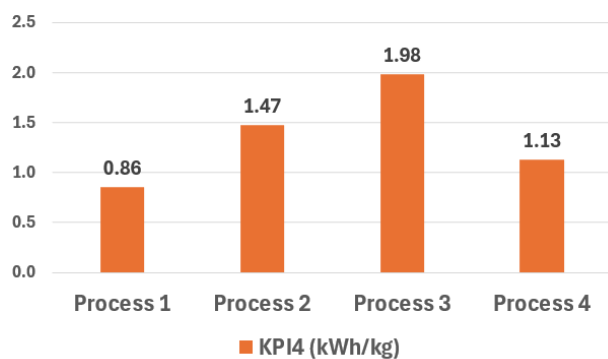


Fig. 5. KPI4 for all processes

Interestingly, despite higher mass flow rates for processes 1 and 4, which result in a higher energy consumption of pumps and centrifuges, the total energy consumption of these two processes is lower than processes 2 and 3. This is due to different operational temperatures of leaching and mineralization in processes 2 and 3, which require heating and cooling at the same time, resulting in a higher overall energy consumption for these two processes.

The mass of captured CO<sub>2</sub> and produced CaCO<sub>3</sub> are the same for all four processes. Therefore, the revenue from sales of CaCO<sub>3</sub> is also the same for all. Figure 6 shows the yearly profit of all four processes per mass of captured CO<sub>2</sub> if the reagents are supplied from East Asia (KPI6), the EU (KPI7), and the US (KPI8).

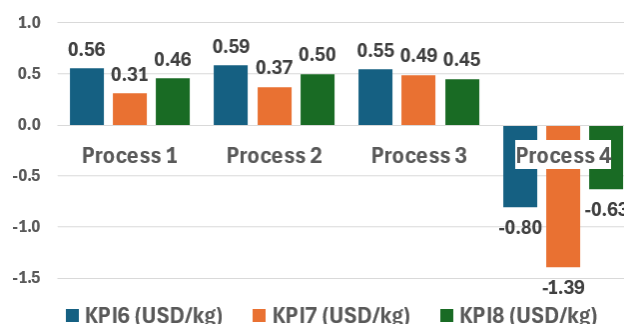


Fig. 6. KPIs 6 to 8, economic assessment

Process 4 shows negative numbers for all three KPIs, meaning that the costs are higher than the revenue, and the process is not profitable. For KPIs 6 to 8, process 1, 2, and 3 are all profitable. The profitability is highest if the reagent is supplied from East Asia (0.59 USD/kg for process 2), and lowest if it is supplied from the EU (0.31 USD/kg for process 1). This can be explained by different chemical prices (see Appendix B).

#### 5.4 Sensitivity analysis

The reactants in the mineralization reaction are CO<sub>2</sub> and CaO. Since the amount of CO<sub>2</sub> to be captured is a design basis value (400 kt/y) in this study, the amount of CaO plays the main role in the mass and energy balance calculations. The amount of CaO, or better Ca<sup>2+</sup> ions in the reaction, is dependent on the extraction efficiency at a given solid-to-liquid ratio. Due to the importance of the extraction efficiency, a sensitivity analysis was conducted to find the impact of this parameter on KPIs 1-4. The extraction efficiency was varied from 8 to 98 %, and the results are shown in Figs. 7 to 10. (The reported efficiencies from the laboratory work are indicated with \* on the horizontal axis in the figures.)

As seen from all four figures, KPIs 1 and 2 have a linearly increasing trend with an increase in the extraction efficiency. This means that the extraction efficiency has a direct impact on the capturing capacity of each process. This trend is reasonable since the extracted Ca is directly reacting with CO<sub>2</sub>. KPIs 3 and 4, on the other hand, show an exponential decay behavior with an increase in the extraction efficiency. The non-linear behavior is seen because there is more than one role-player in the calculation of make-up reagent and the

energy consumption, and these two KPIs are sum of different parameters. The decreasing trend occurs because the reagent and energy consumption are reduced when the extraction efficiency is increased.

Looking into KPIs 1 and 2 for the whole range of extraction efficiencies, process 2 is the more promising one. Doing the same for KPIs 3 and 4, process 1 is the preferred process.

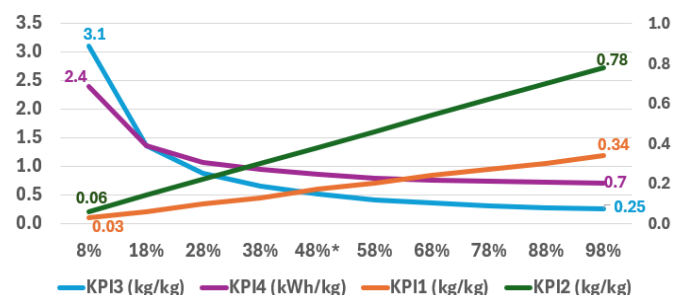


Fig. 7. Sensitivity analysis of process 1

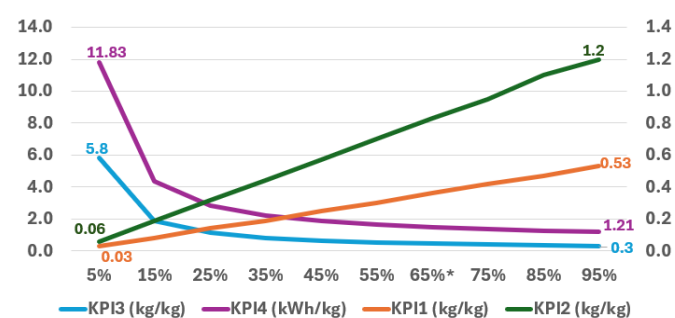


Fig. 8. Sensitivity analysis of process 2

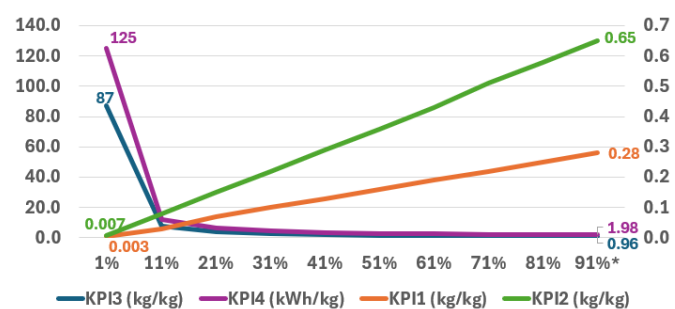


Fig. 9. Sensitivity analysis of process 3

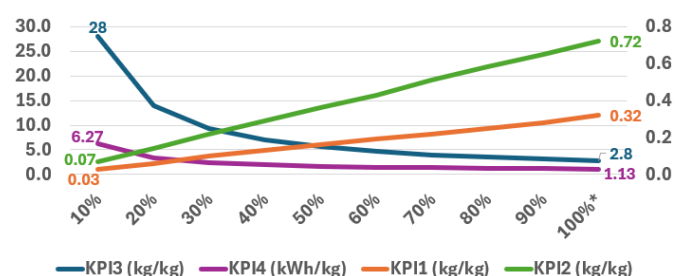


Fig. 10. Sensitivity analysis of process 4

In Fig. 11, the processes are compared at an extraction efficiency of 50%. While process 2 has a better performance for KPIs 1 and 2, process 1 is better for KPIs 3 and 4.

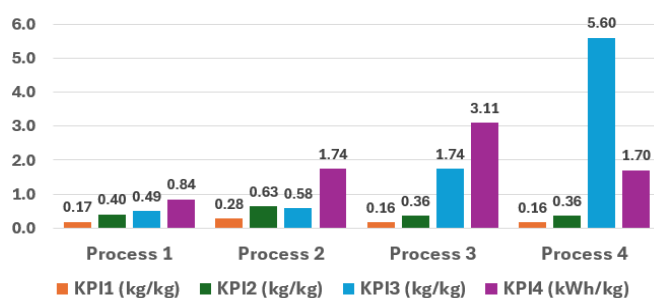


Fig. 11. KPIs 1 to 4 at EF=50% for all processes

## 6. CONCLUSIONS

Many methods to mineralize CO<sub>2</sub> by industrial CaO-rich waste streams have been described in the literature, mainly based on laboratory experiments. In this study, four different processes were selected, scaled up to industrial scale and compared.

The results indicate that a process using recycled concrete fines (RCF) and NH<sub>4</sub>Cl as a solvent (process 2) has the highest specific CO<sub>2</sub> capture (KPI1) and CaCO<sub>3</sub> production (KPI2) and the lowest chemical make-up requirement (KPI3).

However, a process using converter slag (CS) and NH<sub>4</sub>Cl as solvent (process 1) has the lowest specific energy consumption (KPI4).

The two other processes considered were using blast furnace slag (BFS) and either HCl (process 3) or CH<sub>3</sub>COOH as a solvent (process 4). All four processes will generate less than 6% of the CO<sub>2</sub> that is captured (KPI5), meaning that they are all viable from net CO<sub>2</sub> capture point of view.

The analysis indicated that process 4 is not economically viable, not matter if the chemicals are supplied from East Asia (KPI6), the EU (KPI7) or the US (KPI8). Processes 1-3 all appear to be economically viable, irrespective of where the chemical reagent is purchased.

This study indicates that CO<sub>2</sub> mineralization using industrial by-products may be part of a solution going towards a more circular economy. However, as only operational costs were included in the present analysis, the impact of including investment costs could be investigated in a future study.

## ACKNOWLEDGEMENTS

We kindly thank Mr. Zahir Barahmand from University of South-Eastern Norway for his advice on the literature study, Dr. Eugenia Marinou, and Mr. Ole Jakob Sandal from Caux AS for valuable discussions on mineralization and making the opportunity for this study available.

## REFERENCES

- businessanalytiq, Ammonium chloride price index. URL <https://businessanalytiq.com/procurementanalytics/index/ammonium-chloride-price-index/> (accessed 4.24.24).
- Carbon intensity of electricity generation [WWW Document], Our World in Data. URL <https://ourworldindata.org/grapher/carbon-intensity-electricity> (accessed 4.25.24).

- Electricity prices [WWW Document], SSB. URL <https://www.ssb.no/en/energi-og-industri/energi/statistikk/elektrisitetspriser> (accessed 4.24.24).
- Gao, Y., Jin, X., Teng, L., Rohani, S., He, M., Li, J., Ren, S., Liu, Q., Huang, J., Duan, H., Xin, Y., and Liu, W. (2023). CO<sub>2</sub> mineral sequestration and nickel recovery from laterite ore by using waste copperas. *Fuel*, 331, 125750. doi: 10.1016/j.fuel.2022.125750
- Ghazi, A., (2024). CO<sub>2</sub> capture through mineralization of CaO-containing by-products. University of South-Eastern Norway, Porsgrunn, Norway.
- Ho, H.-J., Iizuka, A., Shibata, E., and Ojumu, T. (2022). Circular indirect carbonation of coal fly ash for carbon dioxide capture and utilization. *Journal of Environmental Chemical Engineering*, 10, 108269. doi: 10.1016/j.jece.2022.108269
- Ji, L., Zheng, X., Ren, Y., Wang, Yikun, Wang, Yan, and Yan, S., (2024). CO<sub>2</sub> sequestration and recovery of high-purity CaCO<sub>3</sub> from bottom ash of masson pine combustion using a multifunctional reagent—amino acid. *Separation and Purification Technology* 329, 125171. doi: 10.1016/j.seppur.2023.125171
- Kashefi, K., Pardakhti, A., Shafiepour, M., and Hemmati, A., (2020). Process optimization for integrated mineralization of carbon dioxide and metal recovery of red mud. *Journal of Environmental Chemical Engineering*, 8, 103638. doi:10.1016/j.jece.2019.103638
- Kim, M.-J., Jeon, J., (2020). Effects of Ca-ligand stability constant and chelating agent concentration on the CO<sub>2</sub> storage using paper sludge ash and chelating agent. *Journal of CO<sub>2</sub> Utilization*, b 40, 101202. doi: 10.1016/j.jcou.2020.101202
- Kodama, S., Nishimoto, T., Yamamoto, N., Yogo, K., and Yamada, K., (2008). Development of a new pH-swing CO<sub>2</sub> mineralization process with a recyclable reaction solution. *Energy*, 33, 776–784. doi: 10.1016/j.energy.2008.01.005
- Lin, Y., Yan, B., Mitas, B., Li, C., Fabritius, T., and Shu, Q., (2024). Calcium carbonate synthesis from Kambara reactor desulphurization slag via indirect carbonation for CO<sub>2</sub> capture and utilization. *Journal of Environmental Management*, 351, 119773. doi: 10.1016/j.jenvman.2023.119773
- Liu, L., Gan, M., Fan, X., Sun, Z., Wei, J., Li, J., and Ji, Z., (2023). Synthesis of high-value CaCO<sub>3</sub> via indirect CO<sub>2</sub> fixation utilized blast furnace slag. *Journal of Environmental Chemical Engineering*, 11, 110655. doi: 10.1016/j.jece.2023.110655
- Mehdizadeh, H., Mo, K.H., and Ling, T.-C., (2023). O<sub>2</sub>-fixing and recovery of high-purity vaterite CaCO<sub>3</sub> from recycled concrete fines. *Resources, Conservation and Recycling*, 188, 106695. doi: 10.1016/j.resconrec.2022.106695
- Recycling, 188, 106695. doi: 10.1016/j.resconrec.2022.106695
- Proaño, L., Sarmiento, A.T., Figueredo, M., and Cobo, M. (2020). Techno-economic evaluation of indirect carbonation for CO<sub>2</sub> emissions capture in cement industry: A system dynamics approach. *Journal of Cleaner Production*, 263, 121457. doi: 10.1016/j.jclepro.2020.121457
- Seifritz, W. (1990). CO<sub>2</sub> disposal by means of silicates. *Nature*, 345, 486–486. doi: 10.1038/345486b0
- Song, Q., Guo, M.-Z., and Ling, T.-C. (2024). Synthesis of High-Purity and Stable Vaterite Via Leaching-Carbonation of Basic Oxygen Furnace Slag. *ACS Sustainable Chemistry and Engineering*, 12, 4081–4091. doi: 10.1021/acssuschemeng.3c07375
- Szepessy, S. and Thorwid, P. (2018). Low Energy Consumption of High Speed Centrifuges. *Chemical Engineering & Technology*, 41. doi: 10.1002/ceat.201800292
- Teir, S., Eloneva, S., Fogelholm, C.-J., and Zevenhoven, R. (2007). Dissolution of steelmaking slags in acetic acid for precipitated calcium carbonate production. *Energy, ECOS 05. 18th International Conference on Efficiency, Cost, Optimization, Simulation, and Environmental Impact of Energy Systems* 32, 528–539. doi: 10.1016/j.energy.2006.06.023
- Wang, N., Feng, Y., and Guo, X. (2020). Atomistic mechanisms study of the carbonation reaction of CaO for high-temperature CO<sub>2</sub> capture. *Applied Surface Science*, 532, 147425. doi:10.1016/j.apsusc.2020.147425
- Wu, L., Li, H., Mei, H., Rao, L., Xia, Y., and Dong, Y. (2023). A novel approach to accelerate carbon dioxide sequestration of ladle furnace slag using sodium bicarbonate solution. *Minerals Engineering* 204, 108374. doi: 10.1016/j.mineng.2023.108374
- Yiyimechanical.com [WWW Document], Yiyi Machinery. URL <https://www.yiyimechanical.com/shop/mixing-tank> (accessed 5.18.24).
- Zhang, N., Chai, Y.E., Santos, R.M., and Šiller, L. (2020). Advances in process development of aqueous CO<sub>2</sub> mineralisation towards scalability. *Journal of Environmental Chemical Engineering*, 8, 104453. doi: 10.1016/j.jece.2020.104453
- Zhang, N., Santos, R.M., Smith, S.M., and Šiller, L. (2019). Acceleration of CO<sub>2</sub> mineralisation of alkaline brines with nickel nanoparticles catalysts in continuous tubular reactor. *Chemical Engineering Journal, ISCRE 25 Special Issue: Bridging Science and Technology*, 377, 120479. doi: 10.1016/j.cej.2018.11.177

## Appendix A: Literature review details

Waste Type	Metal	Content (%)	Reagent	C (mol/L)	S:L (g/L)	LT (°C)	TL (min)	EF (%)	MT (°C)	TM (min)	Reference
Biomass ash	Ca	CaO: 36.7	Glycine	2	300	25	60	19.9	25	60	Ji et al., 2024
Kambara reactor desulphurization slag	Ca	CaO: 72.04	NH <sub>4</sub> Cl	0.5	10	25	120	89.76	40	120	Lin et al., 2024
Recycled concrete fines (RCF)	Ca	CaO: 71.11	NH <sub>4</sub> Cl	2	100	85	60	65.7	25	30	Mehdizadeh et al., 2023
Converter slag (CS)	Ca	CaO: 44.5	NH <sub>4</sub> Cl	1	63	80	60	48.1	80	120	Kodama et al., 2008
Blast furnace slag (BFS)	Ca	CaO: 40	HCl	4	80	80	120	91	20	15	Liu et al., 2023
Blast furnace slag (BFS)	Ca	CaO: 40.6	CH <sub>3</sub> COOH	1.7	16.8	25	120	100	N/A	N/A	Teir et al., 2007
Laterite Ore Copperas	Mg	MgO: 37.11	Distilled water	N/A	100	80	60	86	30	120	Gao et al., 2023
Coal fly ash	Ca	Ca: 3.44	HNO <sub>3</sub>	0.3	30	25	150	72.6	25	30	Ho et al., 2022
Ladle furnace slag	Ca	CaO: 46.73	NaHSO <sub>4</sub>	0.5	143	25	90	N/A	25	120	Wu et al., 2023
Basic oxygen furnace slag (BOF)	Ca	CaO: 40	NH <sub>4</sub> Cl	2	50	80	10	60.3	25	10	Song et al., 2024
Red mud	Ca	CaO: 14.21	HCl	1	70	80	120	85	25	N/A	Kashefi et al., 2020
Paper sludge ash (PSA)	Ca	Ca: 67.2	Citrate	0.1	20	25	60	23.5	25	N/A	Kim and Jeon, 2020



## Appendix B: Simulation inputs

Input	Process 1	Process 2	Process 3	Process 4	Reference
Captured CO <sub>2</sub> (t/y)	400	400	400	400	Assumed
CaO content of waste (%)	44.5	71.11	40	40.6	Kodama et al., 2008, Mehdizadeh et al., 2023 Liu et al., 2023, Teir et al., 2007
S/L (g/L)	63	100	80	16.8	Kodama et al., 2008, Mehdizadeh et al., 2023 Liu et al., 2023, Teir et al., 2007
EF (%)	48.1	65.7	91	100	Kodama et al., 2008, Mehdizadeh et al., 2023 Liu et al., 2023, Teir et al., 2007
Purge fraction (%)	10	10	10	10	Assumed
LT (°C)	80	85	80	25	Kodama et al., 2008, Mehdizadeh et al., 2023 (Liu et al., 2023) (Teir et al., 2007)
MT (°C)	80	25	20	25	Kodama et al., 2008, Mehdizadeh et al., 2023 Liu et al., 2023, Teir et al., 2007
Specific heat capacity of streams (kJ/kgK)	4.2	4.2	4.2	4.2	Assumed
Enthalpy of mineralization (kJ/mol)	178	178	178	178	Wang et al., 2020
Centrifuge energy consumption (kWh/m <sup>3</sup> )	1.5	1.5	1.5	1.5	Szepessy and Thorwid, 2018
Density of streams (kg/m <sup>3</sup> )	1000	1000	1000	1000	Assumed
Agitator power (kW)	7.5	7.5	7.5	7.5	“yiyimechanical.com,” n.d.
Head loss (m)	5	5	5	5	Assumed
CO <sub>2</sub> production in Norway (kg/kWh)	0.029	0.029	0.029	0.029	“Carbon intensity of electricity generation,” n.d.
Energy price in Norway (USD/MWh)	42.63	42.63	42.63	42.63	“Electricity prices,” n.d.
Reagent price in East Asia (USD/kg)	0.08	0.08	0.04	0.47	businessanalytiq, n.d.
Reagent price in the EU (USD/kg)	0.56	0.56	0.1	0.68	businessanalytiq, n.d.
Reagent price in the US (USD/kg)	0.27	0.27	0.14	0.41	businessanalytiq, n.d.
Process water cost (USD/kg)	0.00755	0.00755	0.00755	0.00755	Proaño et al., 2020
Price of CaCO <sub>3</sub> (USD/kg)	0.31	0.31	0.31	0.31	Proaño et al., 2020

# Phase Transformations in Steelmaking Slags: A Thermodynamic Approach

Tuomas Alatarvas, Rita Kallio, Eetu-Pekka Heikkinen, Qifeng Shu

*Process Metallurgy Research Unit, Centre for Advanced Steel Research,  
University of Oulu, P.O. Box 4300, FI-90014 Oulu, Finland*

**Abstract:** In addition to solidification, steelmaking slags may undergo phase transformations in solid state during their cooling process. The mineralogy of these oxide slags is significantly influenced by the chemical composition and cooling rate. For the phases forming, two distinct solidification modes can be assumed, depending on the cooling rate: equilibrium cooling and Scheil–Gulliver cooling. Characterization methods, such as scanning electron microscopy (SEM) and electron probe microanalyzer (EPMA) allow analyzing the elemental composition of individual phases. Here, computational thermodynamics were applied in phase identification of crystallized electric arc furnace (EAF) slags. FactSage 8.3 thermodynamic calculation software was used to estimate the composition of stable phases as a function of temperature. Solid solutions with varying compositions were considered in this study. The calculation results from two solidification modes, i.e., equilibrium cooling and Scheil–Gulliver cooling, were saved in Excel spreadsheets. A MATLAB script was developed to go through the results and find the phase with a composition closest to the input values. For both solidification modes, the composition and temperature best fitting the input analysis was determined. The input is the elemental composition of the phase of interest, acquired using EPMA. After the data processing, the results are visualized in graphs, illustrating the analyzed and estimated compositions of the identified solid solution phase and its occurrence temperature.

**Keywords:** computational thermodynamics, mineralogy, slag, solidification

## 1. INTRODUCTION

Various types of ferrous slags are generated as by-products of steel industry. During the melting and reduction of iron ore to produce pig iron in a blast furnace, the most voluminous of these, Blast Furnace Slag (BFS) is produced, whereas steel slag is generated either by converting pig iron to steel in a Basic Oxygen Furnace (BOF) or from the melting of scrap to produce steel in Electric Arc Furnace (EAF). According to EUROSLAG data (European Association representing metallurgical slag producers and processors), 37 million tons of slag were produced in 2018, of which 20.7 million tons are BFS and 16.3 million tons are various types of steelmaking slags. EAF slags from carbon steel production represent 26.2% of the steelmaking slags (Euroslag Statistics, 2018). The main applications of steelmaking slags include production of aggregates for road construction, cement or concrete addition, fertilizer, hydraulic engineering, and internal use for metallurgical purposes (Bru et al., 2021).

The physical and chemical characteristics of the EAF slags are dictated by the raw materials used, the constraints of the production process and the cooling method. The main oxide components of EAF slags are iron oxides ( $\text{FeO}_n$ ), lime ( $\text{CaO}$ ), silica ( $\text{SiO}_2$ ), magnesia ( $\text{MgO}$ ) and alumina ( $\text{Al}_2\text{O}_3$ ), in addition to minor components such as chromium, manganese and phosphorus oxides (Bru et al., 2021). These components form minerals such as dicalciumsilicates (commonly as monoclinic  $\beta$  form, larnite) monoxide solid solution ( $\text{Mg,Fe}$ )O, ( $\text{Mg,Cr,Fe,Al}$ )-spinel, calcium ferrites, calcium

aluminates and melilite group members such as gehlenite (Mombelli et al., 2014, Li et al., 2022).

In addition to solidification, steelmaking slags may undergo phase transformations in solid state during their cooling process. The stability of phases in the oxide system can be assessed with computational thermodynamics software such as FactSage (Bale et al. 2002). For a given total composition of the system, the equilibrium composition can be calculated, for instance as a function of temperature. For the phases forming during solidification of molten slags, two distinct solidification modes can be assumed, depending on the cooling rate: equilibrium cooling and Scheil–Gulliver cooling. Equilibrium cooling assumes that the whole system is in equilibrium during the cooling process. On the other hand, the Scheil–Gulliver model, as described by Durinck et al. (2007), assumes no diffusion in the solid phases, infinitely rapid diffusion in the liquid phase, and local equilibrium at the solid/liquid interface. This means that the solidified phases are considered inert, and the final phase structure is reached when the system has completely solidified. Andersson et al. (2024) utilized FactSage for assessing the Scheil–Gulliver solidification sequence of EAF slags with varying composition. The authors reported that for the unmodified slag, with basicity ( $\text{CaO/SiO}_2$  ratio) of 2.0, the primary crystallizing phase is monoxide. Modification of the slag by decreasing the basicity to 1.5 and 1.0 promoted the formation of spinel type phases.

The current study presents a method to compare the thermodynamic calculation results with the elemental analysis obtained from steelmaking slags. Interpreting the results

allows the estimation of the formation sequence and temperature for the identified phases.

## 2. MATERIALS AND METHODS

### 2.1 Investigated materials

The samples (Slag#1 and Slag#2) provided for this study were produced in two scrap-based EAF steelmaking sites, which manufacture advanced steels for the automotive and engineering industries. Slags are air-cooled with moderate irrigation and processed according to conventional slag-handling procedures including crushing, sieving and magnetic separation stages.

Representative samples from the two sites were delivered to Raahe Research Centre, which conducted the bulk XRF analyses using Malvern Panalytical equipment with Rh X-ray tube and MsDecipher 2.4.17 calculation program. The composition data of the studied slags is presented in Table 1.

Table 1. Chemical composition of the studied slags, in wt.%.

	Slag#1	Slag#2
SiO <sub>2</sub>	12.13	12.96
Al <sub>2</sub> O <sub>3</sub>	7.30	8.81
MgO	3.80	8.71
CaO	39.58	32.41
FeO	29.53	30.21
MnO	5.44	4.96
Cr <sub>2</sub> O <sub>3</sub>	2.22	1.93

### 2.2 Electron probe microanalyzer

Polished blocks (Ø 25 mm) of the slag samples were prepared for electron probe microanalyzer (EPMA) analysis, which was conducted at the Centre for Material Analysis, University of Oulu. The blocks were coated with carbon prior to analyses. A JEOL JXA-8530FPlus electron probe microanalyzer (JEOL Ltd., Tokyo, Japan) equipped with X-ray Wavelength Dispersive Spectrometer (WDS) was employed to characterize mineral chemical compositions in slags, with the analytical conditions including an accelerating voltage of 15 kV, a beam current of 15 nA, and a beam diameter of 1–10 µm. The peak and background counting times were set at 10 s and 5 s, respectively, for all components. The matrix correction with the ZAF method (atomic number—absorption—fluorescence) was applied. Backscattered electron (BSE) detector was used for imaging of the samples.

### 2.3 Computational thermodynamics

Computational thermodynamics software FactSage 8.3 and its *FToxid* database were used to estimate the stable phases in the studied materials. In addition to liquid oxide slag, 29 solid solutions from the *FToxid* database were selected for the calculations. Table 2 lists the selected solutions with explanatory composition or mineralogical name. Stoichiometric, pure compounds were not included in the system. For the studied materials, the phase fractions and their compositions were calculated as a function of temperature. In the equilibrium solidification mode, phase stability was calculated in the temperature range of 1800–300 °C. For the

Scheil–Gulliver solidification mode, the starting temperature was defined as 1800 °C, where both studied compositions are fully liquid. Here, the cooling step was defined as 25 °C. The results were saved as Excel spreadsheet files.

Table 2. Selected solutions for the calculations.

Solution name	Example composition or mineralogical name
A-slag	Liquid oxide slag
B-spinel	Magnetite (Fe <sub>3</sub> O <sub>4</sub> ), Manganoferrite (MnFe <sub>2</sub> O <sub>4</sub> ), Manganochromite (MnCr <sub>2</sub> O <sub>4</sub> ), Chromite (FeCr <sub>2</sub> O <sub>4</sub> ), Galaxite (MnAl <sub>2</sub> O <sub>4</sub> ).
A-monoxide	Wüstite (Fe <sub>x</sub> O), Lime (CaO), Periclase (MgO), Magnesiowüstite (MgO-Fe <sub>x</sub> O), Manganowüstite (Mn-Fe) <sub>x</sub> O, Manganosite (Mn <sub>x</sub> O).
A-Clinopyroxene	Clino-enstatite (MgSiO <sub>3</sub> ), (Metastable) clino-ferrosilite (FeSiO <sub>3</sub> )
A-Orthopyroxene	Ortho-enstatite (MgSiO <sub>3</sub> ), (Metastable) ortho-ferrosilite (FeSiO <sub>3</sub> )
A-Protopyroxene	Proto-enstatite (MgSiO <sub>3</sub> )
LowClinopyroxene	CaMgSi <sub>2</sub> O <sub>6</sub> - Mg <sub>2</sub> Si <sub>2</sub> O <sub>6</sub> solid solution
A-Wollastonite	CaSiO <sub>3</sub>
Bredigite	Ca <sub>3</sub> (Ca,Mg) <sub>4</sub> Mg(SiO <sub>4</sub> ) <sub>4</sub>
bC2SA	Ca <sub>2</sub> SiO <sub>4</sub>
aC2SA	Ca <sub>2</sub> SiO <sub>4</sub>
A-Melilite	Akermanite (Ca <sub>2</sub> MgSi <sub>2</sub> O <sub>7</sub> )
A-Olivine	Forsterite (Mg <sub>2</sub> SiO <sub>4</sub> ), Fayalite (Fe <sub>2</sub> SiO <sub>4</sub> ),
Cordierite	Al <sub>4</sub> Fe <sub>2</sub> Si <sub>5</sub> O <sub>18</sub> - Al <sub>4</sub> Mg <sub>2</sub> Si <sub>5</sub> O <sub>18</sub> solution
Mullite	[Al,Fe] <sub>2</sub> [Al,Si,B,Fe][O,Va] <sub>5</sub>
CAFS	Ca <sub>2</sub> (Al,Fe) <sub>8</sub> SiO <sub>16</sub>
CAF6	Ca(Al, Fe) <sub>12</sub> O <sub>19</sub>
CAF3	Ca(Al,Fe) <sub>6</sub> O <sub>10</sub>
CAF2	Ca(Al,Fe) <sub>4</sub> O <sub>7</sub>
CAF1	Ca(Al,Fe) <sub>2</sub> O <sub>4</sub>
C2AF	Ca <sub>2</sub> (Al,Fe) <sub>2</sub> O <sub>5</sub>
C3AF	Ca <sub>3</sub> (Al,Fe) <sub>2</sub> O <sub>6</sub>
Corundum	Corundum (Al <sub>2</sub> O <sub>3</sub> )
Garnets	Grossularite (Ca <sub>3</sub> Al <sub>2</sub> Si <sub>3</sub> O <sub>12</sub> )
CaSpinel	CaCr <sub>2</sub> O <sub>4</sub> - CaFe <sub>2</sub> O <sub>4</sub> solid solution
A-Tetragonal-Spinel	Low-temperature Mn <sub>3</sub> O <sub>4</sub> dissolving Fe, Mg, Cr, Ti and Al
Bixbyite	Mn <sub>2</sub> O <sub>3</sub>
Braunite	Non-stoichiometric Mn <sub>7</sub> SiO <sub>12</sub>
Rhodonite	MnSiO <sub>3</sub>
Pyroxmangite	MnSiO <sub>3</sub> - MgSiO <sub>3</sub> solution

### 2.4 Data processing

The procedure for data processing is presented in Fig. 1. The Excel spreadsheet files must be generated once for a sample

with a specified composition. After that, the phase identification is carried out using MATLAB scripting. As input values, the WDS point analysis of Mg, Al, Si, Ca, Cr, Mn, and Fe is compared with the compositions of all existing phases at all temperatures, read from the Excel spreadsheets.

The results are then visualized for a given point analysis, showing the best  $R^2$  fit for the calculated vs. analyzed Mg, Al, Si, Ca, Cr, Mn, and Fe contents, along with their occurrence temperature for both equilibrium cooling and Scheil–Gulliver solidification modes.

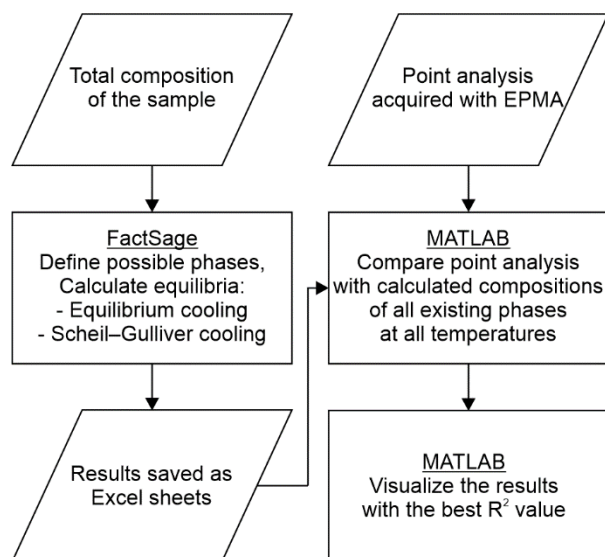


Fig. 1. Flowsheet for the data processing.

### 3. RESULTS AND DISCUSSION

#### 3.1 Slag#1

The fractions of equilibrated phases according to equilibrium cooling in Slag#1 is presented in Fig. 2. At 1800 °C, the slag is fully liquid. As the temperature decreases to approx. 1700 °C (liquidus temperature), tetragonal spinel is the first solid phase to form, followed by monoxide phase at around 1500 °C. As the temperature further decreases to 1050 °C, the system is completely solid (solidus temperature). The results reveal solid state phase transformations of silicate phases bc2SA, bredigite and olivine. In addition, stable phases include calcium aluminate C3AF with Fe solubility, another monoxide phase, and a small amount of another spinel phase (B-spinel) below 400 °C.

Scheil–Gulliver solidification results for Slag#1 are presented in Fig. 3. Tetragonal spinel is the first forming solid phase, followed by monoxide. Other phases to form during solidification are bc2SA and C3AF. In this case, the solidus temperature is comparable to the value with equilibrium cooling. By definition, solid state transformations do not occur in Scheil–Gulliver solidification. Therefore, bc2SA silicate exists in the final solidification phase structure.

BSE image of Slag#1 sample is presented in Fig. 4, acquired with EPMA. The phase identification procedure results are presented for selected points of interest: A, B, and C with distinct grayscale brightness.

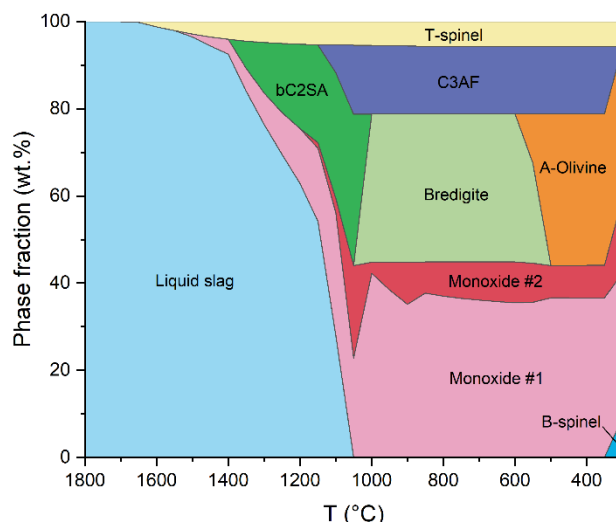


Fig. 2. Phase fractions in Slag#1, equilibrium cooling.

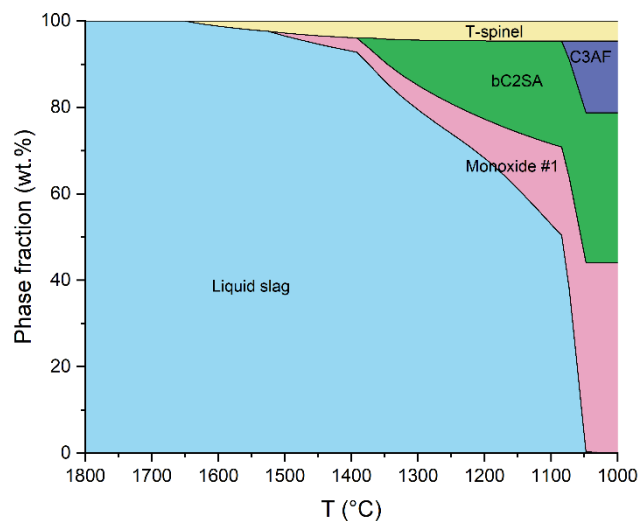


Fig. 3. Phase fractions in Slag#1, Scheil–Gulliver solidification.

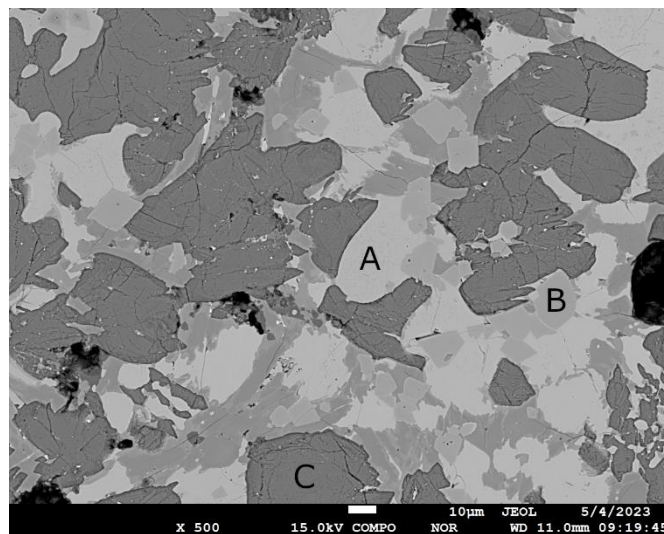


Fig. 4. BSE image of Slag#1.

The graphs presented in Fig. 5 present the WDS analyzed contents of Mg, Al, Si, Ca, Cr, Mn, and Fe for point A in Fig. 4. Comparing to the FactSage calculated values, it is found that the analyzed values correspond best to monoxide phase around 1100–1150 °C both for equilibrium cooling and Scheil–Gulliver solidification. According to the chemical composition, this appears to be a wüstite based solid solution (Fe,Mn,Mg,Cr)O.

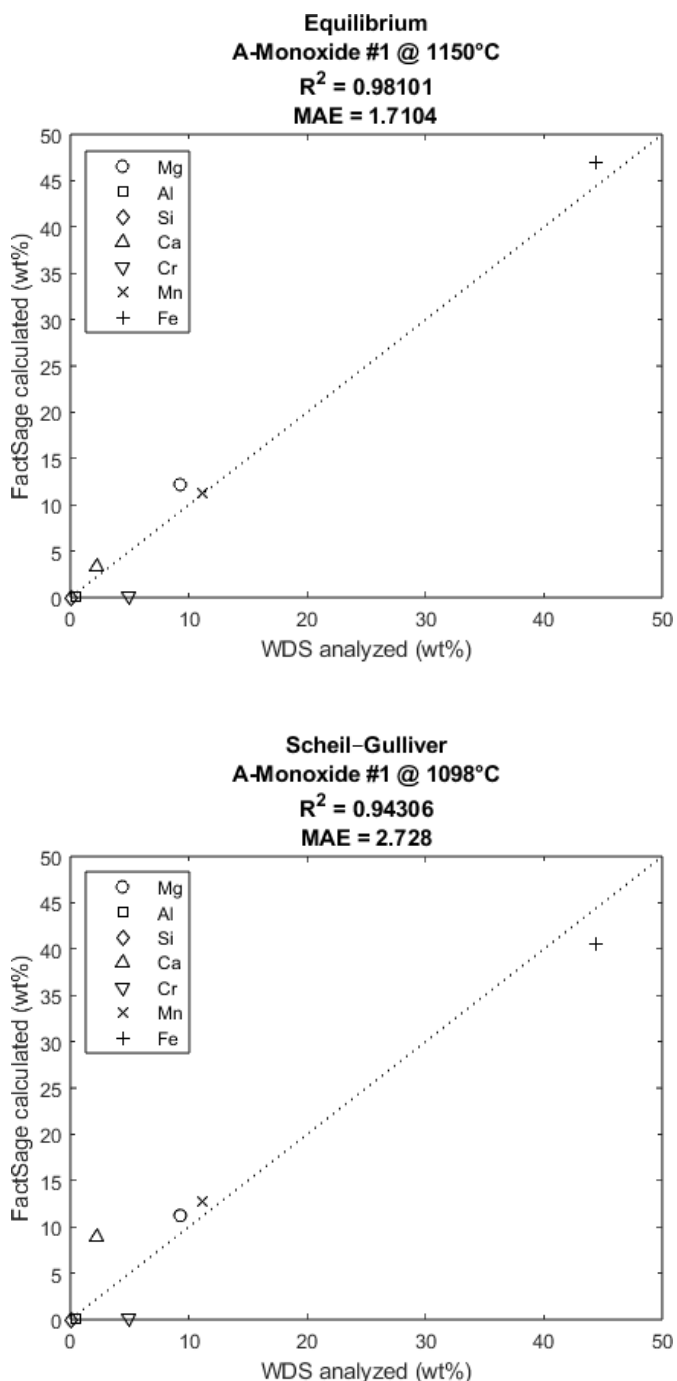


Fig. 5. Elemental composition of point A compared to the calculated stable phases in Slag#1.

For point B observed in Fig. 4, the results are shown in Fig. 6. According to both cooling calculations, the composition is closest to tetragonal spinel occurring at high temperatures, with an  $R^2$  value of approx. 0.63. Noting the analyzed high chromium content, the phase composition is close to chromite (FeCr<sub>2</sub>O<sub>4</sub>), another spinel phase.

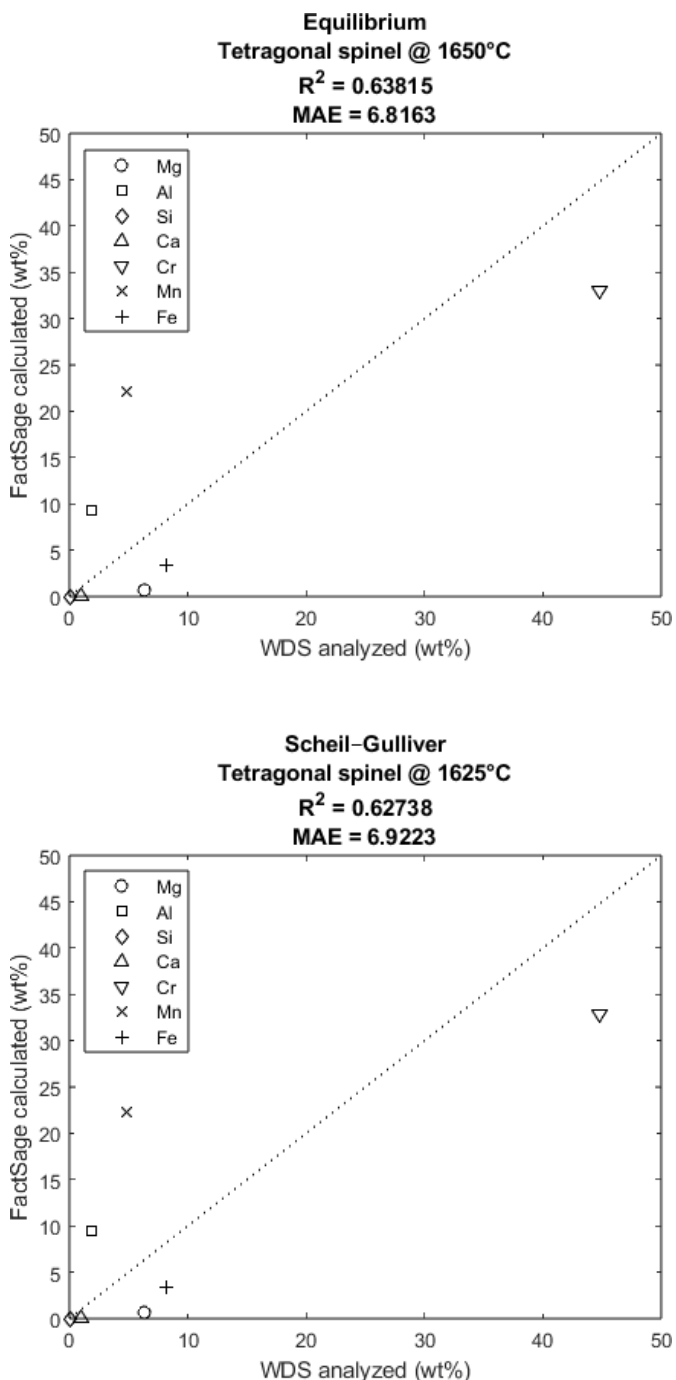


Fig. 6. Elemental composition of point B compared to the calculated stable phases in Slag#1.



It is observed in Fig. 7 that the composition in point C contains mostly calcium and silicon, whereas other considered elements are close to zero, suggesting a silicate phase. The equilibrium cooling calculation indicates that the phase is olivine at the minimum calculation temperature. The Scheil–Gulliver solidification calculation shows another silicate phase, bC2SA.

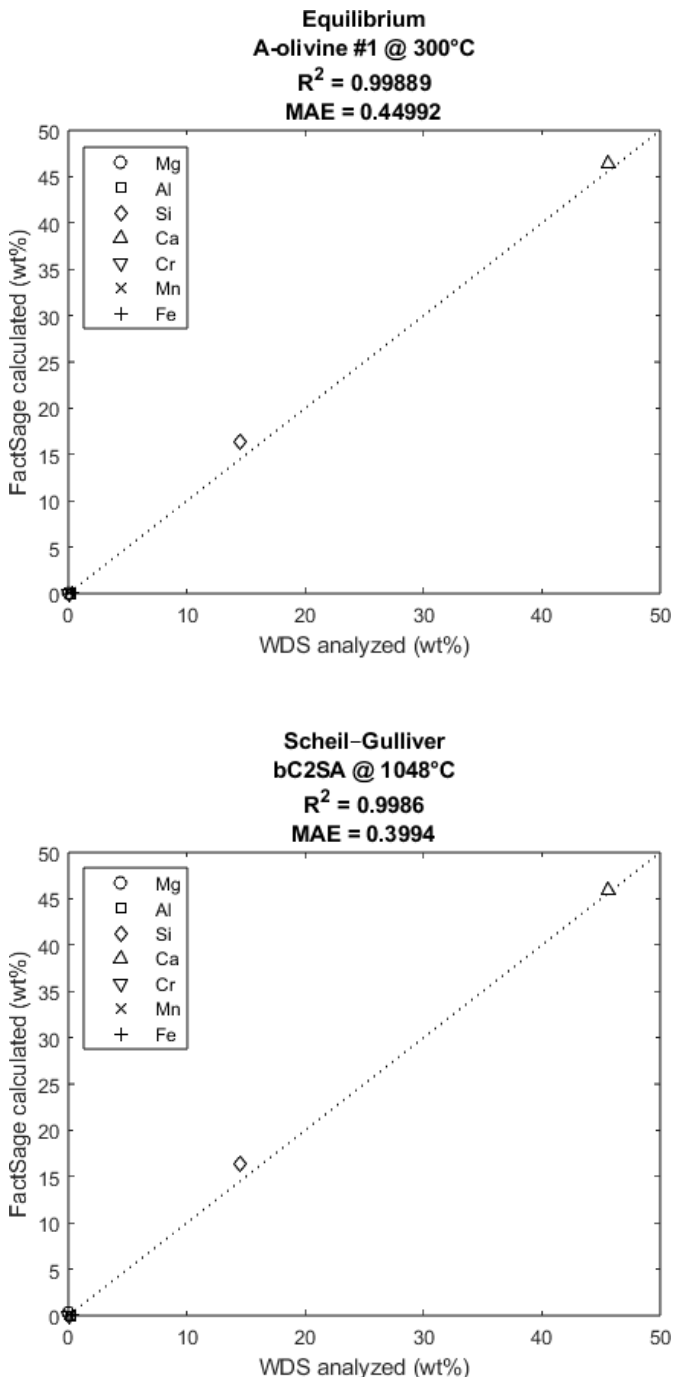


Fig. 7. Elemental composition of point C compared to the calculated stable phases in Slag#1.

### 3.2 Slag#2

The fractions of phases according to equilibrium cooling in Slag#2 is presented in Fig. 8. At 1800 °C, the slag is fully liquid. As the temperature decreases to approx. 1700 °C, monoxide is the first solid phase to form, followed by tetragonal spinel almost immediately. The system is completely solid at ~1100 °C. Similarly to Slag#1, silicate phases bC2SA, bredigite and olivine undergo solid state phase transformation. A stability region of calcium aluminates CAF1 and C3AF is found, both with Fe solubility. Another monoxide phase and B-spinel appear at relatively low temperatures.

The Scheil–Gulliver solidification results for Slag#2 are presented in Fig. 9. As expected from the equilibrium cooling results, the first forming solid phases are monoxide and tetragonal spinel. Other precipitating phases include bC2SA, CAF1, and C3AF. Contrary to equilibrium cooling, melilite phase exists in the results. Here, the solidus temperature is lower (~1050 °C) than in equilibrium cooling (~1100 °C).

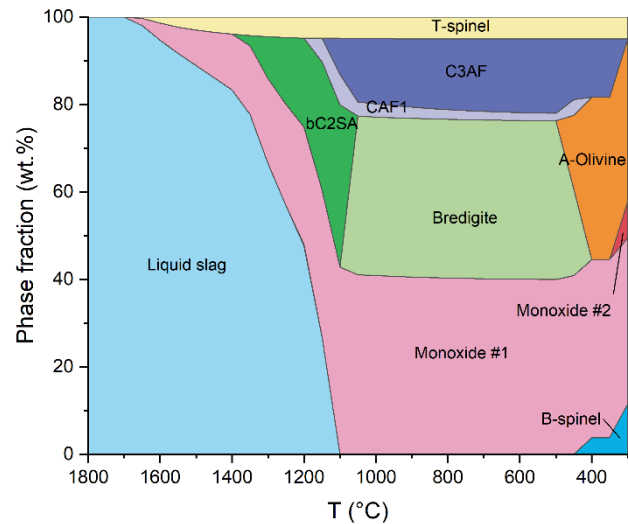


Fig. 8. Phase fractions in Slag#2, equilibrium cooling.

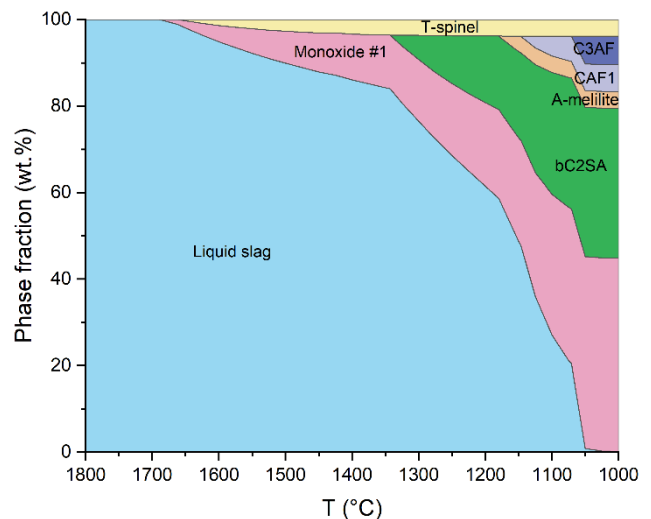


Fig. 9. Phase fractions in Slag#2, Scheil–Gulliver solidification.

BSE image of Slag#2 sample is presented in Fig. 10. For the second slag sample, the procedure results are presented for selected points of interest: D, E, and F.

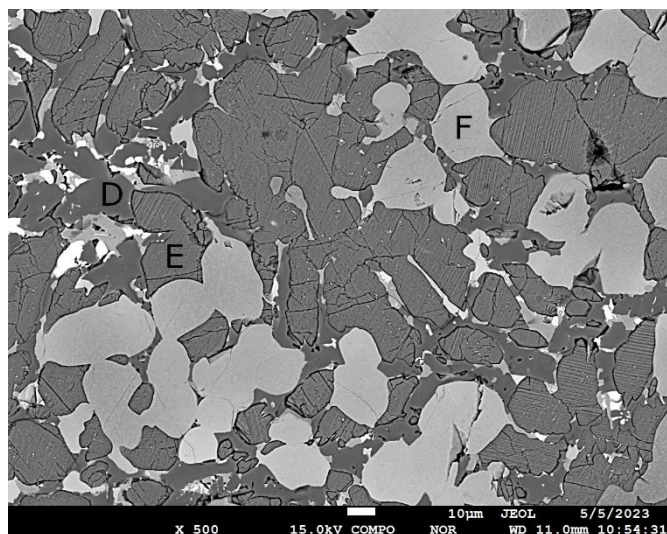


Fig. 10. BSE image of Slag#2.

For point D, the phase identification results are plotted on Fig. 11. Here, the WDS analysis contains mostly calcium and aluminum, with near-zero amount of other considered elements. From both calculation results, the composition corresponds to calcium aluminate C3AF. The calculated composition is very close to stoichiometric  $3\text{CaO}\cdot\text{Al}_2\text{O}_3$  in both cases, explaining the identical resulting  $R^2$  values.

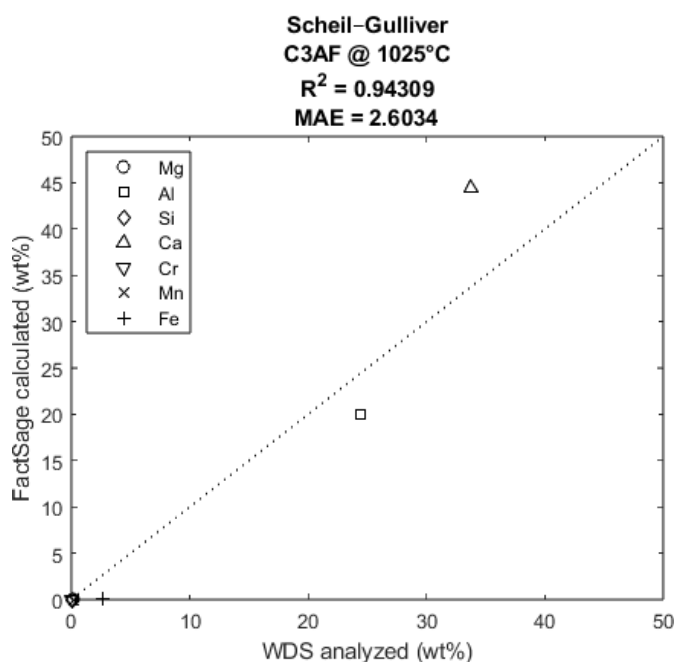
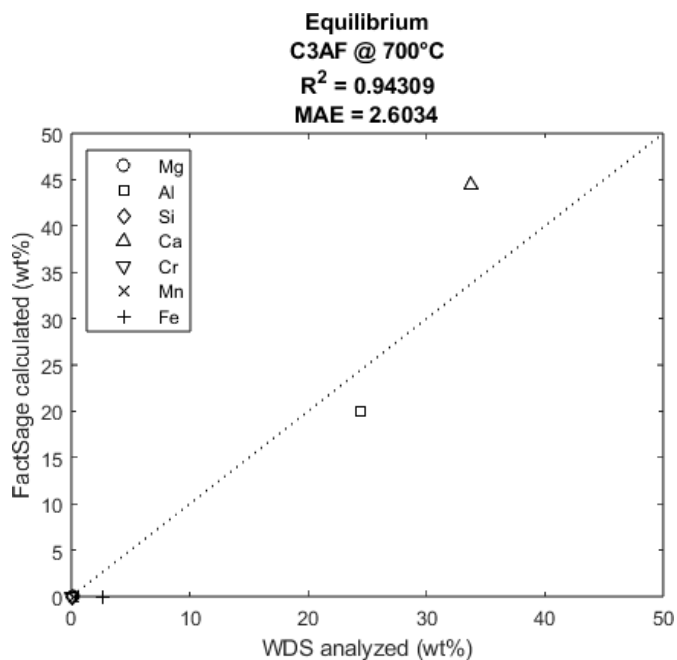


Fig. 11. Elemental composition of point D compared to the calculated stable phases in Slag#2.

It is observed in Fig. 12 that the composition in point E contains mostly calcium and silicon, whereas other considered elements are close to zero, suggesting a silicate phase. The equilibrium cooling calculation indicates that the phase is olivine occurring at low temperatures. The Scheil–Gulliver solidification calculation shows another silicate phase, bC2SA with a composition occurring at 1025 °C.

Figure 13 shows the results for point F in Slag#2. Comparing to the calculated values, it is observed that the analyzed values match a monoxide phase both for equilibrium cooling and Scheil–Gulliver solidification. According to the WDS analysis, this appears to be a wüstite based solid solution (Fe,Mn,Mg,Cr)O.

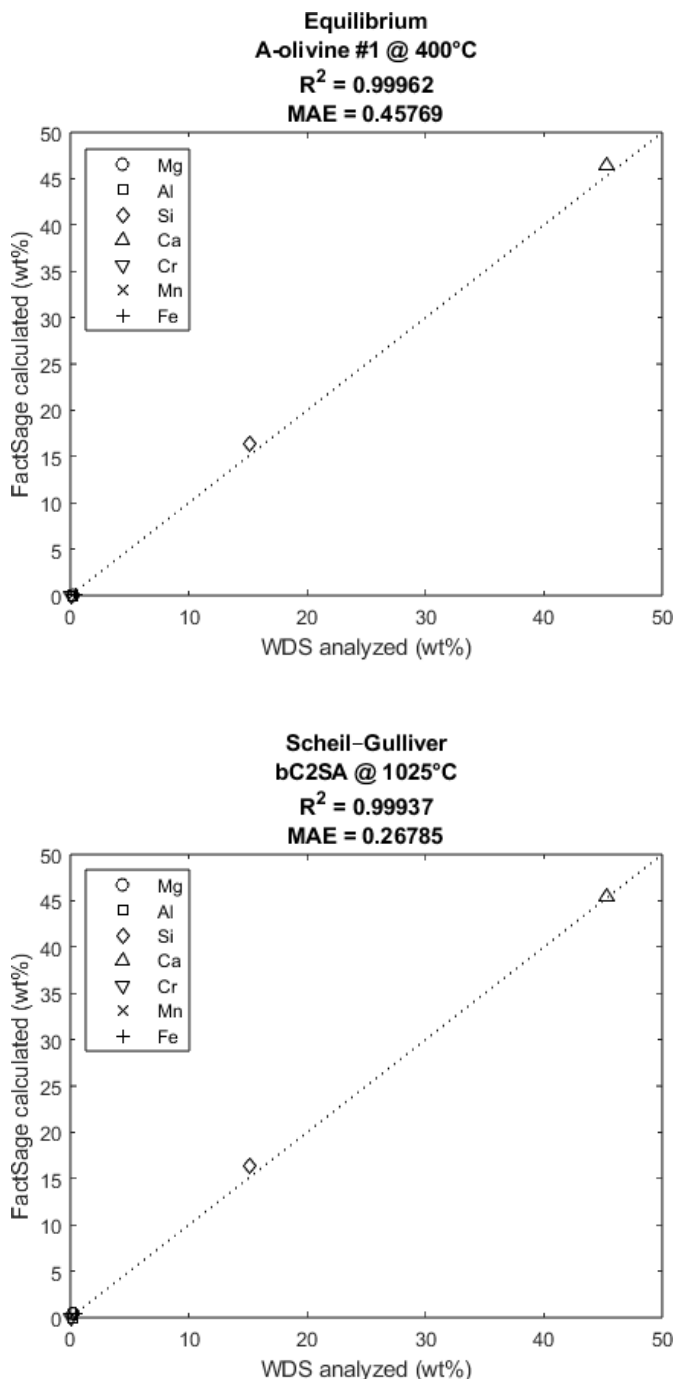


Fig. 12. Elemental composition of point E compared to the calculated stable phases in Slag#2.

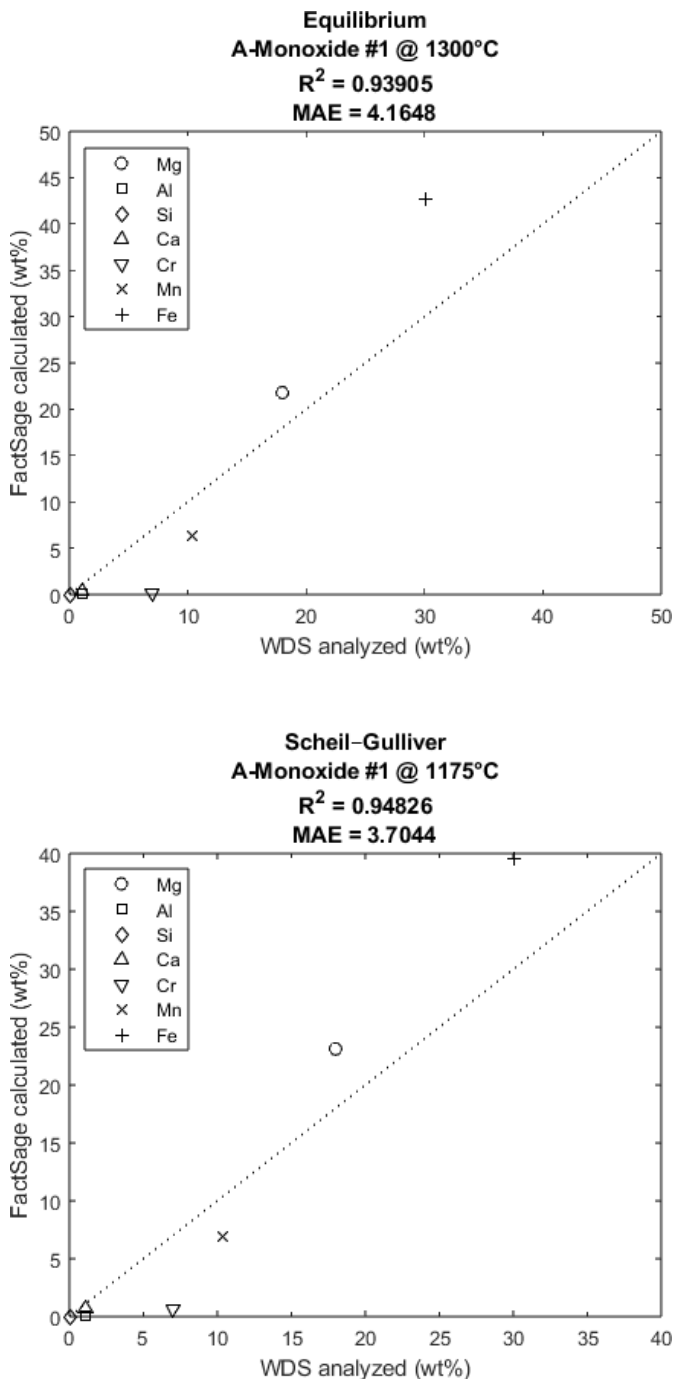


Fig. 13. Elemental composition of point F compared to the calculated stable phases in Slag#2.

## 4. CONCLUSIONS

In addition to chemical composition, the cooling rate affects the phases forming during solidification and cooling of oxide-based steelmaking slags. The thermodynamic calculation software FactSage was used to estimate the phase fractions and their composition as a function of temperature, according to equilibrium cooling and Scheil–Gulliver solidification. The current study presents a method to compare the thermodynamic calculation results with the elemental analysis obtained from steelmaking slags. Interpreting the results allows the estimation of the formation sequence and temperature for the identified phases.

In this study, two EAF slag materials, samples Slag#1 and Slag#2 were investigated. The thermodynamic calculations show some similarities in the materials: the stable phases include monoxide (wüstite), spinel, calcium silicates and calcium aluminate phases. The equilibrium cooling results illustrated solid state phase transformations of calcium silicates, bC2SA–bredigite–olivine with decreasing temperature. The presented phase identification procedure can be used to visualize the elemental composition of the analyzed points and confirm the mineralogical group. For monoxide phases, calcium silicates, and calcium aluminates, the FactSage calculated and WDS analyzed compositions correspond with excellent  $R^2$  values ( $>0.9$ ). However, the FactSage calculated chromium content in spinel phase (chromite  $\text{FeCr}_2\text{O}_4$ ), was lower than analyzed, resulting in moderate  $R^2$  value.

The presented approach can be used to visualize the results for different experimental scenarios: altering the cooling rate of the slag, effect of iron saturation in the system, or the atmosphere effect. In addition, the results using different databases in the FactSage calculations can be compared.

The procedure is not restricted to slag materials, instead, a similar approach can be applied to other systems where solidification and possible solid state phase transformations take place. Besides EPMA, other analysis methods can be employed, for instance Scanning Electron Microscope (SEM).

## ACKNOWLEDGEMENTS

The authors are grateful to the funding of the research program FFS2 (Towards Fossil-Free Steel 2, 5534/31/2023) funded by Business Finland.

## REFERENCES

- Andersson, A., Isaksson, J., Lennartson, A., and Engström, F. (2024). Insights into the Valorization of Electric Arc Furnace Slags as Supplementary Cementitious Materials. *Journal of Sustainable Metallurgy* 10(1), 96–109.
- Bale, C.W., Chartrand, P., Degterov, S.A., Eriksson, G., Hack, K., Mahfoud, R.B., Melançon, J., Pelton, A.D., and Petersen, S. (2002), FactSage Thermochemical Software and Databases. *Calphad* 26(2), 189–228.
- Bru, K., Seron, A., Morillon, A., Algermissen, D., Lerouge, C. and Menad, N. (2021). Characterization of a Chromium-Bearing Carbon Steel Electric Arc Furnace Slag after Magnetic Separation to Determine the Potential for Iron and Chromium Recovery. *Minerals*, 12(1), 47.
- Durinck, D., Jones, P.T., Blanpain, B., and Wollants, P. (2007). Slag Solidification Modeling Using the Scheil–Gulliver Assumptions. *Journal of the American Ceramic Society*, 90 (4), 1177–1185.
- Euroslag. Statistics (2018). Available online: <https://www.euroslag.com/products/statistics/statistics-2018/> (accessed on 8 May 2024).
- Li, Y., Guo, K., Xiang, J., Pei, G., Lv, X. (2022). Effect of cooling method on the mineralogy and stability of steel slag. *ISIJ International*, 62 (11), 2197–2206.
- Mombelli, D., Mapelli, C., Barella, S., Di Cecca, C., Le Saout, G. and Garcia-Diaz, E. (2016). The effect of chemical composition on the leaching behaviour of electric arc furnace (EAF) carbon steel slag during a standard leaching test. *Journal of Environmental Chemical Engineering* 4(1), 1050–1060.

## Utilizing computational thermodynamics in characterization and classification of non-metallic inclusions in Ti-deoxidized steels

Tuomas Alatarvas\*, Henri Tervo\*\*, Antti Kaijalainen\*\*, Qifeng Shu\*

\* Process Metallurgy Research Unit, Centre for Advanced Steel Research,  
University of Oulu, P.O. Box 4300, FI-90014 Oulu, Finland

\*\* Materials and Mechanical Engineering, Centre for Advanced Steel Research,  
University of Oulu, P.O. Box 4200, FI-90014 Oulu, Finland

**Abstract:** Non-metallic inclusions (NMIs) are micrometer-sized particles observed in all steel materials, often considered detrimental. In this study, NMIs in titanium-deoxidized steels were investigated, complemented with thermodynamic assessment for more accurate phase characterization. The NMIs were analyzed with a Jeol JSM-7900F FESEM-EDS (Field Emission Scanning Electron Microscope equipped with Energy Dispersive X-ray Spectroscopy). For automated particle analyses on FESEM, Aztec Feature runs were carried out on polished steel samples, providing the elemental composition, in addition to morphological data, for each observed NMI. Utilizing the obtained EDS analyses, the fractions of oxides ( $\text{Al}_2\text{O}_3$ ,  $\text{MnO}$ ,  $\text{TiO}_x$ ), manganese sulfide ( $\text{MnS}$ ), and titanium nitride ( $\text{TiN}$ ) in each NMI are estimated with a MATLAB script. Based on the estimated phase contents, a composition-based classification method for the NMIs is presented. To visualize the phase contents of the observed NMIs, the calculated compositions are plotted on  $\text{MnO-TiO}_2\text{-Ti}_2\text{O}_3$  ternary diagrams. Computational thermodynamics software FactSage 8.3 was firstly utilized to estimate the fully liquid NMI composition region at steelmaking temperatures in the considered ternary oxide system of  $\text{MnO-TiO}_2\text{-Ti}_2\text{O}_3$ . Secondly, the thermodynamic stability of NMI phases in the steel was assessed with decreasing temperature during the solidification of steel. The current study demonstrates how computational thermodynamics can be utilized in characterization and classification of non-metallic inclusions and giving insight on their formation during solidification of steel.

**Keywords:** computational thermodynamics, non-metallic inclusions, steel

### 1. INTRODUCTION

The composition of non-metallic inclusions (NMIs) in steel is mainly dependent on the total elemental composition of the steel. To enhance steel cleanliness by decreasing the oxygen content, the molten steel must be deoxidized by introducing deoxidation agents. Conventional deoxidation agents include aluminum and silicon, binding the oxygen to  $\text{Al}_2\text{O}_3$  and  $\text{SiO}_2$  based NMIs. Titanium can be used as a deoxidation agent when titanium oxide based NMIs are desired. These types of NMIs are known to affect the microstructure in welded structures, both in weld metal and in heat-affected zones by inducing the formation of acicular ferrite (Tervo et al. 2020, 2021; Loder et al. 2016)

The titanium contents in Ti-deoxidized steels affect the phases observed in NMIs. According to the presentation by Panda et al. (2019), the  $\text{MnO-TiO}_2\text{-Ti}_2\text{O}_3$  oxide system includes spinel, ilmenite, and pseudobrookite solid solutions with varying compositions, as illustrated in Fig. 1. According to Wang and Li (2021) and Wang et al. (2018), the higher Mn content in spinel and ilmenite – compared to pseudobrookite – promotes the formation of beneficial acicular ferrite by depleting the Mn content of the steel matrix in the vicinity of NMIs.

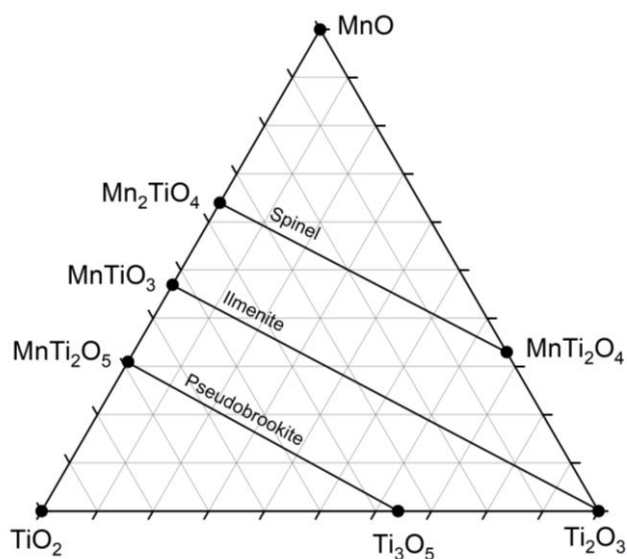


Fig. 1.  $\text{MnO-TiO}_2\text{-Ti}_2\text{O}_3$  ternary diagram with the main solid solutions.



In addition to the primary NMIs formed during the deoxidation process in the molten steel, other NMIs can form during further processing and solidification of the steel. The primary NMIs can act as nucleation sites for forming phases. To estimate the phases forming in the molten steel and during solidification of steel, thermodynamic assessment can be applied. While the rate of the reactions is not taken into account, the amount and composition of stable phases at a given temperature can be estimated. The assumption of thermodynamic equilibrium can be considered reasonable in molten steel (>1500 °C) with homogeneous composition. As the steel solidifies, kinetics constrain achieving the equilibrium. However, the thermodynamic equilibrium points the direction of possible NMI reactions also in the solid state.

The current study demonstrates how computational thermodynamics and data processing can be utilized in characterization and classification of non-metallic inclusions and giving insight on their formation during solidification of steel.

## 2. MATERIALS AND METHODS

### 2.1 Investigated materials

In this study, NMIs in two samples of titanium-deoxidized steels were investigated. The materials were denoted as Ti<sub>Low</sub> and Ti<sub>High</sub>, according to their titanium contents (Tervo et al. 2020, 2021). The chemical compositions of the samples are presented in Table 1. The material was cast in a laboratory to 20 kg ingots. After hot rolling of the ingots, polished cross section samples were prepared for NMI investigation.

Table 1. Composition of the studied steels, in wt%.

	Ti <sub>Low</sub>	Ti <sub>High</sub>
C	0.05	0.05
Si	0.23	0.03
Mn	1.7	1.7
P	0.007	0.005
S	0.003	0.003
Al	0.003	0.002
Nb	0.01	0.01
V	0.07	0.07
Ti	0.016	0.027
N	0.008	0.006
O	0.0047	0.0080

### 2.2 Scanning electron microscopy

Non-metallic inclusions were analyzed with a Jeol JSM-7900F FESEM-EDS (field emission scanning electron microscope with energy dispersive X-ray spectroscopy). The characteristics, including elemental composition and morphology, of particles larger than 2 µm in an equivalent circle diameter were analyzed with *Aztec Feature* automated particle analysis functionality. The scanned section in each sample covered an area of approximately 52 mm<sup>2</sup>, comprising of total 330 fields at 250x magnification. Figure 2 presents one of the fields imaged with backscattered electron (BSE) detector, illustrating gray NMI particles on light background, i.e., steel.

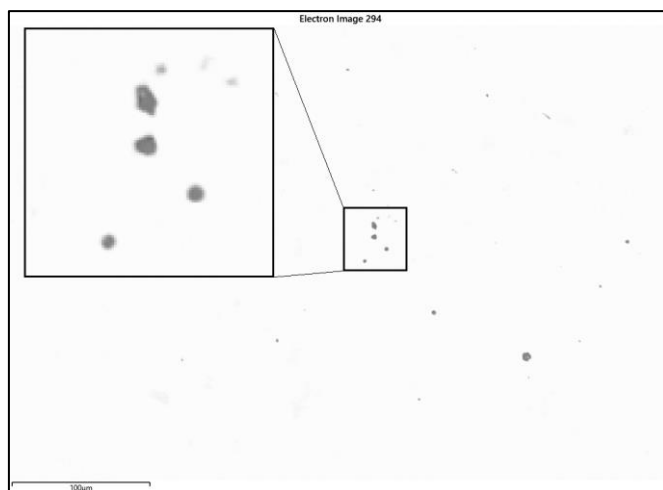


Fig. 2. Single field imaged on FESEM, showing NMI particles in steel.

After the *Feature* runs, representative NMIs were selected from the dataset. High-resolution BSE images and EDS point analyses were acquired after relocating the selected sites.

### 2.3 Classification of non-metallic inclusions

There are various ways to process the raw data acquired with the *Feature* particle analysis runs (Alatarvas et al 2019, Warttainen et al. 2020). Here, all particles containing less than 30 wt.% of Al, Mn, N, Ti and S in total were discarded. The remaining particles were considered true NMIs, comprising of components Al<sub>2</sub>O<sub>3</sub>, MnO, MnS, TiN, TiO<sub>2</sub> and Ti<sub>2</sub>O<sub>3</sub>. To estimate the phase composition, the following stepwise procedure was applied, utilizing the EDS analyses of Al, Mn, Ti, N, S, and O of each NMI.

1. All analyzed Al bound to Al<sub>2</sub>O<sub>3</sub>,
2. All analyzed S bound to MnS,
3. All analyzed N bound to TiN,
4. Remaining Mn (from the 2<sup>nd</sup> step) bound to MnO,
5. Remaining Ti (from the 3<sup>rd</sup> step) and O (from the 1<sup>st</sup> and 4<sup>th</sup> step) determine the proportions of TiO<sub>2</sub> and Ti<sub>2</sub>O<sub>3</sub>.

The total amount of components was scaled to 100 wt.%, after which the NMIs were classified into combinations of Al<sub>2</sub>O<sub>3</sub>, MnO, MnS, TiN, and TiO<sub>x</sub> with a 10 wt.% threshold, totaling 31 combinations, i.e., inclusion classes. For plotting, the oxide system was considered MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub>, allowing the construction of ternary diagrams.

Executed with MATLAB scripts, the presented approach allows the data processing and visualization of the NMI characteristics in the scanned area, as well as for individual NMIs with point analyses.

### 2.4 Computational thermodynamics

Computational thermodynamics software FactSage 8.3 (Bale et al. 2002) and its *FToxid* database was used to assess the composition region in the MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> system, where oxides exist only as fully molten slag. Considering steelmaking temperatures (1500–1700 °C) and plotted into ternary MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> diagram with the *Phase Diagram* module, the region is denoted as liquid window. Here, solid

solutions of monoxide, titania spinel, rutile, ilmenite, and pseudobrookite were considered as precipitating phases, included in the considered system.

Further, FactSage was used to assess the stability of NMIs as a function of temperature with the steel compositions reported, with total mass of 100 g of material. *Equilib* module with the databases *FSstel* and *FToxid* were selected for the calculations, with the following solid solutions in addition to liquid steel and liquid slag: FCC (face-centered cubic, austenite), BCC (body-centered cubic, ferrite), cementite ( $\text{Fe}_3\text{C}$ ), MS-c (MnS–FeS solid solution), A-monoxide, corundum, A-rutile, A-ilmenite, A-olivine, pseudobrookite, B-spinel, and titania spinel were considered. For plotting of the results, also TiN (as FCC phases) was included, known as a possible NMI type, whereas other nitrides and carbides were omitted.

### 3. RESULTS AND DISCUSSION

#### 3.1 Sample-specific NMI characteristics

The *Feature* particle analyses resulted in total of 2623 and 1255 observed particles in  $\text{Ti}_{\text{Low}}$  and  $\text{Ti}_{\text{High}}$  samples, respectively. After applying the compositional criteria, the number of particles were decreased to 2210 and 1177, to represent the true NMIs in  $\text{Ti}_{\text{Low}}$  and  $\text{Ti}_{\text{High}}$  samples, respectively. The results according to the classification procedure are presented in Fig. 3. It can be seen that the number of particles per scanned area is in the range of 22–42 particles per square millimeter. In  $\text{Ti}_{\text{Low}}$  sample, the most dominant NMI class is “MnO MnS TiN TiO<sub>x</sub>”, which translates to oxides in the MnO–TiO<sub>x</sub> system with precipitated sulfide (MnS) and nitride (TiN) phases. Both samples are also characterized by the occurrence of “MnO TiO<sub>x</sub>” type inclusions, i.e., oxides without sulfide and nitride phases.

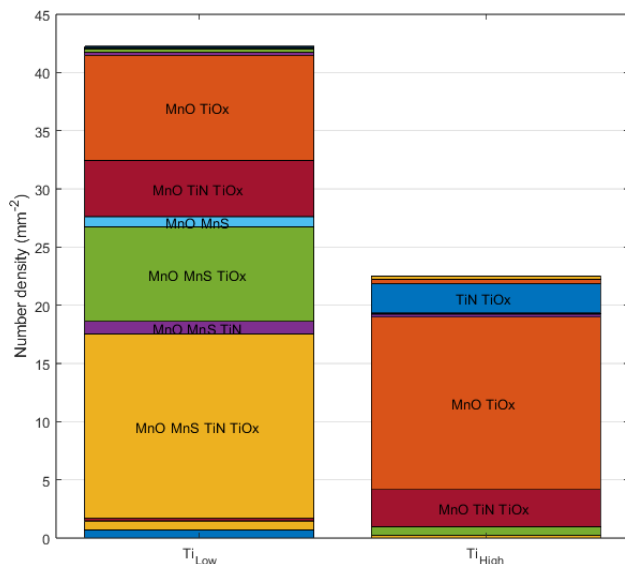


Fig. 3. Number of NMIs per unit area, by NMI classes.

#### 3.2 Selected NMIs with point analyses

Selected NMIs were located from the *Feature* dataset after the run, saving high-resolution BSE images. The BSE images illustrate compositional contrast: high atomic number (heavy)

elements show as brighter areas than low atomic number (light) elements. Therefore, the images can aid in observing the phase complexity of NMIs.

EDS analyses were manually acquired from certain points in the NMIs. Further, the EDS analyses were used as an input in a MATLAB script to visualize the oxide composition in a ternary diagram. In this section, a total of five representative oxide inclusions are presented, with corresponding compositions plotted on ternary diagrams.

In Fig. 4, a BSE image of an oxide NMI in the  $\text{Ti}_{\text{Low}}$  sample is presented. The structure of the NMI comprises of the primary oxide phases (points 8 and 9) with precipitated phases on their surfaces (points 7 and 10). After the calculation procedure and plotting onto ternary diagram, shown in Fig. 5, it is observed that the points 7–9 locate near the pseudobrookite solid solution line ( $\text{Mn}_2\text{TiO}_5$ – $\text{Ti}_3\text{O}_5$ ). On the other hand, the point 10 is closer to the spinel composition on the  $\text{Mn}_2\text{TiO}_4$ – $\text{MnTi}_2\text{O}_4$  line.

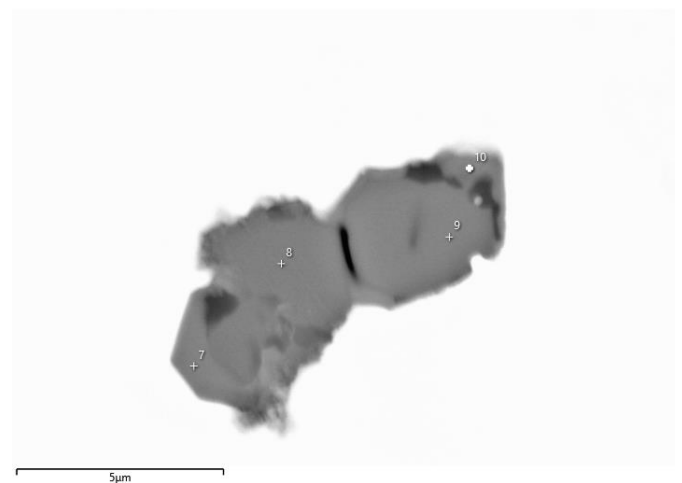


Fig. 4. BSE image of NMI#1, with EDS analyzed points.

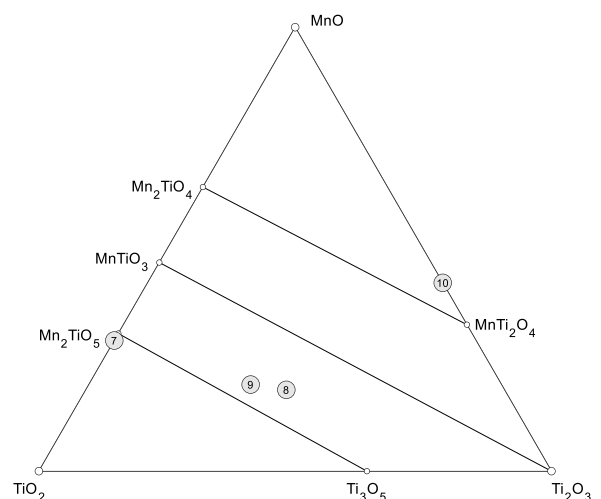


Fig. 5. Estimated oxide compositions of the analyzed points in NMI#1.

Figure 6 presents a BSE image of another oxide NMI in the  $Ti_{Low}$  sample. A layered structure is observed, with pseudobrookite phases in the core, found on points 23 and 24, as shown in Fig. 7. The darker phase at point 22 is spinel, according to the phase identification procedure.

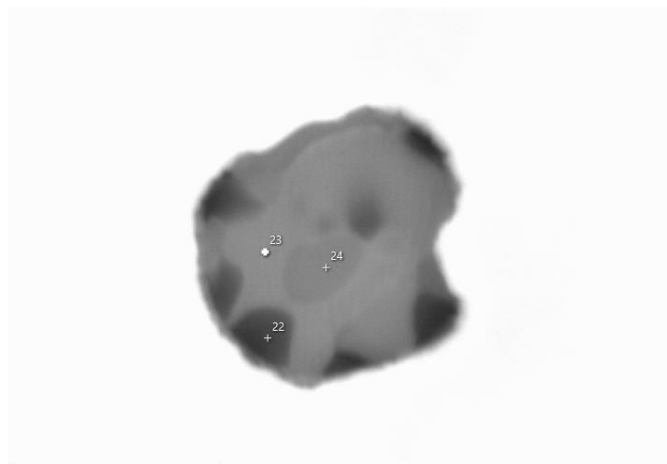


Fig. 6. BSE image of NMI#2, with EDS analyzed points.

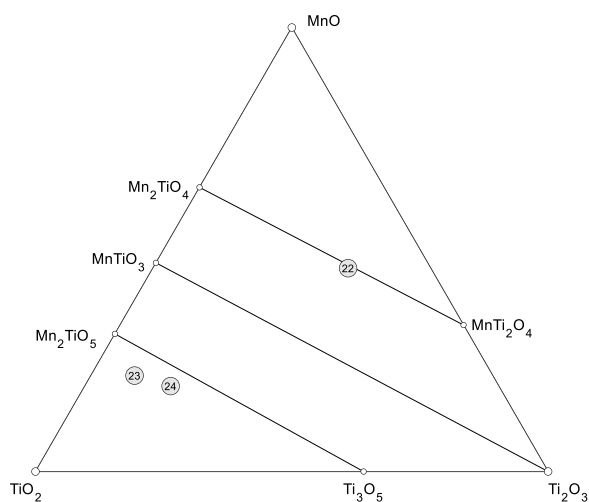


Fig. 7. Estimated oxide compositions of the analyzed points in NMI#2.

A cross section of a spherical NMI is presented in Fig. 8, observed in  $Ti_{High}$  sample. The uniform gray shade and the roundness suggest that the NMI has been fully liquid in the molten steel, before solidification. As observed in Fig. 9, the variation in the calculated oxide composition is negligible, with points 1–3 located near the pseudobrookite solid solution line.

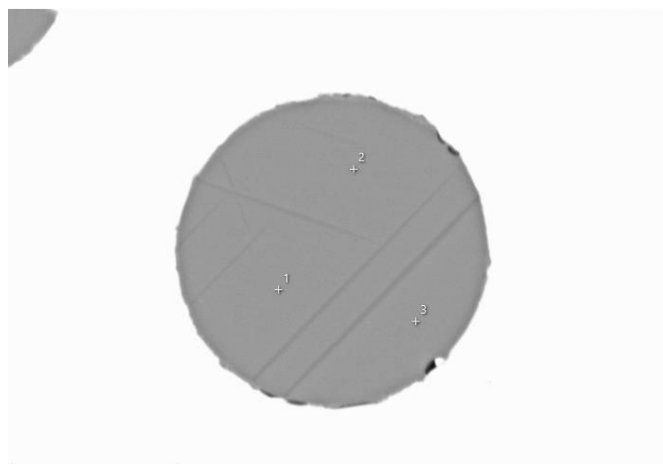


Fig. 8. BSE image of NMI#3, with EDS analyzed points.

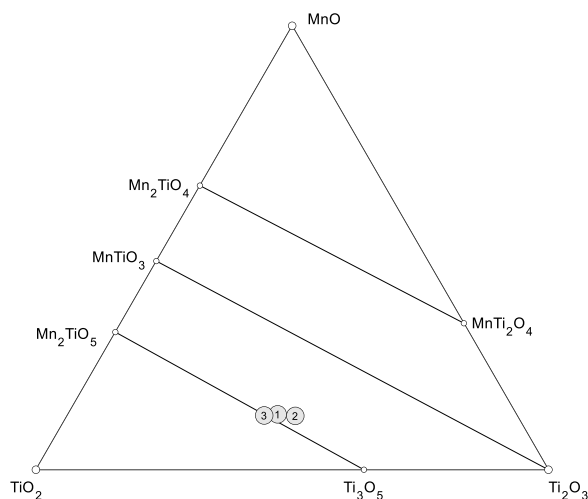


Fig. 9. Estimated oxide compositions of the analyzed points in NMI#3.

In Fig. 10, BSE image of another spherical NMI in the  $Ti_{High}$  sample is presented. Comparing with the estimated compositions of points 8–10 in Fig. 11, it is concluded that the main phase is pseudobrookite, which has been most likely molten at steelmaking temperatures. The enclosed light phase has a composition higher in  $TiO_2$  content, closer to the rutile corner of the presented ternary diagram.

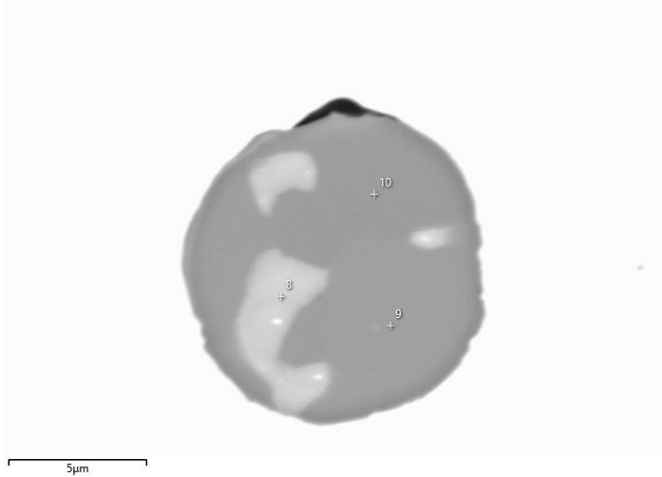


Fig. 10. BSE image of NMI#4, with EDS analyzed points.

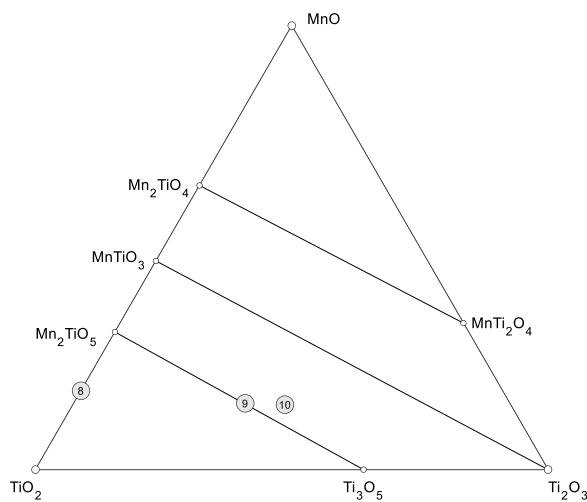


Fig. 11. Estimated oxide compositions of the analyzed points in NMI#4.

Some agglomerated NMIs were observed in the  $Ti_{High}$  sample. The NMI presented in Fig. 12 is separated by a crack, seen as a black stripe in the BSE image. The primary composition of the oxides (points 16–19) are located near the pseudobrookite line. The composition of point 20 is nearly all titanium nitride ( $TiN$ ), thus it is not plotted in the oxide ternary diagram in Fig. 13. Found on the periphery of the NMI, it can be concluded  $TiN$  has precipitated as a secondary phase on the surface of the primary oxide NMI during the solidification of the steel.

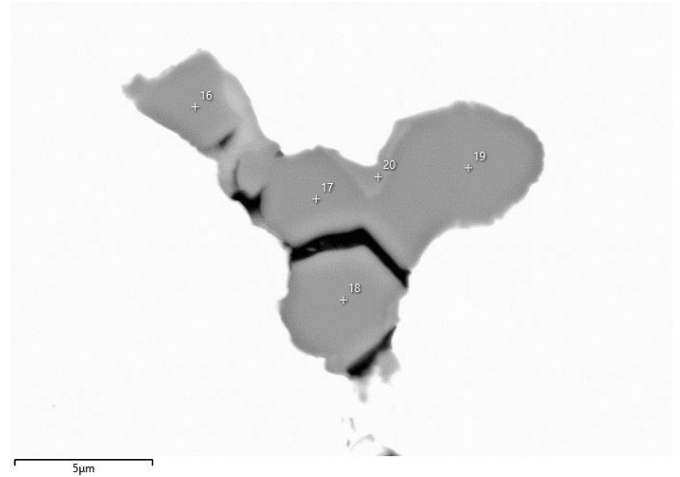


Fig. 12. BSE image of NMI#5, with EDS analyzed points.

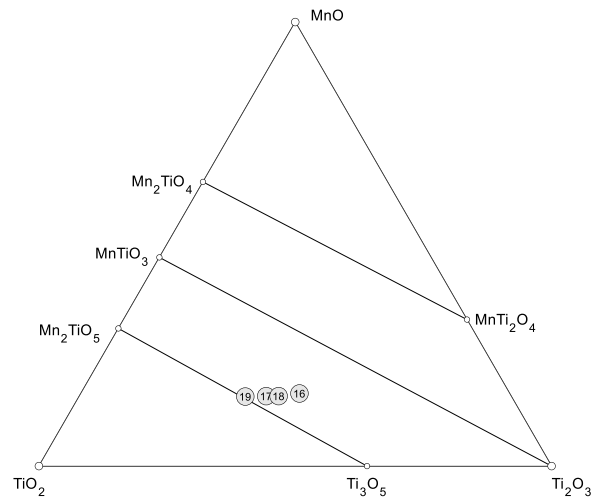


Fig. 13. Estimated oxide compositions of the analyzed points in NMI#5.

### 3.3 Liquid window in the MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> system

Figure 14 presents the colored-line isotherms 1500 °C, 1600 °C, and 1700 °C for all-liquid regions in the MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> oxide system. At 1500 °C, the liquid window is located between the MnO and TiO<sub>2</sub> compositions, extending slightly towards the Ti<sub>2</sub>O<sub>3</sub> corner. Increasing the temperature to 1600 °C and 1700 °C, the all-liquid region expands notably. The univariant phase regions, separated by black lines, show the first-precipitating phases when an all-liquid slag with a certain composition starts to solidify. The chemical composition of the presented spherical NMIs corresponds to pseudobrookite phase, fully molten at 1600 °C.

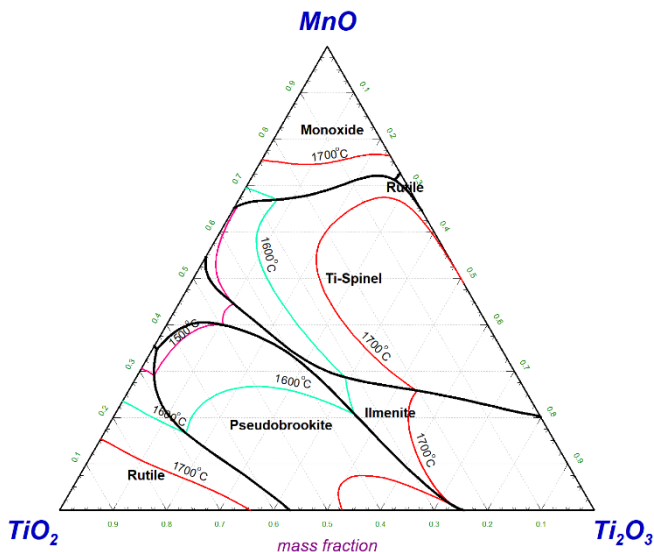


Fig. 14. All-liquid isotherms and univariant phases in the MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> oxide system.

### 3.4 Thermodynamic stability of NMIs

Figures 15 and 16 present the mass of thermodynamically stable phases as a function of decreasing temperature, depicting the solidification of steel. As the total mass of the system is constant 100 g, the mass corresponds to weight percentages in the system. The results show some similarities because of the comparable steel compositions. The slight variations originate from the differences in aluminum, titanium, oxygen, and nitrogen contents.

For both samples, all NMIs are as fully liquid slag until around 1650 °C. Excluding steel, the first solid solution to precipitate is pseudobrookite in both cases. As the solidification of steel progresses, titanium nitride TiN and manganese sulfide MnS start to form as well. As the temperature further decreases, a transformation of oxides can be observed: first pseudobrookite to corundum and ilmenite, then towards spinel and olivine stability. In practice, kinetics constrain the reactions in the solid state, thus the equilibrium at lower temperatures is never achieved. Comparing the calculations to the observed NMI#2 presented in Fig. 6, the results are in line. The core of the NMI is pseudobrookite, since it is the first precipitating solid phase from the fully liquid oxide inclusions. On the periphery of pseudobrookite, spinel phases are observed, further precipitated during solidification.

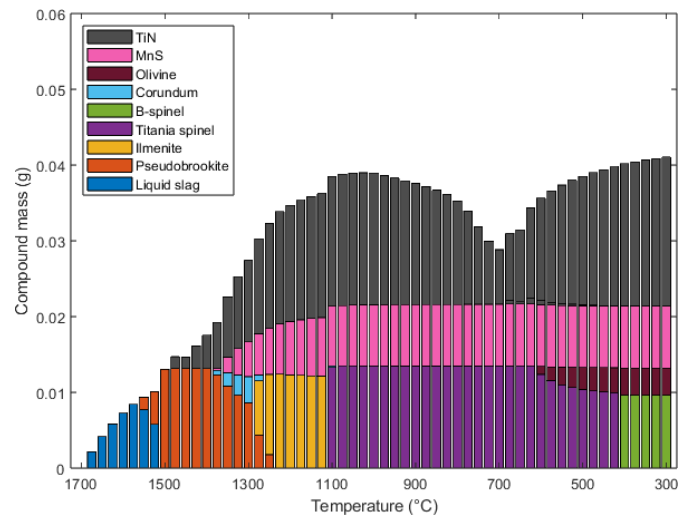


Fig. 15. Mass of stable compounds in Ti<sub>Low</sub> sample.

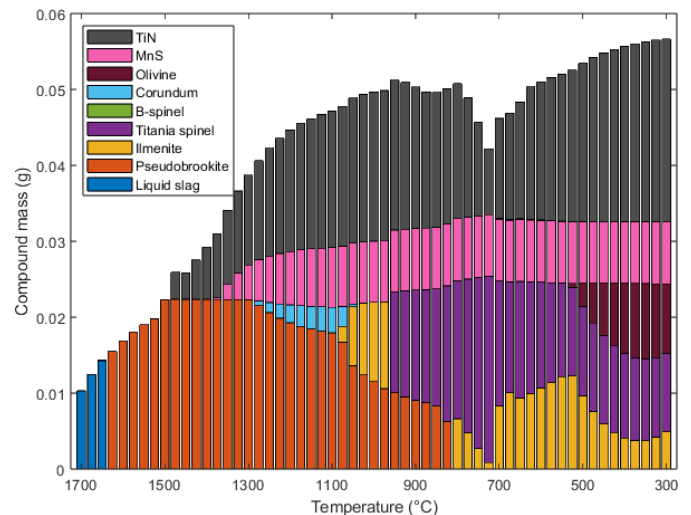


Fig. 16. Mass of stable compounds in Ti<sub>High</sub> sample.

## 4. CONCLUSIONS

A case study of classification and characterization of NMIs in titanium-deoxidized steel samples is presented. The analysis is coupled with computational thermodynamics. Data processing and visualization was executed by MATLAB scripts to automate the procedures. The presented approach allows the data processing and visualization of the total NMI characteristics in the scanned area, as well as for individual NMIs.

Utilizing the presented methods, the sample-specific NMI types were presented, and for individual NMIs, oxide phases in the MnO–TiO<sub>2</sub>–Ti<sub>2</sub>O<sub>3</sub> system were identified. The primary oxide phases were pseudobrookite, a finding supported by the thermodynamic calculations. The spherical shape and homogeneous composition of pseudobrookite NMIs suggest they have been completely liquid in the molten steel. Spinel and TiN phases were observed as precipitated phases on the surface of pseudobrookite NMIs.



## ACKNOWLEDGEMENTS

This research is a part of the CLEAN2STEEL project that benefits from the financial and strategic support of the Kvantum Institute, University of Oulu. H.T. and A.K. are grateful to the funding of the research program FOSSA II (Fossil-Free Steel Applications: Phase II, Dnro 5562/31/2023) funded by Business Finland. SSAB Europe Oy is acknowledged for the provision of the studied materials.

## REFERENCES

- Alatarvas T., Antola, T., and Fabritius, T. (2019). Utilizing inclusion data in characterization of oxide-sulfide stringers in hot-rolled plates. *IOP Conference Series: Materials Science and Engineering*, 529(1), 012066.
- Bale, C.W., Chartrand, P., Degterov, S.A., Eriksson, G., Hack, K., Mahfoud, R.B., Melançon, J., Pelton, A.D., and Petersen, S. (2002). FactSage Thermochemical Software and Databases. *Calphad* 26(2), 189–228.
- Loder, D., Michelic, S.K., and Bernhard, C., (2016). Acicular ferrite formation and its influencing factors - a review. *Journal of Materials Science Research*, 6(1) 24–43.
- Panda, S.K., Hudon, P., and Jung I.-H. (2019). Coupled experimental study and thermodynamic modeling of the MnO-Mn<sub>2</sub>O<sub>3</sub>-Ti<sub>2</sub>O<sub>3</sub>-TiO<sub>2</sub> system. *Calphad*, 66, 101639.
- Tervo, H., Kajjalainen, A., Javaheri, V., Kolli, S., Alatarvas, T., Anttila, S., and Kömi, J. (2020). Characterization of Coarse-Grained Heat-Affected Zones in Al and Ti-Deoxidized Offshore Steels. *Metals*, 10, 1096.
- Tervo, H., Kajjalainen, A., Javaheri, V., Ali, M., Alatarvas, T., Mehtonen, M., Anttila, S., and Kömi, J. (2021). Comparison of Impact Toughness in Simulated Coarse-Grained Heat-Affected Zone of Al-Deoxidized and Ti-Deoxidized Offshore Steels. *Metals*, 11, 1783.
- Wang, B. and Li, J. (2021) Effect of Mn Content on the Characteristics of Inclusions in Ti-Containing Steel Welds. *Metals and Materials International*, 27(8), 2656–2665.
- Wang, B., Liu, X., and Wang G. (2018). Inclusion Characteristics and Acicular Ferrite Nucleation in Ti-Containing Weld Metals of X80 Pipeline Steel. *Metallurgical and Materials Transactions B*, 49(6), 2124–2138.
- Wartiainen A.-M., Harju, M., Tamminen, S., Määttä, L., Alatarvas, T., and Röning, J. (2020). A tool for finding inclusion clusters in steel SEM specimens. *Open Engineering*, 10(1), 642–648.

# Cellular automata model for austenite formation and grain growth during heating and holding above austenization temperature

Aarne Pohjonen.\* Oskari Seppälä.\*\* Olli Väinölä.\*\*  
Jari Larkiola\*\*\*\*

\* *University of Oulu, Pentti Kaiteran Katu 1, Materials and mechanical engineering department (e-mail: Aarne.Pohjonen@Oulu.fi).*

\*\* *e-mail: Oskari.Seppala@Oulu.fi*

\*\*\* *e-mail: Olli.Vainola@Ouraring.com*

\*\*\*\* *e-mail: Jari.Larkiola@Oulu.fi.*

**Abstract:** Understanding the steel microstructure formation during thermal treatments is crucial for controlling the mechanical properties of a steel product. One of the important factors affecting the subsequent microstructure development is the austenite grain size. To gain understanding of the effect of temperature dependent nucleation and growth rates, as well as providing the tools for quantitatively control the austenite grain size distribution, we have implemented a cellular automata (CA) model for describing austenite nucleation and growth during heating, as well as austenite grain growth during holding in temperatures above the austenitization temperature. The model implementation is based on previous study of Sieradzki and Madej for grain growth during recrystallization now augmented with the relevant equations for describing the austenite nucleation and growth. The model parameters and their effect on austenite grain size distributions are tested with numerical experiments. The developed computational tool will serve as a basis that can be parameterized with experimental data in the future, which will then enable quantitative predictions for austenite phase transformation and grain size development.

*Keywords:* cellular automata, austenization, numerical modeling, steel, grain size

## 1. INTRODUCTION

The mechanical properties of metals and alloys depend on their chemical composition and microstructures. During processing, various mechanisms drive microstructural changes, including recovery, recrystallization, and grain growth. Accurately predicting and controlling these evolutions is vital for achieving desired mechanical properties. By using models, engineers and designers can optimize processing parameters (e.g., temperature, deformation, strain rate) to achieve desired microstructures. This optimization leads to improved mechanical properties, such as strength, ductility, and toughness.

During processing of steel, there are several temperature regimes, where different phenomena take place. Sometimes it is of interest to simulate heating and austenization, which is employed to produce suitable austenite grain structure and to facilitate formation of suitable precipitates to enhance the mechanical properties. Naturally, also for forming operations it is also of interest to austenitize the material, which is easier to deform at austenitic state. During cooling, austenite decomposes to ferritic phases. However, in the current study, we restrict to the formation of austenite during heating, and the grain growth of austenite, indicated in Fig. 1.

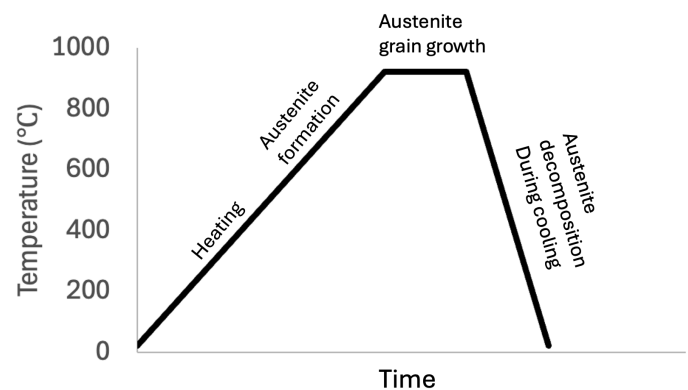


Fig. 1. Schematic illustration of thermal treatment of steel including austenite formation and grain growth stages, which are examined in this work.

Different types of models can be used for simulating microstructure evolution. For example mean field models Pohjonen et al. (2018); Seppälä et al. (2023) provide rapid calculation method which can be coupled with macroscopic heat conduction calculations Pohjonen et al. (2021). Full field models, such as phase field Loginova (2003); Pohjonen (2023), level-set Hallberg (2011) and

cellular automata Sieradzki and Madej (2013) allow for simulating the actual microstructure topology evolution.

Cellular Automata (CA) is the method of choice in this study due to its ability to simulate meso-scale simulation areas with a relatively low computing cost. It is a method which can be used to simulate microstructure evolution during recrystallization. CA is capable of representing topological features and realistically reflecting grain boundary migration. It allows for accurate predictions of grain size distribution and texture evolution. Zhu et al. (2020)

In previous studies by authors, CA has been used to simulate phase transformation from austenite to bainite and martensite. The base model for those phase transformations has been described in Seppälä et al. (2018), martensite growth has been studied in more detail in Kaijalainen et al. (2019), bainite growth has been studied in more detail in Seppälä et al. (2021) and the bainite growth model has been utilized in Seppälä et al. (2023), where in-situ SEM experiments were conducted to study the growth of bainite during cooling.

In the current study, phase transformation from ferrite to austenite during heating will be simulated with another CA model. It is based on Sieradzki and Madej (2013), where static recrystallization and grain growth are simulated. The equiaxial grain growth and grain boundary development algorithms are utilized. The equations for nucleus formation and driving force for growth are modified to suit phase transformation instead of recrystallization.

## 2. THEORY, MODELLING AND SIMULATION

In this section we describe the theory and approximations of the models for austenite nucleation from initial ferrite during heating, the growth of the austenitic regions to the surrounding ferrite and the evolution of the austenitic grain structure due to interaction of the neighbouring distinct austenitic grains. The nucleation model is based on the classical nucleation theory Porter and Easterling (2022) with using simplifications that allow the model to be used with only few fitting parameters Kirkaldy (1983); Luukkonen et al. (2023). The austenite growth is described as a thermally activated process. In the current study the diffusion effects are not explicitly included in the model, but both the nucleation and growth rate are limited by the equilibrium constraints, which dictate the temperature where austenite can form and the maximum austenite fraction.

### 2.1 Austenite nucleation and growth in ferrite

During heating, austenite regions nucleate in the initial ferritic phase. Since it is energetically favourable for austenite to nucleate at certain sites, such as defects, ferrite interfaces, high-carbon regions, etc., the nucleation during heating is heterogeneous. The heterogeneous nucleation rate  $N_{\text{het}}$  can be described by the following equation:

$$N_{\text{het}} = \omega C \exp\left(-\frac{\Delta G^*}{RT}\right) \exp\left(-\frac{\Delta G_m}{RT}\right) \quad (1)$$

where  $R$  is the ideal gas constant,  $C$  is the concentration of nucleation sites,  $\omega$  is the attempt frequency for the nucle-

ation,  $\Delta G_m$  is the activation energy per atom for atomic migration between the austenitic and ferritic regions and  $\Delta G^*$  is the energy barrier for nucleation of austenite from ferrite, which is strongly dependent on the temperature relative to the austenite ferrite equilibrium temperature  $T_{\text{eq}}$ . Theoretically, the energy barrier  $\Delta G^*$  can be obtained by calculating the total Gibbs energy change as a function of nucleating austenite region size Porter and Easterling (2022). Such a detailed approach could be viable, if the energy values associated with the nucleus volume and surface, as well as possible misfit strain energies and the effect of local chemical inhomogeneities could be quantitatively estimated. However, as it is difficult to estimate all the effects, we take in the current study the widely used pragmatic approach and replace  $\omega C \exp\left(-\frac{\Delta G^*}{RT}\right)$  in

Eq. (1) with the expression  $A(T - T_{\text{eq}})^a \left(1 - \frac{\chi}{\chi_{\text{max}}(T)}\right)^b$ , where  $\chi$  is the transformed austenite fraction,  $\chi_{\text{max}}(T)$  is the temperature dependent maximum austenite fraction, which can be obtained from thermodynamic databases.

The factor  $\left(1 - \frac{\chi}{\chi_{\text{max}}(T)}\right)^b$  was added to limit the austenite transformation to the maximum fraction that can be transformed in temperature  $T$ . The parameters  $A$ ,  $a$  and  $b$  are kinetic fitting parameters that can be obtained by fitting the model to experimental data. This expression yields

$$N_{\text{het}} = A(T - T_{\text{eq}})^a \left(1 - \frac{\chi}{\chi_{\text{max}}(T)}\right)^b \exp\left(-\frac{\Delta G_m}{RT}\right) \quad (2)$$

which we have implemented in the CA model.

Austenite growth to the surrounding ferrite matrix is simulated as thermally activated process using the Arrhenius type equation

$$v = F \left(1 - \frac{\chi}{\chi_{\text{max}}(T)}\right)^c \exp\left(-\frac{Q_G}{RT}\right) \quad (3)$$

where  $F$  is a constant prefactor containing the attempt frequency and average atomic sites per unit length as well as the length of austenite interface progression due to one transformation event, and  $Q_G$  is the activation energy for growth.

### 2.2 Austenite-austenite grain interface evolution

The evolution of austenite-austenite interfaces, i.e. the grain growth stage, is calculated based on the reference Sieradzki and Madej (2013). When austenite grains have impinged to each other, grain growth occurs to minimize the surface energy. The effect of crystal orientations is neglected in the current model, and in this case, the surface energy minimization becomes dependent only on the surface curvature  $\kappa$ . The speed of the interface between the two austenite grains  $v$  is described by

$$v = \frac{H}{kT} \exp\left(-\frac{Q_b}{RT}\right) \kappa \quad (4)$$

where  $H$  and  $Q_b$  are parameters for fitting the thermally activated, curvature dependent interface mobility.

### 2.3 Cellular automata method

In Cellular Automata (CA), a simulation area is divided into equal-sized, square-shaped cells. Each cell has an individual value, which is manipulated using neighborhood rules, which change the cell value based on user-defined conditions. In this study, CA is used to simulate microstructural evolution of solid state steel, so the cell value describes its phase instance. The CA model essentially reshapes a graphical two-dimensional representation of the steel microstructure.

The user-defined neighborhood rules are the main way to customize the general CA solver. The neighborhood rules, found in Sieradzki and Madej (2013), are applied for ferrite-to-austenite phase transformation and austenite grain growth. Phase transformation starts at nucleation, which can occur anywhere on the simulation area, but grain boundaries and similar high-energy areas have a higher nucleation chance. From the point-of-view of the CA algorithm, nucleation simply means that a single cell of the simulation area is set to an austenitic phase instance. A pseudo-hexagonal rule is used for equiaxial grain growth. The rule works so that every cell checks its neighboring cells for each time step, using the pseudo-hexagonal pattern. Each neighboring cell that belongs to an austenitic phase instance cumulatively increases the total phase transformation pressure of the checked cell, and after the total transformation pressure increases higher than a set level, the cell transforms to the neighboring phase instance.

With this growth algorithm, the new austenitic grains gradually grow freely into the old ferritic microstructure, until the grains collide with each other. At this point, the second neighborhood rule, which controls austenite grain growth, activates. The basic principle for the rule is that boundaries of two grains strive to be straight. For each cell near a grain boundary, the amount of grain cells are calculated in a long Moore (5x5 cells) neighborhood. In a 2D case, this means that there are a total 25 cells in the neighborhood. If more than 15 of those cells belong to the same grain as the checked cell, nothing happens. Otherwise, the total pressure to change into the neighboring grains cumulatively grows, which increases according to the number of neighboring grain cells, and after the total change pressure reaches a set level, the cell transforms to the neighboring grain. When calculated for all cells, this algorithm causes austenite grain boundaries to gradually become more straight and some grains grow at the expense of other, smaller grains, which gradually disappear.

In a time-dependent simulation like CA, time-step is a crucial parameter. The basic principle in choosing the correct time-step is to use the largest possible time-step that still keeps the simulation stable and produces correct results. Using larger time-steps decreases total simulation time, which obviously is a desirable outcome. On the other hand, simulation stability is required to get reliable results, so it becomes a key limiting factor for the time-step. In this article, the main limitation comes from the relation between austenite grain growth algorithm and cell size. If the growth of austenite in a single time-step exceeds the cell size, then the surplus growth goes to waste. In such

a case, the cell size of the simulation would become an artificial limiter for grain growth. The initial time-step in the following simulations has been chosen so that austenite growth will not exceed cell size, and the solver includes a safety check to decrease time-step size as necessary. This makes sure that the growth rate in a single time-step will not become too large.

## 3. SIMULATION EXAMPLES

To test the operation of the mathematical model, we picked some parameter values in the order of magnitude range, that could be expected in realistic cases. In the current study, the parameter values are not fitted to experimental data, but the aim is to demonstrate the operation of the model and the relative effect of nucleation and growth rates on the simulation outcome. We do not consider specific steel in the current study, and the thermodynamic equilibrium temperatures were chosen based on a textbook example Callister and Rethwisch (2000), p. 381. In that example, the equilibrium austenite formation temperature is 727 °C, and the temperature dependent maximum austenite fraction  $\chi_{\max}$  was calculated using the usual lever rule construction Callister and Rethwisch (2000).

Four simulation test cases were performed, where the effect of changing nucleation and growth rates was examined. In all of the simulation cases, the heating rate was defined as 10 °C/s and initial temperature of the simulation was set to 700 °C, which is well below assumed austenitization temperature of 727 °C. Heating was continued for 25 seconds, until the maximum temperature of 950 °C was reached. During the simulation, all the relevant mechanisms operated, nucleation and growth primarily affected the results until the austenite regions became more impinged, and after this the grain growth phenomena had more influence, and it became the dominating effect in the later stages of the simulation.

To demonstrate the operation of the model, the following parameters were used in the simulations:  $\Delta G_m = 170$  kJ/mol,  $Q_G = 140$  kJ/mol,  $F = 25 \times 10^{-5}$ ,  $a = 1$ ,  $b = 1$ ,  $c = 1$ . Parameter  $A$  is dependent on the local probability of nucleation in the initial ferritic microstructure, and it was set to obtain reasonable effect within simulated times. This approach is capable of taking in to account the fact that the nucleation is heterogeneous, i.e. there usually are sites that are more probable for nucleation, as well as including the randomness to the simulation. In simulating more realistic cases, the parameter values will be fitted to data that can be obtained for example from simplified laboratory experiments.

The model was tested for four different cases, scaling the nucleation rate with factor of  $N$  and growth rate with factor of  $G$ . This shows what kind of effect the relative change in nucleation and growth rates have in the formation of the microstructure.

Figure 2 shows overall transformed austenite fraction as function of temperature during heating for the four examined cases. The deflection of the curve at a range of 50-70 % is mainly due to the temperature dependent maximum austenite fraction  $\chi_{\max}$ , which is a realistic

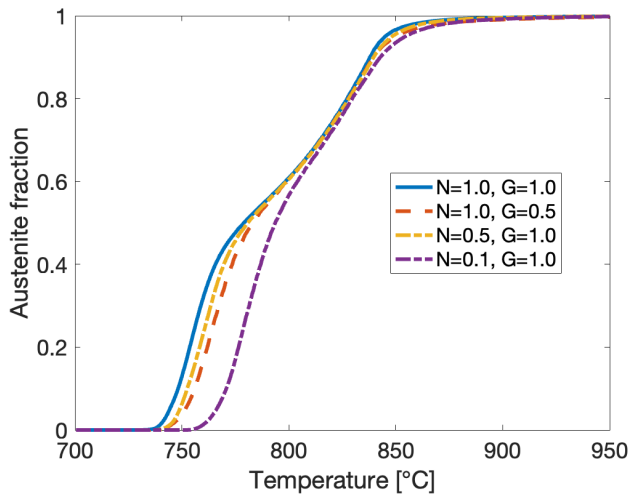


Fig. 2. Overall transformed austenite fraction as function of temperature during heating for the four examined cases where the nucleation and growth rate were scaled with parameters  $N$  and  $G$  respectively.

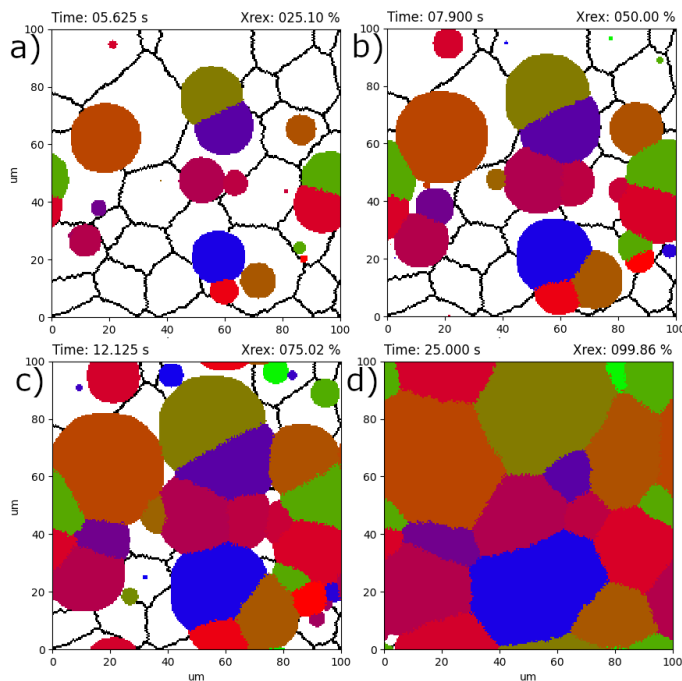


Fig. 3. The baseline simulation with scaling parameters  $N = 1$  and  $G = 1$ , where the other simulation cases can be compared. Simulation snapshots where a) 25 %, b) 50 % and c) 75 % austenite has formed and d) at the end of the heating stage  $950^\circ$

effect similar to that shown in Savran et al. (2010). Figure 3 shows the development of austenite grain structure during heating for the case 1 ( $N=1$ ,  $G=1$ ). This case serves as the baseline where the other simulation cases can be compared.

Figure 4 shows the development of austenite grain structure during heating for the case 2 ( $N=1$ ,  $G=0.5$ ). This case shows that when the growth rate is halved, there is much

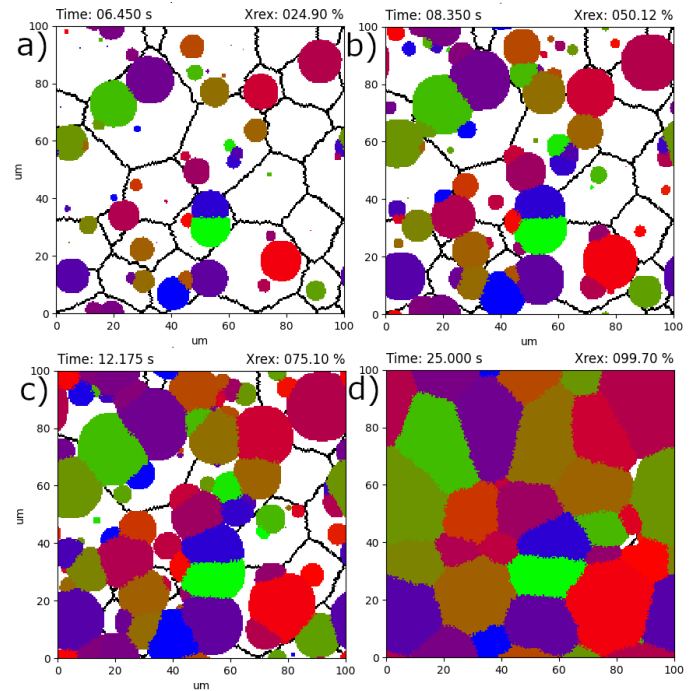


Fig. 4. The development of austenite grain structure during heating for the case 2 ( $N=1$ ,  $G=0.5$ ). a), b), c), d) same fractions as in Fig. 3

more available nucleation sites at higher temperature, and the resulting austenite structure becomes more refined.

Figure 5 shows the development of austenite grain structure during heating for the case 3 ( $N=0.5$ ,  $G=1$ ). This case shows not much difference to the baseline case (case 1). The reason appears to be that the growth rate has more effect to the nucleation by removal of the available nucleation sites than actually halving the nucleation rate.

Figure 6 shows the development of austenite grain structure during heating for the case 4 ( $N=0.1$ ,  $G=1$ ). This case shows that when the nucleation rate is more drastically diminished, it has significant effect to the resulting austenite structure, as less austenite grains are formed and they grow bigger.

#### 4. CONCLUSIONS AND OUTLOOK

To form a basis for future research efforts relating to austenite formation during heating and grain growth during holding at high temperature, it was desired to construct a model that includes the most important effects pertaining to these processes.

The nucleation and growth rates for austenite formation during heating are based on thermally activated processes, taking account the temperature dependent variation of the net energy gain due to the transformation, i.e. the driving force, by using suitable parameterized functions. The implemented model for grain growth during holding at high temperature is based on the earlier article presented in ref. Sieradzki and Madej (2013).

The model equations were used in the cellular automata model implementation. The model operation was demon-



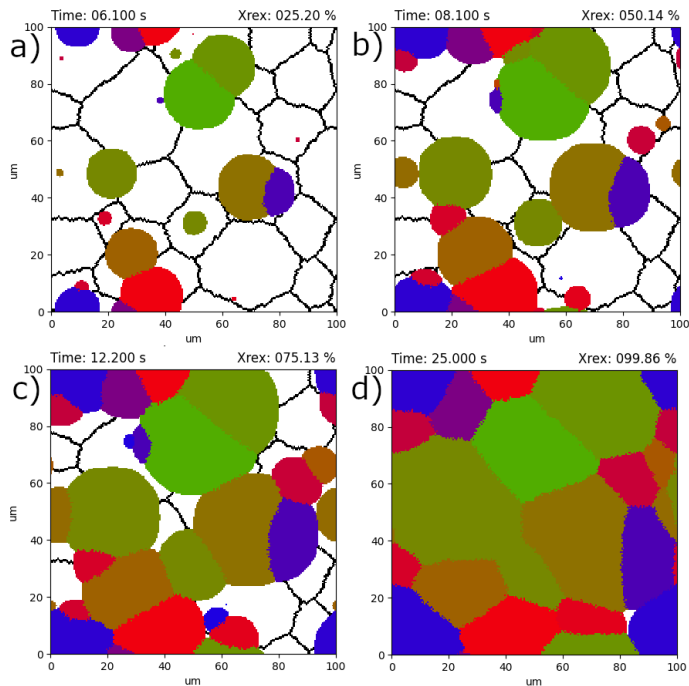


Fig. 5. The development of austenite grain structure during heating for the case 3 ( $N=0.5$ ,  $G=1$ ). a), b), c), d) same fractions as in Fig. 3

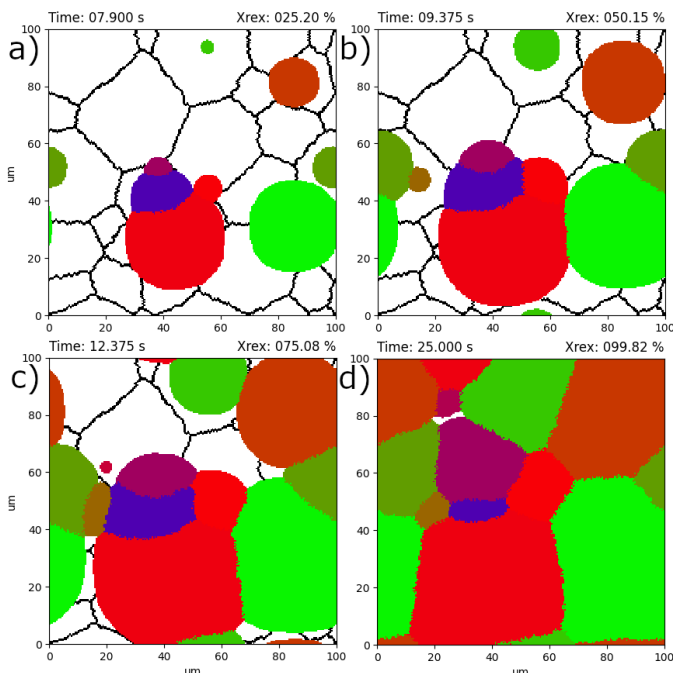


Fig. 6. The development of austenite grain structure during heating for the case 4 ( $N=0.1$ ,  $G=1$ ). a), b), c), d) same fractions as in Fig. 3

strated with four examples by altering the nucleation and growth rate. The purpose of the current article is to present the theory and the numerical modelling aspects pertaining to the examined phenomena. In future, the model will be parameterized using experimental data and compared to experimental microscopy results.

## ACKNOWLEDGEMENTS

Partial funding from AID4GREENEST project is gratefully acknowledged: This research has received funding from the European Commission under the European Union's Horizon Research and Innovation programme (Grant Agreement No. 101091912). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

## REFERENCES

- Callister, W.D. and Rethwisch, D.G. (2000). *Fundamentals of materials science and engineering*, volume 471660817. Wiley London.
- Hallberg, H. (2011). Approaches to modeling of recrystallization. *Metals*, 1(1), 16–48. doi:10.3390/met1010016. URL <https://www.mdpi.com/2075-4701/1/1/16>.
- Kaijalainen, A., Seppälä, O., Javaheri, V., Pohjonen, A., Porter, D., and Kömi, J. (2019). Comparison between experimental data and a cellular automata simulation of martensite formation during cooling. In *Journal of Physics: Conference Series*, volume 1270, 012035. IOP Publishing.
- Kirkaldy, J. (1983). Prediction of microstructure and hardenability in low alloy steels. In *Proceedings of the International Conference on Phase Transformation in Ferrous Alloys, 1983*. AIME.
- Loginova, I. (2003). *Phase-field modeling of diffusion controlled phase transformations*. Ph.D. thesis, Kungliga tekniska högskolan.
- Luukkonen, J., Pohjonen, A., Louhenkilpi, S., Miettinen, J., Sillanpää, M.J., and Laitinen, E. (2023). Gradient boosted regression trees for modelling onset of austenite decomposition during cooling of steels. *Metallurgical and Materials Transactions B*, 54(4), 1705–1724.
- Pohjonen, A. (2023). Full field model describing phase front propagation, transformation strains, chemical partitioning, and diffusion in solid–solid phase transformations. *Advanced Theory and Simulations*, 6(3), 2200771. doi:10.1002/adts.202200771.
- Pohjonen, A., Kaikkonen, P., Seppälä, O., Ilmola, J., Javaheri, V., Manninen, T., and Somani, M. (2021). Numerical and experimental study on thermo-mechanical processing of medium-carbon steels at low temperatures for achieving ultrafine-structured bainite. *Materialia*, 18, 101150.
- Pohjonen, A., Paananen, J., Mourujärvi, J., Manninen, T., Larkiola, J., and Porter, D. (2018). Computer simulations of austenite decomposition of microalloyed 700 MPa steel during cooling. *AIP Conference Proceedings*, 1960(1), 090010. doi:10.1063/1.5034936.

- Porter, D.A. and Easterling, K.E. (2022). *Phase transformations in metals and alloys (4th edition)*. CRC press.
- Savran, V., Offerman, S., and Sietsma, J. (2010). Austenite nucleation and growth observed on the level of individual grains by three-dimensional x-ray diffraction microscopy. *Metallurgical and Materials Transactions A*, 41, 583–591.
- Seppälä, O., Pohjonen, A., Kaijalainen, A., Larkiola, J., and Porter, D. (2018). Simulation of bainite and martensite formation using a novel cellular automata method. *Procedia Manufacturing*, 15, 1856–1863.
- Seppälä, O., Pohjonen, A., and Larkiola, J. (2021). Effect of anisotropic growth and grain boundary impingement on bainite transformation models. *Proceedings of The 61st SIMS Conference on Simulation and Modelling SIMS 2020, September 22-24, Virtual Conference, Finland*. doi:10.3384/ecp20176146.
- Seppälä, O., Pohjonen, A., Mendonça, J., Javaheri, V., Podor, R., Singh, H., and Larkiola, J. (2023). In-situ sem characterization and numerical modelling of bainite formation and impingement of a medium-carbon, low-alloy steel. *Materials & Design*, 230, 111956.
- Sieradzki, L. and Madej, L. (2013). A perceptive comparison of the cellular automata and monte carlo techniques in application to static recrystallization modeling in polycrystalline materials. *Computational Materials Science*, 67, 156–173.
- Zhu, H., Chen, F., Zhang, H., and Cui, Z. (2020). Review on modeling and simulation of microstructure evolution during dynamic recrystallization using cellular automaton method. *Science China Technological Sciences*, 63(3), 357–396.

# Non-interacting lattice random walks for calculating diffusion controlled growth in solid state for dilute concentrations

Aarne Pohjonen\* Touko Puro\*\*  
Assa Aravindh Sasikala Devi\*\*\*

\* *University of Oulu, Materials and mechanical engineering department  
(e-mail: Aarne.Pohjonen@Oulu.fi).*

\*\* *Aalto University, Department of Computer Science, Konemiehentie  
2, 02150 Espoo (e-mail: Touko.Puro@Aalto.fi)*

\*\*\* *University of Oulu, (e-mail: Assa.Sasikala.Devi@Oulu.fi)*

---

**Abstract:** To connect the molecular length scale phenomena to the macroscopic length scale in diffusion controlled growth in solid state, there is need to consider the movement of individual atoms in the crystal lattice and examine the length scale effect where the average density of the atoms approaches to the continuum macro scale. For this purpose a lattice random walk model has been constructed to represent the diffusion of atoms to form a precipitate. Once the atom is in contact with the precipitate surface, the precipitate grows and the atom is not anymore contributing to the random walk. Through the model, it is possible to evaluate the concentration fluctuations at different length scales in diffusion controlled growth and connect the continuum description of diffusion to the atomic level description. We connect the different length scales in theoretical description from atomistic scale through random atom movements to macroscale. In the current study, two-dimensional lattice random walks and growth are considered. The study contributes to the modelling efforts of understanding diffusion controlled precipitate growth in steels.

*Keywords:* Diffusion, random walk, scale bridging, atom level, continuum level, random movement probability, neb method, saddle points.

---

## 1. INTRODUCTION

Diffusion is an important phenomenon, which affects drastically the formation of several microstructural features in materials, such as formation of precipitates (Pohjonen et al., 2022), movement of phase boundaries (Pohjonen 2023), segregation of atoms to crystal defects (Cottrell and Bilby, 1949; Macchi et al., 2024; Pohjonen et al., 2022), etc. Therefore, it is of utmost practical importance to obtain the highest possible level of thorough understanding of the phenomena that affects the diffusion in atomic lattice in different conditions.

Traditionally, diffusion in macroscopic scales can be modelled using the Fick's laws of diffusion (Porter et al., 2022). The connection of the probability of molecular movements and their macroscopic effects dates back to Brownian motion and the theoretical explanation by Einstein in ref. (Einstein (1905, 1906) and Smoluchowski in ref. (von Smoluchowski 1906) as described in (Kac, 1947). Previously, the diffusional growth of a precipitate in steels has been examined through random movements of atoms in the atomic lattice. (Larsson and Ågren, 2003) Also the activation energy that relates to the atomic movement in steels has been calculated for austenitic and ferritic/martensitic structures (Wang et al., 2021). In the current study, we describe the theory connecting these different length scales through the implementation of a

random walk algorithm for diffusional atom movement in the atomic lattice. This approach provides initial step for bridging the atomistic energy, length, and time scales to macroscopic description, and it provides information on the transition, where the inherent concentration fluctuations in atomic scale diminish when increasing the length scale.

## 2. THEORY

The basic connection between flux of atoms and the random movement is reasonably straightforward. (Porter et al., 2022) Consider neighbouring planes of atomic sites containing diffusing atoms. If the diffusing atom on plane 1 has probability of  $p_{x+}$  to move in positive  $x$  direction to plane 2 and atom on plane 2 has probability  $p_{x-}$  to move in the negative direction to plane 1, then the net flux in  $x$ -direction  $\mathbf{f}_x = \hat{n}_x(p_{x+n_1} - p_{x-n_2})/A$  where  $\hat{n}_x$  is the unit normal vector in  $x$  direction. For simplicity, let us consider the movement of atoms in cubic lattice, where the atom hops from one cube to another with probability  $p$ . If the probability of the random atom movement is independent of direction  $x, y, z$ , one obtains the Fick's first law of diffusion:

$$\mathbf{J} = -D\nabla C \quad (1)$$

where  $C$  is the concentration of atoms, and the temperature dependent diffusion coefficient  $D$  has the connection to the atomic level probability of movement  $p$  through

$$D = \frac{1}{2N} \Gamma a^2 \quad (2)$$

where  $N$  is the dimension (for 2 dimensional diffusion  $N = 2$  and 3 dimensional diffusion  $N = 3$ )  $a$  is the lattice constant and  $\Gamma$  is the frequency of atom movement, which is related to the direction-independent probability  $p$  for atom to move per unit time, and the number of dimensions the atom can move during time-step. (Porter et al., 2022) For example consider  $k$  timesteps. If atom has probability  $p_x$  to move in  $x$  direction and probability  $p_y$  to move in  $y$ -direction during one time-step  $\Delta t$ , then the frequency of the atom movement is  $\Gamma = p_x k / (k \Delta t) + p_y k / (k \Delta t) = (p_x + p_y) / \Delta t$ . If probability  $p_x = p_y = p$ , then  $\Gamma = Np / \Delta t$ . The equations (1) and (2) provide the connection between average random atomic movement in bulk material and the flux of atoms. The time evolution of the concentration field can be obtained from the continuity equation by the divergence of the flux, which yields the Fick's 2nd law:

$$\frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \nabla \cdot (D \nabla C) \quad (3)$$

At the atomistic scale, there is considerable fluctuation of concentration about the average value due to the random movement of atoms, which can be examined with random walk simulations for different cases.

In case the probability of the random atomic movement  $p$  is independent of the position  $\mathbf{x}$ , the time evolution of the concentration field is only affected by the gradient  $\nabla C$ . However, if the probability  $p(\mathbf{x})$  is function of position, the drift (i.e. advection) of atoms occurs according to the following equation:

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C) = \nabla D \cdot \nabla C + D \nabla^2 C \quad (4)$$

which is the advection-diffusion equation, where the advection velocity  $\mathbf{v} = -\nabla D = -\nabla \frac{\Gamma a^2}{2N} = -\nabla \frac{p a^2}{2 \Delta t}$ . Considering the position dependent probability for random atom movements  $p(\mathbf{x})$  and compression/tension of the lattice, it becomes possible to link the atomistic phenomena to the emergent macroscopic diffusion and advection phenomena (Cottrell and Bilby, 1949).

An analytical solution is available for diffusion from point concentration (MIT, 2024), which serves as a useful test case for diffusion models (Pohjonen, 2024b,a), which we shall compare to also in the current study. For two-dimensions, the analytical solution is described by

$$C(x, y, t) = \frac{M}{4\pi Dt} \exp\left(-\frac{x^2 + y^2}{4Dt}\right) \quad (5)$$

where  $M$  is the number of atoms.

The atom movement is thermally activated process, which is caused by the random vibrations of atoms. If the diffusing atom gains enough energy, it has high probability of moving in the lattice. The probability for the atom movement from a stable lattice site to another stable site can be obtained from Arrhenius type equation:

$$p = A \exp\left(-\frac{E_A}{RT}\right) \quad (6)$$

where the activation energy barrier  $E_A$  can be calculated using atomistic simulations using the nudged elastic band method (NEB) (Wang et al., 2021). NEB method is a powerful tool to identify the microstructural evolution of a

system in which defects or impurity atoms are present and they evolve interactively (Jónsson et al., 1998; Henkelman et al., 2000, 2002) The atomic scale information such as the energies of the initial, final and transition states, can be used to identify the energy barriers and can serve as inputs to the description of mesoscale phenomena. Basically NEB method can provide a minimum energy path that describe the energy variation of the atomic movement from an initial to final state. It is a chain of states method, to determine the minimum energy path on the potential energy surface. Each atomic configuration will be at a potential energy of 0 K, represented by a point in the configuration space, and can be determined either by empirical potentials or first principles calculations. In the NEB method the initial and final configurations will be calculated by minimizing the energy and then a linear interpolation will be carried out between the two end states to generate a finite number of replicas. Two nearby replicas will be connected by a spring, resembling an elastic band made of beads and springs. To solve the corner cutting and sliding that can arise, a force projection, such as "nudging" is employed. This procedure followed by proper optimization ensures that the elastic band converges to the minimum energy path. Further, after optimization, both the position and energy information of the configurations can be obtained. There are different variations of the basic NEB method, adapted to suit the needs of the system in use, such as extended three dimensional defects which requires a large model system with a long reaction path. This is to ensure that enough replicas are included to map the long trajectory between the saddle point and final state (Zhu et al., 2007).

The effect of elastic lattice distortions can bias the random movement of the atoms and give rise to net drift of interstitial atoms towards tensile stress and away from compression. The dependence of the random movement probability on the local strain can be quantitatively examined with ab-initio based NEB methods, and the emerging flux and the random fluctuations can be examined with the random walk simulations.

In certain temperature range, it is energetically favourable for the atoms to coalesce and form a precipitate, which then grows due to diffusion of more atoms to the surface of the precipitate (Larsson and Ågren, 2003; Pohjonen et al., 2022).

### 3. NUMERICAL ALGORITHM

A random walk algorithm was implemented to simulate the random movements of atoms in two-dimensional lattice, and their coalescence to a precipitate, which is located in the center of the simulation domain. When the atom coalesces to the precipitate surface, the precipitate radius will grow and the coalescing atoms are removed from the random walk simulation. As a result of this, the concentration is lowered near the precipitate, which then implicitly causes net flux of atoms towards the negative of the concentration gradient. Periodic boundaries were applied in the simulation.

The random walk algorithm has two main stages. First, it moves the atoms based on the random probability of movement, and secondly, it will remove the atoms from calculation when they are within the precipitate radius,

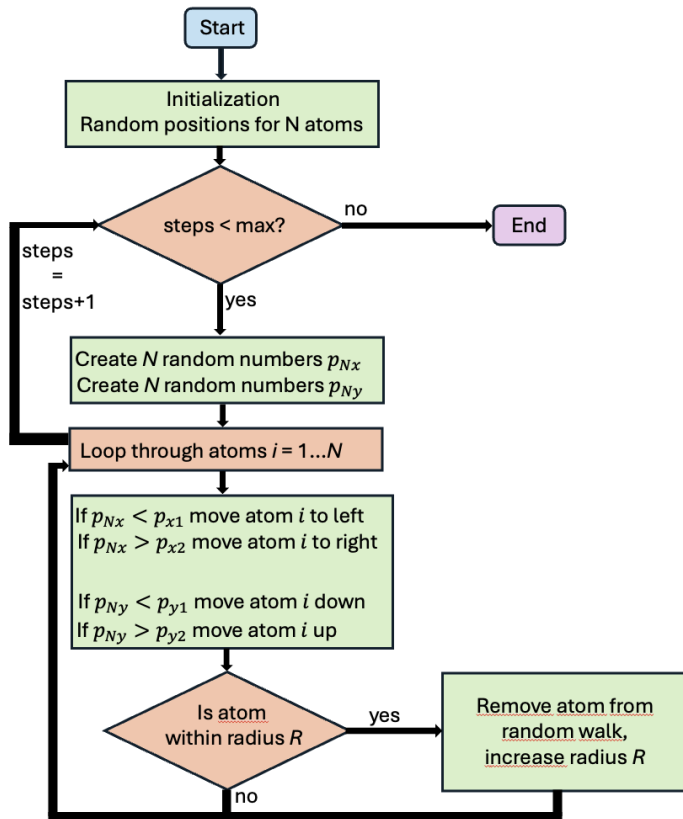


Fig. 1. Flow chart depicting the operation of the random walk and precipitate growth algorithm.

which grows each time an atom is located there within. A flow chart depicting the operation of the algorithm is shown in Fig. 1.

The local density of atoms was calculated by dividing the simulation domain to rectangular subdomains and calculating the number of atoms within each subdomain, and dividing the number of atoms by the size of the subdomain (area in 2D).

To speedup the simulations we accelerated the random walk algorithm to run on GPUs by writing our algorithm directly in CUDA. In the following, we refer readers unfamiliar with CUDA and GPU terminology to (NVIDIA, 2024). The random walk algorithm is straightforward to parallelize and accelerate, because except for growing the precipitate radius the atoms do not affect each others. In the performed simulations it was found that atoms coalescing to the precipitate surface is rare enough that performance-wise the radius expansion can be performed with GPU atomic operations.

Atomic operations guarantee that no other atomic operations happen while the atomic operation is being performed. Thus they prevent race conditions that would otherwise happen when multiple threads try to write to the same variable. Naturally atomic operations are more costly than normal operations, especially if many of them are being performed on the same memory address, since the operations have to be synchronized in some manner.

To minimize the costs of the atomic operations we use a common technique of first combining all atomic updates

inside a threadblock into a variable in shared memory and make a single atomic update from the threadblock into the variable in global memory that is shared across all threads. This cuts down the amount of atomic operations to variables in global memory significantly and atomic operations to variables in shared memory are significantly faster due to them being closer in memory and due to the need of synchronizing only between the threads in the current threadblock.

Due to this optimization and the radius updates being rare we found the atomic operations to be a solution that is easy to implement for growing the radius. This implementation did not have measurable effect on the performance of the kernels. The rarity of the coalescing motivated optimization, where instead of synchronizing between a single update of each atom, we update each atom  $n$  steps and then synchronize. This cuts down memory traffic by a factor of  $n$  since we can reuse the values loaded to local memory. The only difference between synchronizing between each step and every  $n$  steps is that the radius size lags after the first update done after synchronizing in the second scheme, but if the number of radius updates is small enough this difference is negligible. We found that on the tested hardware, a single RTX A2000 8GB Laptop GPU, synchronizing after every second step gave a performance improvement of 30 percent.

## 4. NUMERICAL TEST CASES

### 4.1 Diffusion from initial point source

To test the connection between the random movements and the macroscopic diffusion equation, we compared the random walk simulation from initial point concentration to corresponding diffusion calculation. Consider a two-dimensional case where all atoms are initially located at the origin. The random movement is assumed to occur in two dimensional square lattice with lattice constant  $a = 1$ . Also, timestep is chosen as  $\Delta t = 1$ , i.e. dimensionless units were used, as this is mathematical study, not directly connected with physical properties. The atoms have 50 % probability to move within time-step in x-direction, and the same probability for y-direction. If they move, they have equal probability to move either in positive or negative direction. Since the diffusing atom can move both  $x$  and  $y$  direction within timestep with 50% chance in each direction, the frequency for the atom movement during timestep is  $\Gamma = (0.5 + 0.5)/\Delta t = 1/\Delta t$ , and according to equation (2) the diffusion coefficient in this case becomes  $D = \Gamma/(2N) = 1/4$ . The simulation results and comparison to analytical continuum solution, equation (5), are shown in Fig. 2. To see the effect of the random fluctuations, the simulation was repeated ten times and the standard deviation for each plotted datapoint was calculated. The standard deviations are indicated as error bars in the figure. The small difference between the simulated results and the analytical solution arises because of the differences between the analytical continuum description and the discrete random walk results, and also due to chosen area where the atoms are averaged over when calculating the results from the simulations.



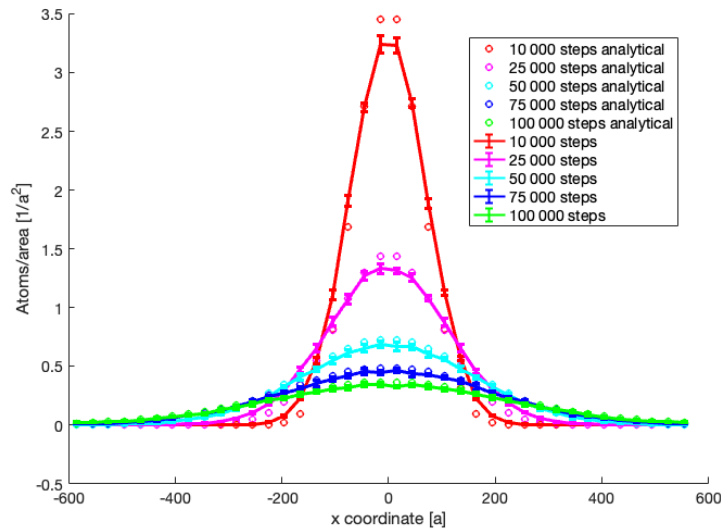


Fig. 2. Diffusion from initial point concentration. The lines show the simulation results averaged over 10 separate simulations and the error bars indicate the standard deviation. The markers show the corresponding analytical solution described by equation (5).

#### 4.2 Concentration fluctuations as function of system size

Fluctuations in small systems are inherent due to random movement of discrete particles. It was examined how the concentration fluctuations in a two dimensional system depend on the system size. To achieve this, the length of the square domain and the number of atoms were both scaled by scaling factor  $sf$ , which was altered in the simulations. The atoms were initialized to random positions and the random walk simulation was ran for 1000 timesteps. In each simulation the domain was divided in to  $40 \times 40$  square subdomains. The 2d concentration was calculated as number of atoms/area for each subdomain. Then standard deviation of the subdomain concentrations was calculated. The standard deviation as function of the scaling factor is shown in Fig. 3. The result shows that the standard deviation depends on the scaling factor approximately proportional to  $1/sf$ . Increasing the system size in the random walk simulations up to  $sf = 12$  still showed noticeable fluctuations, and was not much different from the case  $sf = 6$ .

#### 4.3 Diffusion controlled growth of precipitate

To test the random walk simulation in a more interesting case, a coupled precipitate growth and diffusion simulation was performed. The atoms, which were initially located randomly at the domain, were moved randomly similar to the previous case, but if they arrived within a radius, the precipitate radius grew and the atom was removed from the random walk. The simulation results from small scale simulations are shown in Fig. 4, where a) shows the concentration of atoms in the whole simulated two dimensional domain and b) shows the plot of concentration of atoms along the horizontal line which passes through the origin, where the growing precipitate is located. The area increase of precipitate due to attachment of atom was set as  $A_{at} = \pi(a/6)^2$ .

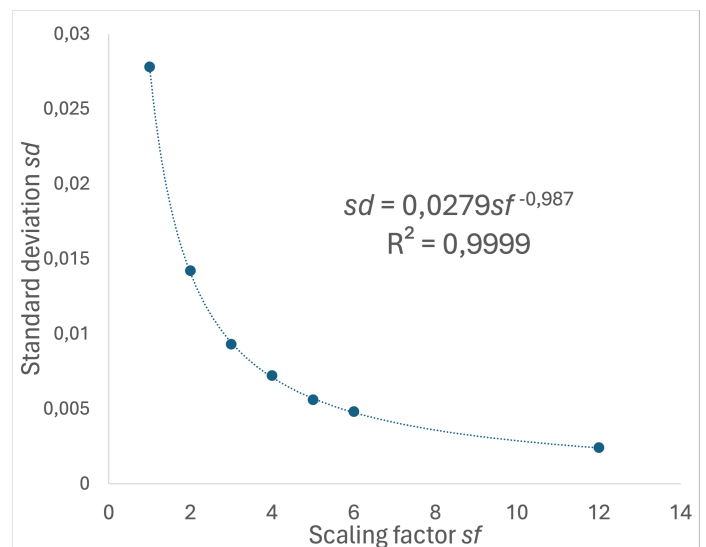


Fig. 3. The standard deviation  $sd$  of the concentrations of 40 by 40 subdomains as function of system scaling factor  $sf$

The result shows that the concentration near the precipitate becomes depleted as atoms are being removed from the diffusion to increase the precipitate radius, which is a realistic effect (Porter et al., 2022).

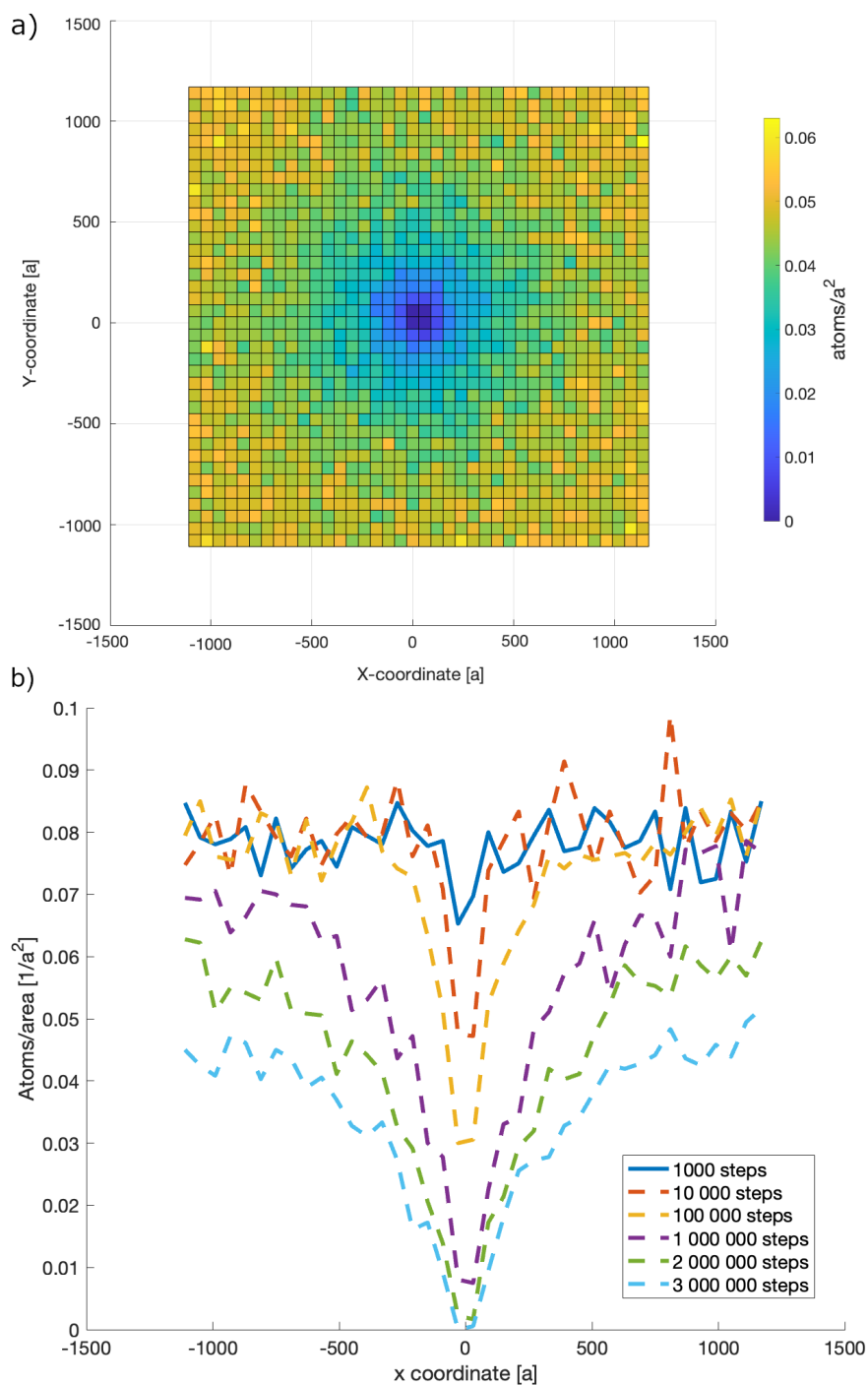


Fig. 4. a) The concentration of atoms per subdomain area after 3 000 000 timesteps, b) Plot of concentration along the horizontal line which passes the origin at several different time-steps.

## 5. CONCLUSIONS

A theory linking random movement (i.e. random walk) of atoms in crystalline material to the activation barrier of atom hopping from stable lattice site to another was described and also the link between the diffusion equation and the random movement probability was presented. The random random walk model describing diffusional movement of atoms in a lattice was constructed and parallelized using GPU. The model was compared to a continuum analytical solution for diffusion from an initial point concentration, and it was applied for mathematical calculation of diffusion controlled growth of a precipitate.

In future studies, the probabilities and their dependence on different factors, such as local stress/strain state, can be obtained from nudged elastic band (NEB) calculations and the model can be parameterized using physical data. The NEB calculations combined with atomistic ab-initio density functional theory simulations have the capability for providing the energy barriers for atomic mobility, and hence they could be used in future for theoretic prediction of the of atomic scale effects of diffusion, precipitation and partitioning phenomena. Such effects are very difficult to observe directly experimentally, which makes phenomenological parameterization of models difficult. The theoretical calculations can also provide explanations connected with these practically very important metallurgical phenomena.

## ACKNOWLEDGEMENTS

Authors acknowledge financial support from the AID4GREENEST project (AI powered characterization and modelling for green steel technology) funded by Horizon Europe Research and Innovation Program, Grant Agreement No. 101091912. We would also like to thank Jane and Aatos Erkkö (J&AE) Foundation and Tiina and Antti Herlin (TAH) Foundation for their financial supports on Advanced Steels for Green Planet project.

## REFERENCES

- Cottrell, A.H. and Bilby, B.A. (1949). Dislocation theory of yielding and strain ageing of iron. *Proceedings of the Physical Society. Section A*, 62(1), 49.
- Einstein, A. (1906). Zur theorie der brownischen bewegung. *Annalen der Physik*, 324(2), 371–381. doi: 10.1002/andp.19063240208.
- Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4.
- Henkelman, G., Jóhannesson, G., and Jónsson, H. (2002). Methods for finding saddle points and minimum energy paths. *Theoretical methods in condensed phase chemistry*, 269–302.
- Henkelman, G., Uberuaga, B.P., and Jónsson, H. (2000). A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of chemical physics*, 113(22), 9901–9904.
- Jónsson, H., Mills, G., and Jacobsen, K.W. (1998). Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and quantum dynamics in condensed phase simulations*, 385–404. World Scientific.
- Kac, M. (1947). Random walk and the theory of brownian motion. *The American Mathematical Monthly*, 54(7P1), 369–391.
- Larsson, H. and Ågren, J. (2003). A random-walk approach to diffusion controlled growth. *Scripta Materialia*, 49(6), 521–526. doi:10.1016/S1359-6462(03)00364-6.
- Macchi, J., Teixeira, J., Danoix, F., Geandier, G., Denis, S., Bonnet, F., and Allain, S.Y. (2024). Impact of carbon segregation on transition carbides and cementite precipitation during tempering of low carbon steels: Experiments and modeling. *Acta Materialia*, 272, 119919. doi:10.1016/j.actamat.2024.119919.
- MIT (2024). Lecture notes on diffusion, accessed 10.6.2024, massachusetts institute of technology. URL <http://web.mit.edu/1.061/www/dream/THREE/THREETHEORY.PDF>.
- NVIDIA (2024). Cuda c++ programming guide. URL [https://docs.nvidia.com/cuda/pdf/CUDA.C\\_Programming\\_Guide.pdf](https://docs.nvidia.com/cuda/pdf/CUDA.C_Programming_Guide.pdf).
- Pohjonen, A. (2023). Full field model describing phase front propagation, transformation strains, chemical partitioning, and diffusion in solid–solid phase transformations. *Advanced Theory and Simulations*, 6(3), 2200771.
- Pohjonen, A. (2024a). Application of two-grid interpolation to enhance average gradient method for solving partial differential equations. *The Journal of Physics: Conference series*, 2701(2), 97–111.
- Pohjonen, A. (2024b). Solving partial differential equations in deformed grids by estimating local average gradients with planes. *The Journal of Physics: Conference series*, 2701(2), 97–111.
- Pohjonen, A., Babu, S.R., and Visuri, V.V. (2022). Coupled model for carbon partitioning, diffusion, cottrell atmosphere formation and cementite precipitation in martensite during quenching. *Computational Materials Science*, 209, 111413.
- Porter, D.A., Easterling, K.E., and Sherif, M.Y. (2022). *Phase transformations in metals and alloys 4th Edition*. CRC press.
- von Smoluchowski, M. (1906). Zur kinetischen theorie der brownischen molekularbewegung und der suspensionen. *Annalen der Physik*, 326(14), 756–780. doi: 10.1002/andp.19063261405.
- Wang, S., Kistanov, A.A., King, G., Ghosh, S., Singh, H., Pallaspuro, S., Rahemtulla, A., Somani, M., Kömi, J., Cao, W., et al. (2021). In-situ quantification and density functional theory elucidation of phase transformation in carbon steel during quenching and partitioning. *Acta Materialia*, 221, 117361.
- Zhu, T., Li, J., Samanta, A., Kim, H.G., and Suresh, S. (2007). Interfacial plasticity governs strain rate sensitivity and ductility in nanostructured metals. *Proceedings of the National Academy of Sciences*, 104(9), 3031–3036.

# On the Growth Kinetics of Lamellar and Blocky Austenite during Intercritical Annealing of Hot-Rolled Medium Manganese Steel: Thermodynamic and Diffusion-Controlled Transformation Simulations

R. Surki Aliabad, S. Sadeghpour, P. Karjalainen, J. Kömi, and V. Javaheri

\* *Materials and Mechanical Engineering, Centre for Advanced Steels Research, University of Oulu Finland (Tel: +358415741507; e-mail: [roohallah.surkialiabad@oulu.fi](mailto:roohallah.surkialiabad@oulu.fi)).*

**Abstract:** Metastable austenite significantly impacts the mechanical properties of Advanced High-Strength Steels (AHSS), especially Medium Mn Steel (MMnS), where its formation rate during intercritical annealing depends strongly on the initial microstructure. This study employs thermodynamic and diffusion-controlled simulations to investigate the formation of two distinct morphologies of retained austenite—lamellar and blocky known also as globular—commonly observed in an intercritically annealed hot-rolled MMnS. Utilizing Thermo-Calc software, coupled with its Diffusion-Controlled TRAnSformation module (DICTRA), phase equilibria are computed, and one-dimensional diffusion equations are solved. Characterization of the microstructure of a medium manganese steel (MMnS) with a nominal composition of Fe–0.4C–6Mn–2Al–1Si–0.05Nb (in wt. %), hot rolled and intercritically annealed for 1 hour at 680°C, was performed using Energy Dispersive Spectroscopy coupled with Transmission Electron Microscopy (EDS-TEM) and Transmission Kikuchi Diffraction (TKD). These techniques were used for experimental validation and verification of the simulations.

Simulations explore the competition between cementite and austenite growth. Specifically, the growth of austenite starting on various interphase boundaries was examined using spherical and planar geometries. This approach resulted in the formation of blocky and lamellar austenite morphologies, respectively. The findings indicate that austenite first nucleates at the BCC/BCC interface and transforms 40% of the BCC phase within 1 second at 680°C. Cementite then starts to form, limiting further austenite transformation. Finally, cementite particles continue to grow to a size of about 100 nm. These simulation results align well with experimental findings.

**Keywords:** Thermo-Calc, DICTRA, lamellar and blocky austenite, intercritical annealing, Medium Mn Steel.

## 1. INTRODUCTION

Over the last decade, the demand for strong yet formable steels using cost-efficient designing approaches, contributing to a more sustainable steel advancement, has driven the development of third-generation AHSS (Guo et al., 2022). Among them, MMnS containing 4-12wt% Mn (Han, 2023; Hu et al., 2024; Sun et al., 2023a), is considered the most promising candidate for future steel owing to its unique combination of strength and ductility (Han, 2023). Since its introduction by Miller (Miller, 1972), various alloy design concepts for MMnS have been developed (Han, 2023; Hu et al., 2024; Sun et al., 2023a).

After hot rolling, MMnS exhibits mainly lath martensitic microstructure along with small amounts of retained austenite depending on the Mn content (Han, 2023). Lath martensite represents the predominant morphology in practical high-strength steels. This structure comprises a hierarchical arrangement: martensite packets, blocks, and laths, (Luo et al., 2020; X. Zhang, Miyamoto, Toji, et al., 2018).

When the initial as-rolled martensitic structure undergoes Intercritical Annealing Treatment (IAT), the martensite shows minimal recrystallization throughout the applied thermal

cycle. Instead, phase transformations predominantly involve formation of cementite or reversion of austenite (Han, 2023). Martensitic, prior austenitic and BCC/FCC boundaries may endure during intercritical annealing, resulting in several possible nucleation sites for new austenite to form. Additionally, in the tempered martensite microstructure, carbides, particularly cementite, can serve as austenite nucleation sites (Sun et al., 2023b). Austenitic nucleation at these boundaries might result from the co-segregation of C and

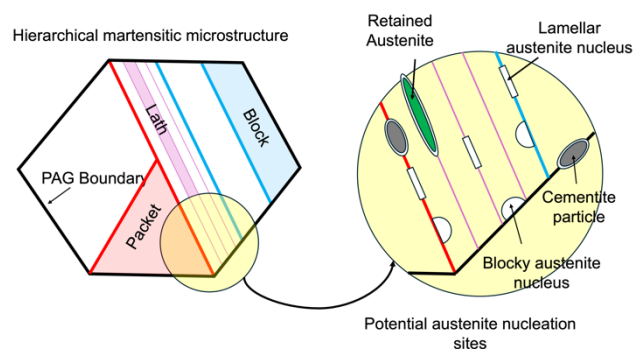


Fig. 1. Illustration of the martensitic hierarchical microstructure and the corresponding preferential austenite nucleation sites.

Mn, which makes them thermodynamically preferred for nucleation (Luo et al., 2011). These nucleation sites can generally be categorized into two groups, as shown in Fig. 1 and described below:

Depending on the location of the nucleus and its orientation with the adjacent martensitic matrix interfaces, two distinct morphologies of austenite may emerge: lamellar and blocky. Lamellar-type austenite typically nucleates at martensite packet, block, or lath boundaries, whereas blocky-type austenite forms at prior austenite grain boundaries, as well as martensite packet and block boundaries (Zhang et al., 2018a). Lamellar austenite grains often display the same orientation as prior austenite grains, which originally form at high temperatures during hot rolling and later coalesce with each other, leading to the reconstitution of the prior austenite grain structure (Han, 2023; Zhang et al., 2018a, 2018b).

Recent multiple-stages MMnS processing routes employ the potential of Mn-rich cementite particles to enhance mechanical properties by providing additional nucleation sites for austenite (Sadeghpour et al., 2021). This tailored austenite differs from conventional formed one, with higher Mn-enrichment and finer size (Ye et al., 2024). Cementite plays complex roles in austenitic nucleation, competing with austenite formation and potentially acting as nucleation sites (Guo et al., 2022). However, not all cementite particles facilitate nucleation; preferential nucleation occurs at intergranular rather than intragranular cementite (Lai et al., 2016). However, Enomoto and Hayashi (Enomoto and Hayashi, 2020) observed a higher fraction of austenite nucleated on cementite existing at prior austenite boundaries and martensite packet boundaries compared to inter-lath boundaries. Additionally, Zhang et al. (Zhang et al., 2021a) suggested that the nucleation potency of cementite particles may depend on their size, with coarser particles expected to have higher potency.

During the IAT, the reverted austenite phase becomes enriched through the partitioning of Mn and C from the surrounding martensite. This gradual enrichment leads to a reduction in the alloy content of nearby martensite, ultimately resulting in its transformation into ferrite. Consequently, after quenching, a significant portion of austenite remains untransformed at ambient temperature (Bhattacharya et al., 2024). Extensive studies have been conducted on various heat treatment parameters such as IA temperature and time (Bansal et al., 2018), as well as tailoring the characteristics of retained austenite (e.g., stability, volume fraction and morphology, etc.) by different IATs (Zhang et al., 2024).

To better understand aspects relevant to intercritical annealing, thermodynamic and kinetic simulations for austenite growth during IAT in the Fe–C–Mn and Fe–C–Mn–Si systems have commonly been conducted. In most of the simulations, a diffusion couple of austenite and martensite was used (Huyan et al., 2018). In this context, the majority of simulations can be summarized into three systems.

#### ***FCC/BCC System***

Assuming local equilibrium, moving-boundary simulations between austenite and ferrite can provide a simplified

representation of austenite growth kinetics. In this system, the diffusional mechanisms of the two phases govern distinct stages of austenite growth. Initially, austenite transformation undergoes rapid increase driven by C diffusion in both ferrite and austenite under negligible partitioning local equilibrium (NPLE) interfacial conditions, with minimal Mn partitioning. Subsequently, a stage dominated by Mn diffusion in ferrite ensues, transitioning from NPLE to partitioning local equilibrium (PLE) interfacial conditions (Sun et al., 2023b). Finally, a slow equilibration of Mn in austenite, controlled by long-range Mn diffusion within austenite, completes the process (Wu et al., 2020).

#### ***BCC/FCC/Cementite System***

Considering cementite during austenite formation alters the interfacial thermodynamics for austenite growth significantly compared to considering only ferrite and austenite. According to Sun et al. (Sun et al., 2023b), at relatively low intercritical temperatures, when cementite is present in the initial microstructure, austenite growth is primarily controlled by Mn diffusion. Conversely, at higher temperatures, C diffusion governs austenite growth. However, if the initial cementite is enriched with more than 30% Mn, austenite growth remains consistently controlled by Mn throughout the entire intercritical range.

#### ***FCC/BCC/Cementite System***

Within this system, austenite nucleates apart from the majority of cementite, with the ferrite matrix serving as a pathway for carbon diffusion. Enomoto and Hayashi (Enomoto and Hayashi, 2018) simulated the growth of austenite during continuous heating in plain low-carbon martensite and reported that due to slow carbon diffusivity in austenite, the cementite free of austenite tends to dissolve faster than the cementite on which austenite was nucleated except when the particle size of cementite and/or the number of austenite nuclei is small.

Despite extensive studies on austenite growth and the chemistry of austenite formation on various interfaces in MMnS (e.g., (Dai et al., 2018; Huyan et al., 2018; Mehrabi et al., 2024; Sun et al., 2023b; Zhang et al., 2018b, 2018a), there remains a lack of comprehensive understanding the instance when cementite, austenite and BCC phases are all present during IAT. Hence, we investigated this coexisting scenario in a hot-rolled sample subjected to IAT at 680°C for 1 hour, where two austenite morphologies with cementite particles were observed by microscopic means, by simulating the phase changes by the DICTRA calculations. Further, this study aims to simulate solute partitioning within these microstructures to elucidate the initial source of cementite particles and compare the growth rate of cementite as well as those of lamellar and blocky austenite.

## 2. MATERIALS AND METHODS

The steel material with the nominal composition of 0.40C, 1Si, 6 Mn, 2Al and 0.05 Nb (wt.%) which employed in this study was produced by OCAS (Onderzoeks Centrum voor de Aanwending van Staal) using a vacuum induction furnace. Following casting, the material underwent austenization at 1200°C for 2 hours, followed by hot rolling to 90 percent



reduction in thickness. The resulting sheet, with a finish rolling temperature of 900°C, was cooled down in ambient air. Subsequently, the sheet underwent intercritical annealing at 680°C for 1 hour.

After standard metallographic sample preparation, the microstructural features of the sample were analyzed using an TKD mappings using a JEOL JSM-7900F Field Emission Scanning Electron Microscope (FESEM) under an accelerating voltage of 30 kV, a working distance of 15 mm, and a step size of 6 nm. For more comprehensive characterization, elemental analysis of the specimens was carried out using Energy Dispersive x-ray Spectroscopy (EDS) integrated into a 200-kV Energy Filtered Scanning Transmission Electron Microscope (JEOL JEM-2200FS EFTEM/STEM). All of TKD and EDS observations were conducted on a thin lamella sample (8×10 μm) prepared using a Focused Ion Beam-Field Emission Scanning Electron Microscope (FIB-FESEM).

In this study, the DICTRA module of Thermo-Calc 2024a was utilized to model the kinetics of austenite growth, encompassing both lamellar and blocky morphologies with various nucleation sites outlined in Fig. 1. For the simulations, thermodynamic (TCFE7) and mobility (MOBFE2) databases for one-dimensional analyses were used. Establishing the simulations necessitated specifications regarding morphology, compositions, and phase sizes. The diffusion of C, Mn, Si, and Al was factored into the simulation framework. The dimensions of the simulation system were determined based on observed length scales within TKD micrographs obtained from the intercritical annealed sample. The chemical composition of each microconstituent was estimated using EDS method excluding for C.

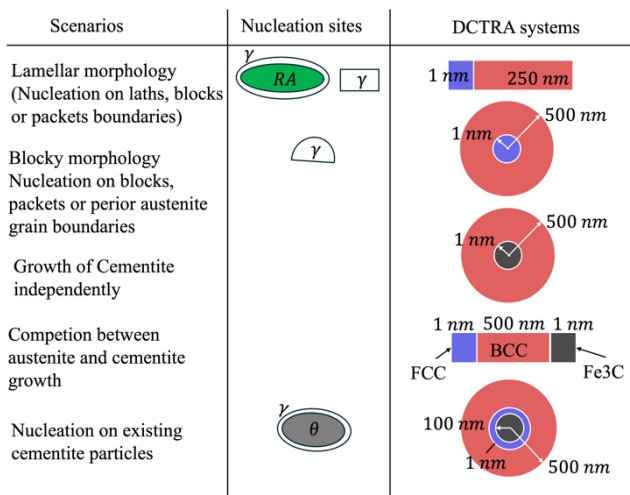


Fig. 2. Schematic illustration of DICTRA simulation configurations based on potential austenite nucleation sites (symbols are described in Fig. 1).

For simulating lamellar and blocky type austenite, planar and cylindrical geometries were respectively adopted. As martensite is not treated as a distinct phase in DICTRA, it is represented similarly to other studies in the literature (Mehrabi et al., 2024), employing a BCC phase with high C and Mn levels. Different simulation configurations based on possible

nucleation sites were considered for the DICTRA, as shown in Fig. 2.

### 3. RESULTS AND DISCUSSION

To model the potential DICTRA systems shown in Fig. 2, the microstructure of the studied steel after the soaking at 680°C for 1 hour were analyzed. The selected timeframe was determined based on our simulations, which suggested that beyond this point, negligible changes in the chemical composition of the solute atoms, particularly Mn, could happen. EBSD-TKD method were employed to obtain the phase map of the microstructure, aiming to discern the distribution and shape of phases and cementite particles. The outcomes of the TKD analysis are depicted in Fig. 2. In this figure, three regions of interest representing the entire microstructure were identified. Fig. 2(b) clearly shows a lamellar microstructure, with the austenite lamellae approximately 200 nm in width and the BCC phase lamella with widths of 300 nm. The half-width dimensions as a preliminary estimation of the system size were used. Additionally in Fig. 3(c), a typical illustration of blocky austenite morphology is depicted. This structure is encompassed by adjacent BCC grains, forming an outer circle with a radius of 500 nm and an inner circle with a radius of 300 nm. Essentially, around 60% of the BCC phase transforms into austenite within this system.

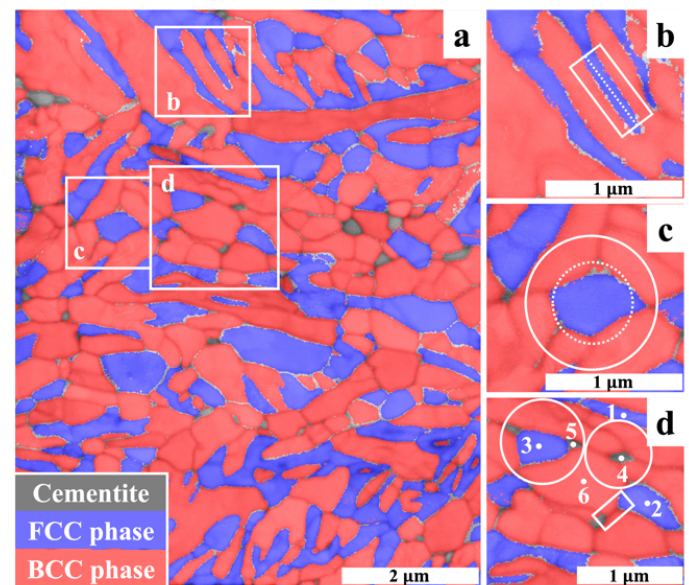


Fig. 3. a) TKD phase map with band contrast image highlighting three regions of interest (b, c, and d) used to define five different DICTRA systems introduced in Fig. 2. In (d), the numbers indicate the regions which their chemical composition determined by EDS (Table 1).

In Fig. 3(d), two distinct configurations of cementite particles depending on their position with the BCC and austenite phases are highlighted. In one arrangement, austenite and cementite are juxtaposed, possibly indicating the nucleation of austenite on cementite surfaces. Additionally, isolated cementite particles are observable within the BCC matrix, some reaching diameters of several hundred nanometers. Close observation of the particles adjacent to austenite grains reveals that most of



them have diameters less than 100 nm after the IAT, whereas those farther from austenite regions generally have diameters exceeding 200 nm.

The efficacy of cementite particles in promoting austenite nucleation has been shown highly dependent on their size, with maximum nucleation potency observed in particles with diameters between 100 and 200 nm (Zhang et al., 2021b). Particles larger or smaller than this range do not significantly contribute to the nucleation process. This suggests that, at the beginning of IAT, we might have had particles within the nucleation size range, but they partially dissolved during IAT. Conversely, the particles now larger than 200 nm and isolated in the BCC matrix likely did not engage in the austenite nucleation process.

Figure 4(a) illustrates the time-dependent variation in the fraction of transformed austenite under different scenarios by according to the DICTRA simulation. Curves b (red) and c (blue) depict results for lamellar and blocky austenite, respectively, which nucleated and grew at the BCC/BCC interface. Additionally, curves d (shown in solid and dashed black lines) indicates the growth kinetics of cementite particles that formed at the BCC/BCC interface but grew independently of austenite.

In curves b and c, the initial rapid increase in the fraction transformed is attributed to the growth of austenite under NPLE conditions. However, this NPLE stage is usually not observed experimentally under typical intercritical annealing conditions (Sun et al., 2023b). During this stage, growth is governed by carbon diffusion in ferrite. Over time, the carbon distribution in the ferrite matrix becomes progressively more homogeneous (Wei et al., 2013). Consequently, the growth of austenite becomes increasingly controlled by the carbon diffusion within the austenite until the diffusion of substitutional atoms, particularly Mn, begins to occur under the PLE conditions. The later stage comprises three separate steps (Luo et al., 2011). The corresponding interfacial boundary position at this stage can be considered as the size of the intercritical austenite observable experimentally.

Numerical simulations on the diffusion couples FCC/BCC (Fig. 4(a), curves b and c) indicate that the phase fraction of the formed austenite (for both the lamellar and blocky morphologies) after 1 hour of annealing at 680°C should increase up to 0.7. However, this fraction is generally much higher than what is calculated based on Fig. 3(a), showing about 0.38 area fraction.

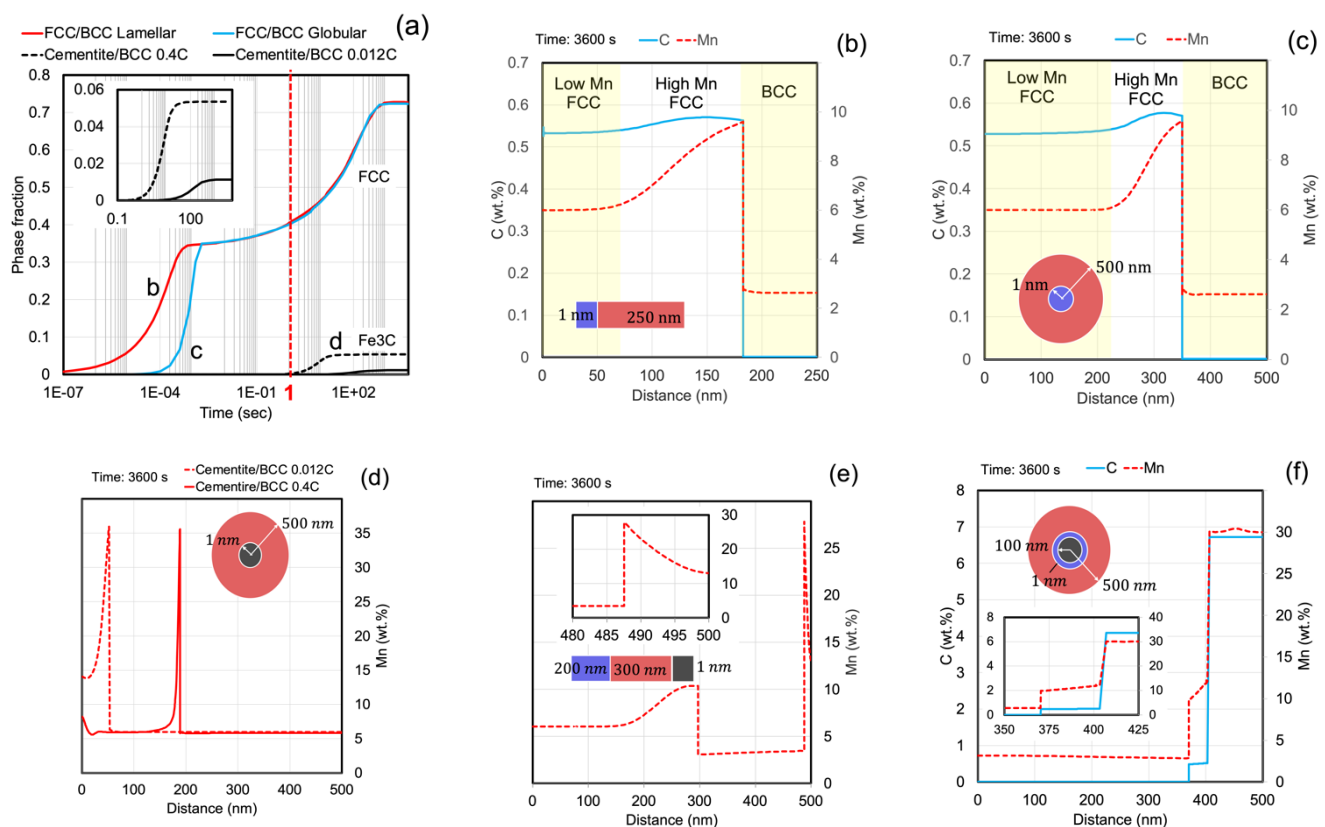


Fig. 4. a) Comparison of simulated volume fractions of austenite and cementite using DICTRA. Concentration profiles of C and Mn are shown for different growth scenarios: (b) lamellar austenite growth from the BCC/BCC interface, (c) blocky austenite growth from the BCC/BCC interface, (d) comparison of cementite growth and Mn distribution in the BCC phase with two different carbon contents 0.4 and 0.012 wt.%, (e) simultaneous growth of pre-existing austenite and cementite into the BCC matrix, and (f) growth of austenite at the Cementite/BCC interface based on chemical composition measured by EDS (Table 1).

There could be two explanations for this deviation: the interfacial dissipation energy due to finite interface mobility at the austenite/martensite interface (Huyan et al., 2018), and consumption of C and Mn due to formation of cementite after one second of annealing. The latter will be discussed in the following sections.

Comparing the growth rates of lamellar and blocky austenite during the NPLE stage, it is observed that lamellar austenite grows more quickly than blocky austenite does, even when simulations are conducted for both morphologies with the same system size (250 nm). This difference in the growth rate can be attributed to the larger surface area provided by the lamellar structure, which facilitates more efficient carbon diffusion and thereby accelerates the austenite growth. However, both morphologies exhibit the same PLE stage due to the sluggish diffusion of Mn, which cannot be compensated for even with a higher diffusion area. This slow Mn diffusion controls the overall transformation kinetics at this stage.

According to the simulation results, the concentration of Mn should be highly non-uniform in the austenite, particularly in the instance of the blocky morphology. Fig. 4(b) and (c) also indicate that the Mn flux into the austenite remains confined to the region near the boundary, with a maximum of 9.6 wt.%, even after a long holding time. However, the experimental results show a deviation in both Mn concentration and distribution profile. A comparison between chemical composition of different microconstituents, which are marked by numbers in Fig. 3(d), is shown in Table 1.

Table 1. Chemical composition of the regions defined in Fig. 3d

Point	Region of interest	Mn	Al	Si
1	Lamellar austenite	12.13	1.21	0.56
2	Blocky austenite	11.16	1.15	0.69
3	austenite formed on cementite	11.12	1.11	0.71
4	Cementite in BCC matrix	29.74	0.27	0.19
5	Cementite in the vicinity of FCC	29.24	0.19	0.19
6	BCC phase	3.91	1.79	0.98

Figure 4(d) shows two different scenarios of cementite growth for 1 hour annealing. As shown in this figure, two initial carbon levels for the BCC matrix were considered for the simulation. First, we supposed cementite forms in a region far from austenite, where all alloying elements only engage in its growth. It can be seen that there is a very marginal concentration of Mn at the cementite/BCC interface. Consequently, a thickness of 10 to 20 nm and a total cementite diameter of about 400 nm should be formed in the microstructure. However, in the experimental results, we could not find any cementite with such a large size and high Mn gradient.

On the other hand, Fig. 4(a) shows that from a BCC matrix with 0.4% C, cementite would start to grow after about 1 second of annealing, during which most of the carbon in the BCC matrix carbon is consumed by austenite formation. At this point, the BCC carbon concentration lowers to about 0.012 wt.%, but other elements still retain their nominal composition far from

the interface in both sides. For the second scenario, we used 0.012% C for the simulation of the growth of new cementite simulation growth, assuming that at first 40 percent of the BCC matrix transformed into austenite and then cementite separately forms in the remaining matrix. In this scenario, the austenite, cementite, and BCC phases seem to have their chemical compositions closer to the experimental results (Table 1), and also their sizes correspond to those depicted in Fig. 3(a) and (d).

At this point (after 1 second of annealing), another scenario could be considered: a competition between the formation of cementite and the simultaneous growth of pre-existing austenite. Fig. 4(e) illustrates the results for this co-existing scenario and the corresponding system size. In this scenario, cementite and austenite are both forming and growing at the same time. The simulation reveals that after 1 hour of holding at 680°C, cementite can only grow 10 nanometers.

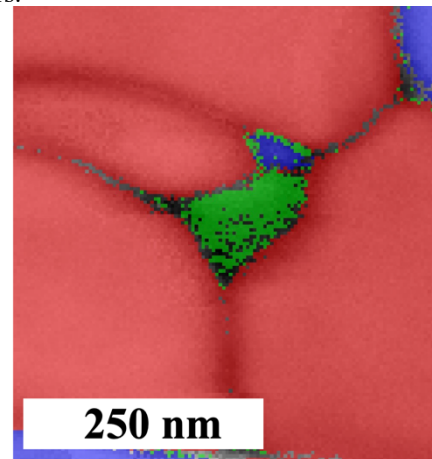


Fig. 5. EBSD-TKD phase map with band contrast image showing austenite, less than 50 nm in size, formed on a cementite particle with a 200 nm diameter after 1 hour of soaking at 680°C (austenite: blue, cementite: green and BCC in red color).

This limited growth could be due to the simultaneous consumption of carbon and other alloying elements by both cementite and austenite. As a result, the growth rate of austenite is also reduced compared to a system where only FCC/BCC transformations occur without the presence of cementite. Moreover, the simulation indicates that the Mn content in cementite, under these conditions, would be less than 20 wt.% in average. This is significantly lower than the Mn content measured by EDS, suggesting that the actual distribution of Mn is more complex than the simulation predicts. This discrepancy further implies that the interaction between cementite and austenite formation is more intricate, affecting the distribution and concentration of alloying elements. In summary, the co-existence scenario shown in Fig. 4e suggests that while cementite formation limits its own growth and the growth of austenite, it does not fully account for the experimentally observed austenite sizes and Mn concentrations. The competition between cementite and austenite growth, along with the resultant elemental distributions, highlights the complexity of phase transformations in these materials.

Additionally, Fig. 4f illustrates a scenario based on experimental results where the cementite particle size and the chemical composition of BCC and cementite were measured directly using TKD and EDS methods. In this scenario, we assumed that austenite would form at the Cementite/BCC interface, as observed in Fig. 5. However, the simulation results shown in Fig. 4f suggest that this scenario does not account for the formation of large blocky austenite grains. The simulation indicates that within holding for 1 hour, the austenite can only grow to a size of 30 nm. In contrast, the measured austenite diameter in Fig. 3d is approximately 700 nm, demonstrating a significant discrepancy. This discrepancy suggests that the big austenite grain does not primarily form at the Cementite/BCC interface. Instead, austenite likely forms initially within the FCC/BCC system separately. Over time, these austenite grains may coalesce with austenite that nucleates at the Cementite/BCC interface, leading to the larger blocky austenite grains observed experimentally.

## 5. CONCLUSIONS

In this study, various scenarios were simulated to better understand the microstructure observed and analyzed using EBSD-TKD and EDS methods after annealing of a hot-rolled 0.40C-6 Mn-1Si-2Al-0.05 Nb steel at 680°C for 1 hour. The aim was to approximate the system size and chemical composition of each microconstituent, including lamellar austenite with a thickness of 100 to 200 nm, blocky austenite with a radius of approximately 500 nm, and cementite particles, either adjacent to austenite or dispersed in the BCC matrix.

The DICTRA software was used to simulate five potential systems to determine which of them could be responsible for the existing microstructure. The following systems were simulated:

- FCC/BCC system
- Cementite/BCC with 0.4% C in the matrix and Cementite/BCC with 0.012% C in the matrix (both assuming cementite growth without the influence of austenite)
- Two co-existing austenite and cementite systems (adjacent and separate from each other, considering the competition between austenite and cementite).

Based on comparison of the results from the simulations and experimental observations, it seems highly likely that austenite nucleates first at the BCC/BCC interfaces and grows, leading to the transformation of 40% of the BCC phase to austenite within 1 s at 680°C. Then, the cementite formation occurs, which ceases the further growth of austenite as a competitor. Following this, only cementite continues to grow under these circumstances. Notably, after 1 hour of annealing at 680°C, the experimental results show that 38% austenite exists in the microstructure, and the size of cementite particles is about 100–200 nm, which align well with the above hypothesis.

## ACKNOWLEDGMENTS

The authors would like to thank Jane ja Aatos Erkon säätiö (JAES) and Tiina ja Antti Herlinin säätiö (TAHS) for their

financial supports on Advanced Steels for Green Planet project.

## REFERENCES

- Bansal, G.K., Madhukar, D.A., Chandan, A.K., Ashok, K., Mandal, G.K., Srivastava, V.C., (2018). On the intercritical annealing parameters and ensuing mechanical properties of low-carbon medium-Mn steel. *Materials Science and Engineering: A* 733, 246–256. doi: 10.1016/j.msea.2018.07.055
- Bhattacharya, A., Biswal, S., Barik, R.K., Mahato, B., Ghosh, M., Mitra, R., Chakrabarti, D., (2024). Comparative interplay of C and Mn on austenite stabilization and low temperature impact toughness of low C medium Mn steels. *Mater Charact* 208. doi: 10.1016/j.matchar.2024.113658
- Dai, Z., Ding, R., Yang, Z., Zhang, C., Chen, H., (2018). Elucidating the effect of Mn partitioning on interface migration and carbon partitioning during Quenching and Partitioning of the Fe-C-Mn-Si steels: Modeling and experiments. *Acta Mater* 144, 666–678. doi: 10.1016/j.actamat.2017.11.025
- Enomoto, M., Hayashi, K., (2020). Simulation of Austenite Formation During Continuous Heating from Low Carbon Martensite with Poly-dispersed Cementite. *Metall Mater Trans A Phys Metall Mater Sci* 51, 618–630. doi: 10.1007/s11661-019-05569-3
- Enomoto, M., Hayashi, K., (2018). Modeling the growth of austenite in association with cementite during continuous heating in low-carbon martensite. *J Mater Sci* 53, 6911–6921. doi: 10.1007/s10853-018-2020-2
- Guo, Q., Hu, B., Luo, H., (2022). Mechanism and Application of Reverse Austenitic Transformation in Medium Mn Steels: A Systematic Review. *Steel Res Int* 93, 2200257. doi: 10.1002/srin.202200257
- Han, J., (2023). A Critical Review on Medium-Mn Steels: Mechanical Properties Governed by Microstructural Morphology. *Steel Res Int* 94, 2200238. doi: 10.1002/srin.202200238
- Hu, B., Sui, H., Wen, Q., Wang, Z., Gramlich, A., and Luo, H. (2024). Review on the plastic instability of medium-Mn steels for identifying the formation mechanisms of Lüders and Portevin–Le Chatelier bands. *International Journal of Minerals, Metallurgy and Materials*. doi: 10.1007/s12613-023-2751-1
- Huyan, F., Yan, J.Y., Höglund, L., Ågren, J., and Borgenstam, A., (2018). Simulation of the Growth of Austenite from As-Quenched Martensite in Medium Mn Steels. *Metall Mater Trans A Phys Metall Mater Sci* 49, 1053–1060. doi: 10.1007/s11661-018-4497-3
- Lai, Q., Gouné, M., Perlade, A., Pardoën, T., Jacques, P., Bouaziz, O., and Bréchet, Y., (2016). Mechanism of Austenite Formation from Spheroidized Microstructure in an Intermediate Fe-0.1C-3.5Mn Steel. *Metall Mater Trans A Phys Metall Mater Sci* 47, 3375–3386. doi: 10.1007/s11661-016-3547-y

- Luo, H., Shi, J., Wang, C., Cao, W., Sun, X., and Dong, H. (2011). Experimental and numerical analysis on formation of stable austenite during the intercritical annealing of 5Mn steel. *Acta Mater* 59, 4002–4014. doi: 10.1016/j.actamat.2011.03.025
- Mehrabi, A., Zurob, H.S., and McDermid, J.R., (2024). Process Maps for Predicting Austenite Fraction (vol.%) in Medium-Mn Third-Generation Advanced High-Strength Steels. *Materials* 17. doi: 10.3390/ma17050993
- Miller, R.L. (1972). Ultrafine-grained microstructures and mechanical properties of alloy steels. *Metallurgical Transactions* 3, 905–912. doi: 10.1007/BF02647665
- Sadeghpour, S., Somani, M.C., Kömi, J., and Karjalainen, L.P. (2021). A new combinatorial processing route to achieve an ultrafine-grained, multiphase microstructure in a medium Mn steel. *Journal of Materials Research and Technology* 15, 3426–3446. doi: 10.1016/j.jmrt.2021.09.152
- Sun, B., Kwiatkowski da Silva, A., Wu, Y., Ma, Y., Chen, H., Scott, C., Ponge, D., and Raabe, D. (2023). Physical metallurgy of medium-Mn advanced high-strength steels. *International Materials Reviews* 68, 786–824. doi: 10.1080/09506608.2022.2153220
- Wei, R., Enomoto, M., Hadian, R., Zurob, H.S., and Purdy, G.R. (2013). Growth of austenite from as-quenched martensite during intercritical annealing in an Fe-0.1C-3Mn-1.5Si alloy. *Acta Mater* 61, 697–707. doi: 10.1016/j.actamat.2012.10.019
- Wu, Y.X., Wang, L.Y., Sun, W.W., Styles, M.J., Studer, A.J., Bréchet, Y., Arlazarov, A., and Hutchinson, C.R. (2020). Austenite formation kinetics from multicomponent cementite-ferrite aggregates. *Acta Mater* 196, 470–487. doi:10.1016/j.actamat.2020.07.001
- Ye, Q., Dong, H., Guo, Q., Yu, Y., Qiao, L., and Yan, Y. (2024). Tailoring the austenite characteristics via dual nanoparticles to synergistically optimize the strength-ductility in cold rolled medium Mn steel. *J Mater Sci Technol* 169, 158–171. doi: 10.1016/j.jmst.2023.07.002
- Zhang, X., Miyamoto, G., Kaneshita, T., Yoshida, Y., Toji, Y., and Furuhashi, T., (2018). Growth mode of austenite during reversion from martensite in Fe-2Mn-1.5Si-0.3C alloy: A transition in kinetics and morphology. *Acta Mater* 154, 1–13. doi: 10.1016/j.actamat.2018.05.035
- Zhang, X., Miyamoto, G., Toji, Y., Nambu, S., Koseki, T., and Furuhashi, T., (2018). Orientation of austenite reverted from martensite in Fe-2Mn-1.5Si-0.3C alloy. *Acta Mater* 144, 601–612. doi: 10.1016/j.actamat.2017.11.003
- Zhang, X., Miyamoto, G., Toji, Y., Zhang, Y., and Furuhashi, T., (2021). Role of cementite and retained austenite on austenite reversion from martensite and bainite in Fe-2Mn-1.5Si-0.3C alloy. *Acta Mater* 209. doi: 10.1016/j.actamat.2021.116772
- Zhang, Y., Ye, Q., and Yan, Y., (2024). Processing, microstructure, mechanical properties, and hydrogen embrittlement of medium-Mn steels: A review. *J Mater Sci Technol*. doi: 10.1016/j.jmst.2024.03.014

# Effect of Slag Particle Diameter on the Re-melting of Ferrochrome Slag by means of Steelmaking Liquid Slag

Reza Safavi Nick, Elisa Olson, Matti Aula, Severi Anttila and Esa Puukko

*Outokumpu Research and Development, Tornio, Finland (e-mail: [reza.safavinick@outokumpu.com](mailto:reza.safavinick@outokumpu.com)).*

---

**Abstract:** Stainless steelmaking slags are, currently, one of the most common non-utilized slags in steelmaking. Hence, in an integrated stainless steelmaking process with a ferrochrome submerged arc furnace, this means not only losing iron to the slag but also valuable chrome. Hence, recovery of iron and chrome have a business incentive and an important function for green industry initiative by reducing the requirement of virgin material. However, one of the challenges of slag recycling can be the energy-intensive nature of such a practice. Therefore, an energy efficient approach in material recovery could enhance the incentive of recycling of slag instead of the current practice of land field storage; one such approach is mixing the solid ferrochrome slag into liquid slag from the steelmaking production line. To that end, a static model of a suspended slag particle inside a melt has been developed to investigate the effect of particle size on evolution of temperature within the solid particles. The simulation showed that changes in the diameter of particle can have a significant effect on energy diffusion from the melt into the slag particle. As an example, the simulation suggests that the temperature magnitude at the centre of a 2mm-in-diameter particle reaches 1200 °C after 1s simulation time while, with 5mm particles the temperature magnitude is less than 200 °C. This behaviour is amplified further when the diameter of particle increases further showing a delaying behaviour of particle's diameter on energy diffusion and, consequently, remelting of solid particles.

*Keywords:* Slag, Ferrochrome, remelting, modelling, solid particle.

---

## 1. INTRODUCTION

As the European Green Deal has become a baseline to make the European Union (EU) climate neutral in 2050, the role of industries and more specifically iron- and steel-producers have become increasingly critical (European Commission. Directorate General for Research and Innovation. 2021). This is due to the fact that this industry is not only energy intensive but also produces a significant amount of by-product mainly in the form of slag which is larger than 1.1 tons of slag for each ton of ferrochrome (Kauppi 2007).

The ferrochrome slag contains not only SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, MgO or CaO but also some sizable percentage of chrome and iron (Karhu et al. 2020). However, a portion of the slag produced by steelmakers will be stored in slag yard ( open space storage facility for unused by-product ) while some is used for road construction or by concrete producers (Gao et al. 2023).

Considering the amount of slag produced for every ton of ferrochrome and the percentage of chrome and iron available in the slag ( 8 and 4, respectively ) (Karhu et al. 2020) it is not difficult to calculate the loss of valuable material to the slag. Obviously, improving the share of unused slag material would result in reduction of the virgin material requirement. This reduction in the raw material, then, would have a direct impact

on reducing the carbon footprint by reducing the need for raw material mining.

As an example, Karhu et. al. (Karhu et al. 2020) have shown the positive impact of the ferrochrome slag when used as an aggregate to produce refractories. It is shown that the refractory bricks produced using such a slag not only have an improved strength but also show better thermal insulation.

With these considerations, the current study tries a new approach where the solid ferrochrome slag aggregate will be mixed and remelted using AOD-slag. This mixed melt, later on, can be used as a feed material to recover the chrome and iron deposit of solid slag material. To that end, a CFD model of single solid slag particle within a liquid slag is modelled to investigate the effect of particle density on heat diffusion within the particle and, consequently, remelting process of solid ferrochrome slag.

## 2. MODEL SPECIFICATION

### 2.1 Modelling Approach

As mentioned, the project tries to melt the solid ferrochrome slag by pouring liquid AOD slag over a packed solid slag bed in a slag pot. However, the initial research question could be formulated as how the heat is transferred through a single pellet inside a melt and if the liquid temperature in the



immediate vicinity of the pellet could drop to the solidus temperature.

With this objective in mind, the current model investigates the temperature evolution of single solid particle suspended in a liquid bath. Moreover, since the liquid assumed to be at rest, it is possible to further simplify the system by considering the liquid slag as a solid phase. This, of course, disregards the natural convection may occur around the pellet which in comparison to the liquid bath size could be of a minimal effect.

The thermal properties used for the solid ferrochrome slag particle are derived from the work of (Karhu et al. 2020) as reported below:

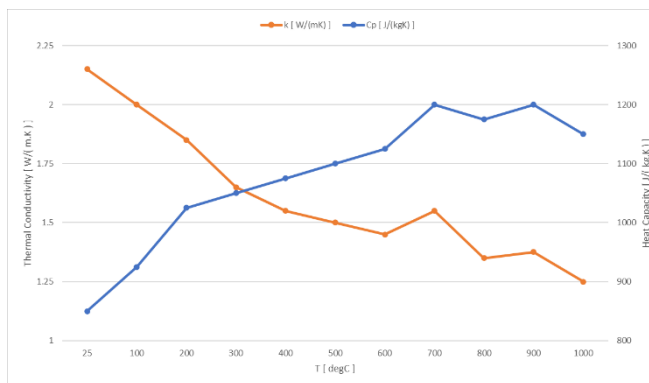


Fig. 1. Thermal properties of ferro chrome solid slag: Castable 1 modified from (Karhu et al. 2020).

## 2.2 Geometry and Boundary Conditions

As explained, the model looked into practice of suspended solid slag particle inside a liquid melt which is shown in Fig. 2. The domain is formed of a cubic box of  $0.5 \times 0.5 \times 0.5 \text{ m}^3$  which is open to the atmosphere from the top while the sides and bottom faces were assumed to be adiabatic walls. This is to assume a solid particle inside a crucible in a controlled environment.

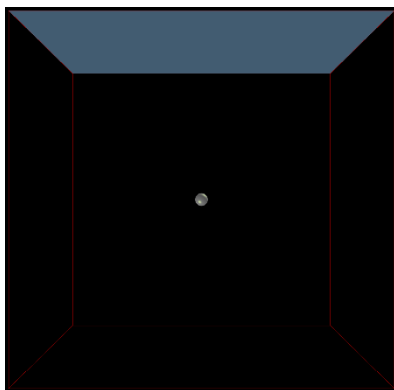


Fig. 2. Domain of calculation with suspended solid slag particle.

As mentioned, the objective of this work is to investigate the changes in the temperature profile of the solid slag particle where it is suspended in a liquid slag. Hence, to be able to paint a clear picture, it seems logical to not only look into changes in the magnitude of temperature within the solid particle but also evolution of temperature magnitude of the liquid bath surrounding the particle. This is due to the fact that if the liquid in contact with the solid particle freezes, this phenomenon could hamper the transfer of energy into the solid material.

Naturally, the solid slag particles are produced as a result of crushing solid material hence, the particles do not have a uniform diameter and it can significantly vary from a small round particle to a solid lumpy material. Therefore, it is important to investigate the changes in the solid particle temperature profile due to changes in the particle size.

### 3.1 Temperature Profile at and around the Solid Particle

Figure 3 shows the evolution of temperature in time at different distance from the centre of the pellet. As can be seen, the x-axis range varies from zero to 5 mm while the particle diameter is equal to 2 mm. Hence, the figure shows the changes in the temperature magnitude of the liquid slag in contact with the pellet till 4 mm away from the surface. Moreover, the vertical dashed line in the figure shows the interface of solid particle to the liquid slag.

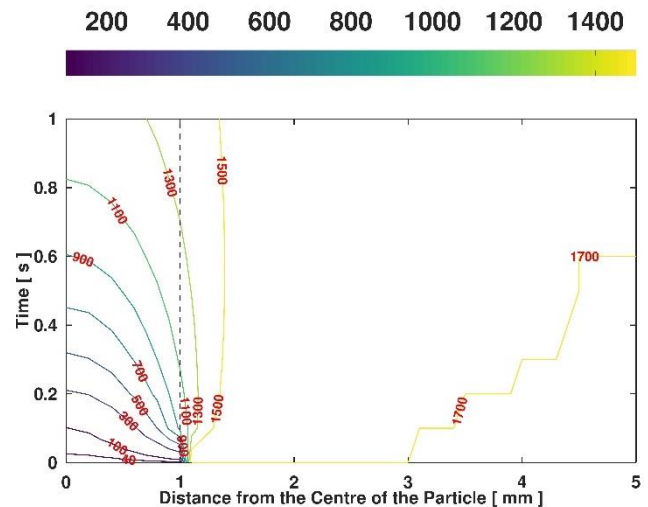


Fig. 3. Contour plot of temperature evolution from the centre of pellet till 4 mm in the liquid bath.

For the pellet, as the figure shows, the temperature at the centre reaches  $1200 \text{ }^\circ\text{C}$  where the liquid temperature in the vicinity of the pellet is around  $1400 \text{ }^\circ\text{C}$ . These values are around  $800$  and  $1200 \text{ }^\circ\text{C}$  respectively,  $0.5 \text{ s}$  into the simulation. This suggest that the energy diffuses faster within the particle in the first  $0.5 \text{ s}$  of the simulation in comparison to the next.

Furthermore, Figure 5 shows the evolution of temperature within the pellet and inside the liquid phase with respect to time for  $1 \text{ s}$ , similar to Fig. 3. It can be seen that the temperature magnitude of pellet centre, when the diameter

increases to 5 mm, drops to 200 °C, more than 80% drop in the magnitude.

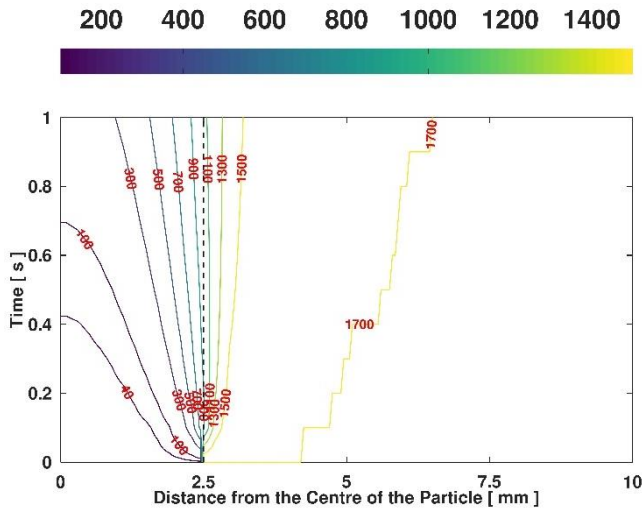


Fig. 4. Contour plot of temperature evolution from the centre of pellet till 7.5 mm in the liquid bath.

Considering the significant change in the magnitude of temperature, it seemed reasonable to try to run the simulation for longer time interval to investigate the effect of particle size on the evolution of temperature.

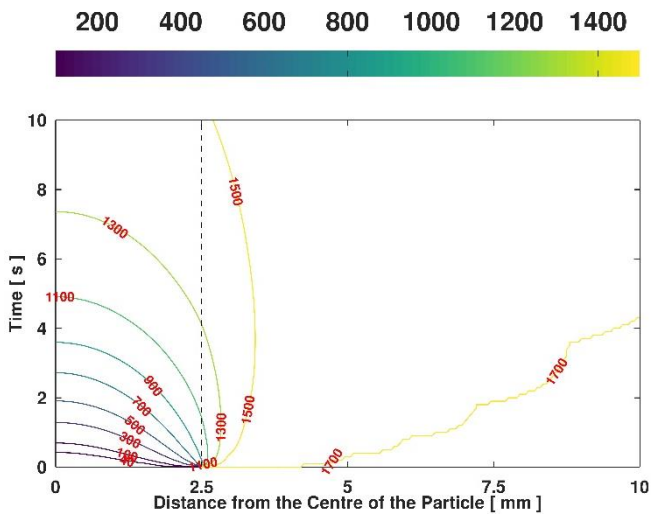


Fig. 5. Contour plot of temperature evolution from the centre of pellet till 7.5 mm in the liquid bath for 10 s.

Figure 5 shows the changes in the temperature magnitude within the pellet and inside the liquid phase for the extended time interval of 10 s. As can be seen, by increasing the diameter of particle to 5 mm, it will take 6 s till the centre of the pellet registers the 1200 °C mark.

Considering such a significant change in the temperature evolution in and around the solid slag pellet, the next logical step was to investigate the temperature profile if the particle size increases.

Figure 6 shows the evolution of the temperature within the pellet and inside the liquid phase with respect to time. It can

be seen that the temperature magnitude of pellet centre, when the diameter increases to 10 mm and after 10 s simulation time, is below 700 °C.

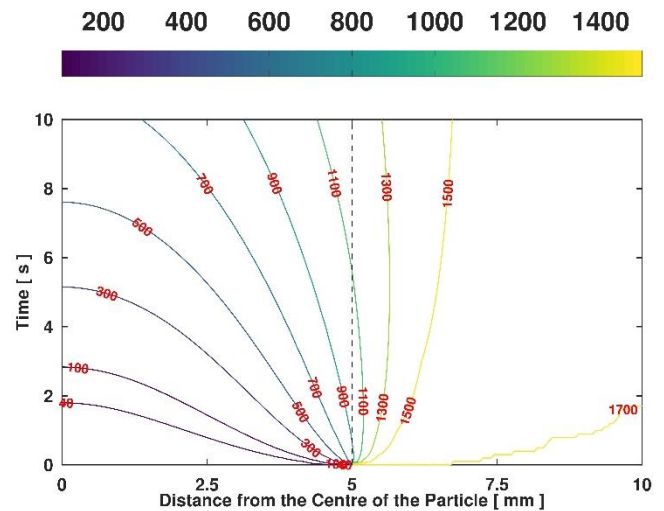


Fig. 6. Contour plot of temperature evolution from the centre of pellet till 5 mm in the liquid bath for 10 s.

The figure also shows that the liquid temperature at the beginning of the simulation is around 1000 °C and after 10 s simulation time it reaches 1200 °C.

Figures 7 and 8 show the evolution of the temperature within the pellet and inside the liquid phase with respect to time for solid particles of 20 and 40 mm, respectively.

It can be seen that the temperature magnitude of pellet centre, when the diameter increases to 20 mm and after 10 s simulation time, has not even reached 100 °C. Moreover, it can be seen that the liquid phase temperature in contact with the particle is almost 1000 °C during the entire simulation time.

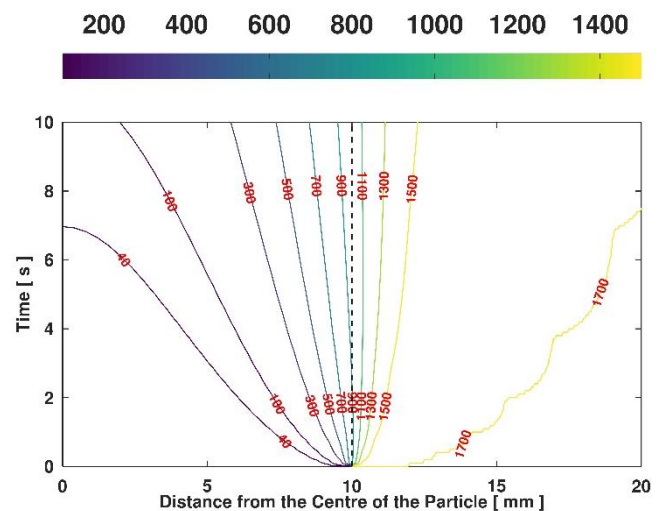


Fig. 7. Contour plot of temperature evolution from the centre of pellet till 10 mm in the liquid bath.

Similarly, it can be seen that, even after 10 s and more than 50% away from the centre of pellet, the magnitude of temperature has not even reached 40 °C and at the near surface

region, the magnitude of temperature is hardly 900 °C. The figure also shows that the liquid phase temperature in the vicinity of solid particle registers a constant magnitude of 900 °C.

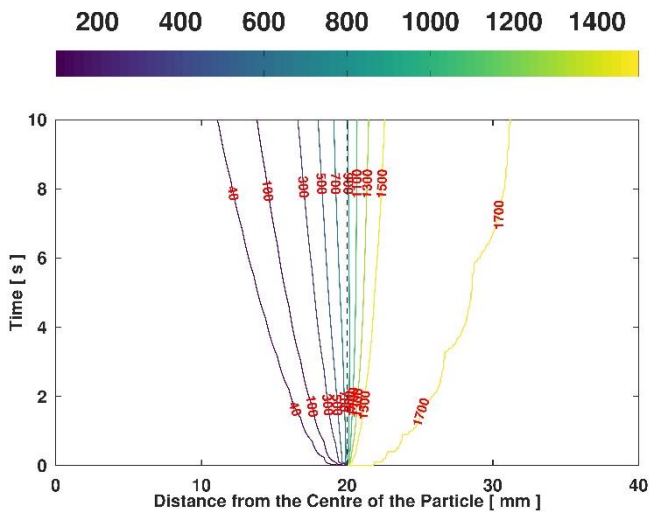


Fig. 8. Contour plot of temperature evolution from the centre of pellet till 20 mm in the liquid bath

### 3.2 Temperature Magnitude

Table 1 shows the changes in the minimum temperature of the liquid slag at the vicinity of the slag particle. This, in another word, the lowest magnitude of temperature the liquid slag in contact with the slag particle registers.

Table 1 Minimum Liquid Film Temperature around the Particle

Diameter	time	$T_t^{Film}$	$T_{Final}^{Film}$
2	0.13	1013	1377
5	0.20	930	1486
10	0.25	881	1197
20	0.78	884	1006
40	3.07	885	914

The importance of such a parameter is that if the liquid slag solidifies around the slag particle, then such a solid shell could hamper the transfer of energy from the melt to the cold particle. This is, of course, due to the fact that the solid slag has a very poor thermal properties and could act as an insulation layer around the solid particle.

In the table, the first column shows the particle diameter, while the second column shows when the liquid file in contact with the solid particle register the lowest magnitude. Then, the third column shows such a magnitude i.e., the lowest magnitude the liquid file registers, and the final column shows the magnitude of the same variable at the end of the simulation.

As can be seen, when the particle diameter increases, the lowest temperature magnitude drops while it occurs later in time. However, when the particle diameter increases from 10 to 20 to 40 mm, the simulation suggests that the minimum magnitude would not drop but occurs significantly further in time ( 3.07 s compare to 0.78 s to 0.25 s )

The final column, on the other hand, shows the lowest magnitude of temperature at the end of simulation is significantly lower for particle of 40 mm diameter compare to the particle of 20 mm.

## 4. DISCUSSION

As it was seen, the temperature magnitude of solid particle with 2 mm diameter at the centre of pellet reached 1200 °C after only one second with average temperature magnitude larger than 1300 °C with liquid phase temperature magnitude around 1400 °C. Therefore, it can be concluded that the solid ferrochrome slag will be melted in such a condition. However, Figure 4 seems to suggest a drastic drop in the temperature magnitude when the diameter of particle is set to 5 mm.

Figure 9 shows the temperature within the slag particle for two diameters of 2 and 5 mm. It can be seen that the minimum temperature magnitudes within the solid slag are 1200 and less than 200 °C for particles with diameters 2 ad 5 mm, respectively.

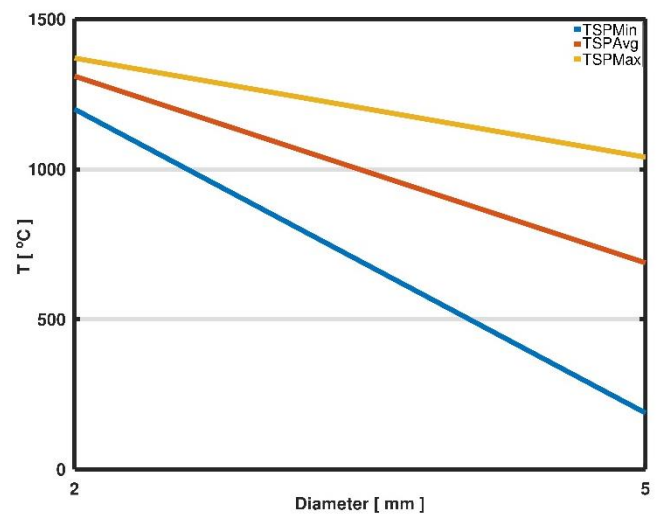


Fig. 9. Minimum, average and maximum temperature magnitude of solid slag pellet at 1 s for diameters 2 and 5 mm.

As mentioned before, the drastic change in the magnitude of temperature suggested the simulation should be run for longer time interval where the particle diameter is larger than or equal 5 mm. Figure 10 shows the final magnitude of temperature for particles diameters 5, 10, 20 and 40 mm within the solid slag and the liquid phase after 10 s simulation time.

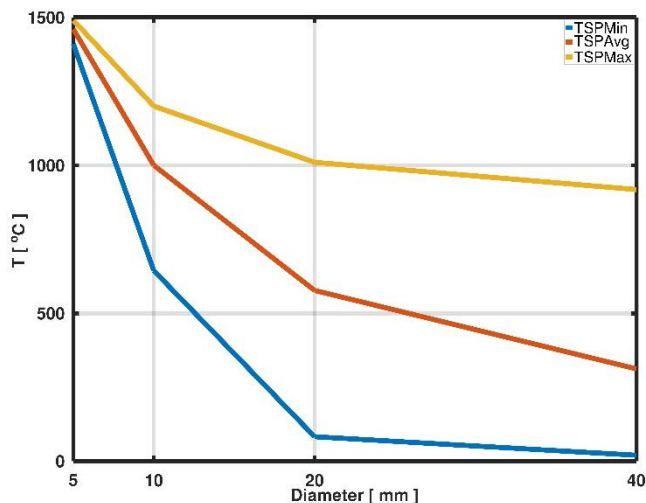


Fig. 10 Minimum, average and maximum temperature magnitude of solid pellet at 10 s for diameters 5, 10, 20 and 40 mm.

It can be seen that the minimum temperature magnitude within the particle drop from 1400 °C to 20 °C for particles diameters of 5 to 40 mm (blue curve in Fig. 10). This, of course, suggests that the temperature magnitude at the centre of particles is not affected by the liquid phase energy even after 10 s simulation time when the particle diameter is equal 40 mm. The figure also shows that the gap between the average temperature magnitude to the maximum increases with the increment of solid particle diameter. This behaviour suggests that a larger portion of the solid particle at larger diameter is yet to be affected by the liquid phase energy.

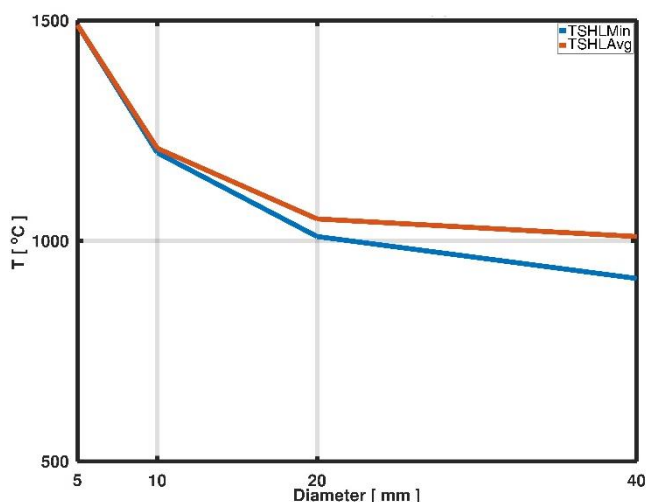


Figure 11 Minimum and average temperature magnitude of liquid phase at 10 s for diameters 5, 10, 20 and 40 mm.

Moreover, Figure 10 shows that for the two cases of 20 and 40 mm diameters, the temperature magnitude of liquid phase in the vicinity of the solid particle is around 1000 [°C]. Therefore, the probability of liquid slag to solidify at the surface of the particle to form a shell increase which, in turn, has the potential to hinder the energy diffusion from the liquid phase into the solid particle.

The cause of such behaviour could be explained through thermal properties of the solid slag. As mentioned, the solid ferrochrome slag properties were reported by Karhu et.al. (Karhu et al. 2020) and Figure 1 shows the thermal conductivity and heat capacity of such a material.

As can be seen, the figure suggests an increase in the temperature magnitude of such a material will lead to decrement of thermal conductivity and increment of specific heat, simultaneously. This, of course, imply that the solid ferrochrome slag hampers the diffusion of energy at the higher magnitude of temperature while it can absorb and store larger amount of energy. These two actions together could be a likely explanation as to why an increase in the size of the particle, even only by a factor of 2.5 from 2 mm to 5, causes such a slowdown in the transfer of energy from the environment towards the centre of solid particle.

## 5. CONCLUSIONS

The current study investigates the exchange of energy between the solid ferrochrome slag and the liquid AOD slag. This objective is to study the re-melting of the solid ferrochrome slag using the liquid slag collected at the AOD station. To that end, the thermal properties of the materials in use play a significant role.

As shows, the simulation suggests that after 1 s simulation time, the temperature magnitude at the centre of solid slag particle reaches 1200 °C. However, an increase in the diameter to 5 mm ( by 2.5 factor ) causes the temperature magnitude at the centre drops significantly to the value of less than 200 °C for a 1 s simulation.

Same behaviour can be observed with respect to increases in the diameter of solid slag particle where for a particle of 40 mm diameter, the magnitude of temperature at the centre of the particle is barely higher than initial temperature of 20 °C.

At the same time, it can be observed that the temperature magnitude of the liquid phase in the vicinity of solid particle drop to 1000 °C for a particle of diameter 40 mm. This, of course, suggest that there is a higher probability for the liquid AOD slag to solidify around the solid particle to form a shell which then hampers the transfer of energy from the environment into the solid ferrochrome slag to a larger degree.

## REFERENCES

- European Commission. Directorate General for Research and Innovation. (2021). *European Green Deal: Research & Innovation Call*. LU: Publications Office. <https://data.europa.eu/doi/10.2777/33415> (June 3, 2024).
- Gao, Wenhao, Wentao Zhou, Xianjun Lyu, Xiao Liu, Huili Su, Chuanming Li, and Hui Wang. (2023). Comprehensive Utilization of Steel Slag: A Review. *Powder Technology* 422: 118449. doi:10.1016/j.powtec.2023.118449.
- Karhu, Marjaana, Bob Talling, Patrycja Piotrowska, Alba Matas Adams, Abirami Sengottuvelan, Elina

Huttunen-Saarivirta, Aldo R. Boccaccini, and Pertti Lintunen. (2020). Ferrochrome Slag Feasibility as a Raw Material in Refractories: Evaluation of Thermo-Physical and High Temperature Mechanical Properties. *Waste and Biomass Valorization* 11(12): 7147–57. doi:10.1007/s12649-020-01092-4.

Kauppi, M. (2007). Production, Characteristics and use of Ferrochromium Slags.  
<https://api.semanticscholar.org/CorpusID:48465274>.



# Optimizing Energy Consumption in Hydrogen Reduction of Iron Ore Pellet: Insights from HSC Chemistry Analysis

A. Heidari, T. Fabritius

Process Metallurgy Research Unit, University of Oulu, Oulu, Finland (e-mail: [aidin.heidari@oulu.fi](mailto:aidin.heidari@oulu.fi), [timo.fabritius@oulu.fi](mailto:timo.fabritius@oulu.fi)).

**Abstract:** Iron ore pellet reduction in shaft furnaces represents a critical process in the steelmaking industry, with energy consumption being a key factor influencing both economic viability and environmental sustainability. This study employs HSC Chemistry software to model and simulate the energy consumption of hydrogen reduction of iron ore pellets under varying water vapor content within the shaft furnace. Thermodynamic modeling was carried out as the first step to analyze the effect of water vapor on the thermodynamic equilibrium, determining the possible range of water vapor content. Subsequently, energy consumption of the process was modeled based on heat and mass balance. Through comprehensive analysis, we investigate the impact of water vapor on the overall energy efficiency of the process based on the two scenarios of supplying the required heat by preheating the feed materials or injection of oxygen to the furnace. Our findings reveal significant insights into optimizing energy consumption and operational parameters to enhance the sustainability and cost-effectiveness of iron ore pellet reduction. This research contributes to the ongoing efforts towards achieving greater efficiency and reduced environmental footprint in the steelmaking industry.

**Keywords:** Hydrogen Reduction of Iron ore pellets, Shaft furnace, Energy consumption, HSC Chemistry analysis

## 1. INTRODUCTION

Currently, more than 70% of the world's iron production is derived from the integrated blast furnace (BF) and basic oxygen furnace (BOF) process, which emits around 1.9 tons of CO<sub>2</sub> per ton of crude steel, as coke is the primary reducing agent. Consequently, the iron and steel industries account for about 7% of global CO<sub>2</sub> emissions. Using hydrogen as a reducing agent presents a solution to reduce CO<sub>2</sub> emissions, replacing CO<sub>2</sub> with H<sub>2</sub>O as a by-product (Özgün et al., 2023; Souza Filho et al., 2023; Spreitzer and Schenk, 2019).

Hydrogen can be employed in ironmaking through three methods: hydrogen injection in blast furnaces, hydrogen direct reduction, and hydrogen plasma reduction (Ahmed et al., 2020; Raabe et al., 2023; Souza Filho et al., 2022, 2021). Among these, hydrogen direct reduction is closest to industrialization, with several companies initiating the construction of their first hydrogen direct reduction plants (Sun et al., 2023). The schematic of process is shown in Fig. 1.

The thermodynamics and kinetics of hydrogen reduction of iron ores have been extensively studied (Fradet et al., 2023; Heidari et al., 2021). Unlike carbon monoxide reduction, hydrogen reduction is entirely endothermic, making increased temperatures thermodynamically favorable but requiring additional energy input. Hydrogen reduction is faster than carbon monoxide reduction due to hydrogen's smaller molecule size, and higher mobility and diffusivity. However, many factors, including feed characteristics (such as ore type, mineralogy, and porosity) and system parameters (such as

reduction temperature, gas flow rate, and pressure), influence the process kinetics (Spreitzer and Schenk, 2019).

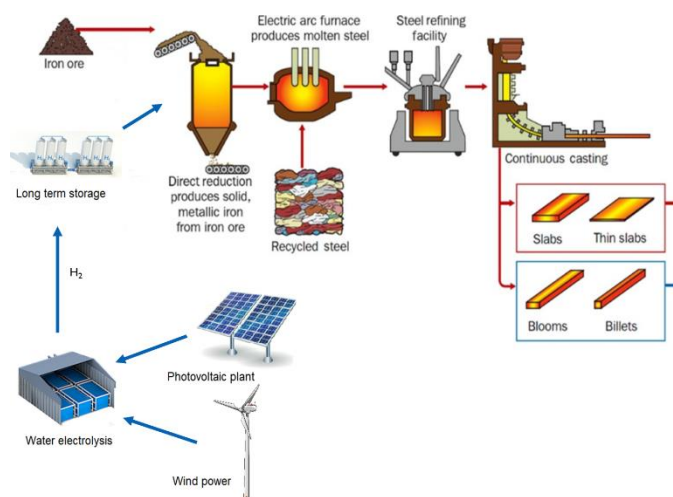


Fig. 1. The schematic of hydrogen-based fossil-free iron and steel making.

Water vapor in the shaft furnace also affects the thermodynamics, kinetics, and energy consumption of the process (El-Zoka et al., 2023). Water vapor can originate from various sources, including as a by-product of reduction reactions, moisture in the pellets from the pelletizing process, and oxygen injection for energy supply, which reacts with hydrogen to produce water vapor. Moreover, hydrogen can be produced through methods like steam methane reforming,

electrolysis of water, and natural gas pyrolysis, all of which may introduce water vapor as an impurity.

Given the presence of water vapor in the shaft furnace, it is essential to study its effects on the process from multiple perspectives. This study examines the energy consumption of the hydrogen reduction process in a shaft furnace with varying water vapor content using HSC Chemistry software.

## 2. HEAT AND MASS BALANCE MODEL

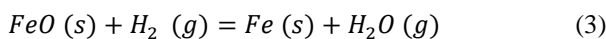
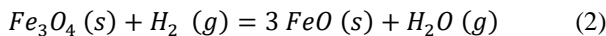
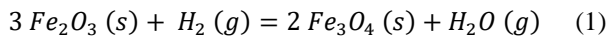
Hydrogen is currently utilized as a reducing agent in certain direct reduction processes like MIDREX and HYL, where hydrogen and carbon monoxide are produced by natural gas cracking in the reformer before being introduced to the shaft furnace. However, no shaft furnace operates with 100% hydrogen for real data application in our model. Thus, this study uses MIDREX process data, the leading direct reduction technology, and adapts it for the hydrogen reduction process.

In this static model, system components are divided into gas and solid phases. The gas phase includes reducing gas and top gas, while the solid phase consists of feed material and produced sponge iron (DRI). The chemical compositions of these phases are detailed in Table 1.

Table 1. Phases and chemical components considered in shaft furnace.

Phases	Chemical composition
Gas	H <sub>2</sub> , H <sub>2</sub> O, O <sub>2</sub>
Solid	Fe <sub>2</sub> O <sub>3</sub> , Fe <sub>3</sub> O <sub>4</sub> , Fe, SiO <sub>2</sub> , CaO, MgO, Al <sub>2</sub> O <sub>3</sub> , TiO <sub>2</sub> , MnO

Hydrogen reduction process includes three stages of reduction of hematite to magnetite, magnetite to wüstite, and wüstite to metallic iron as it is presented by the reactions 1, 2, and 3 respectively.



Moreover, oxygen can be introduced to the shaft furnace and the combustion of hydrogen takes place through the reaction 4 to supply the required heat of reduction reactions.



### 2.1 Mass balance

The mass balance model that has been used in this study is based on the elemental distribution. It means that the amount of each element in input ( $W_{i,input}$ ) should be equal to its amount in the output ( $W_{i,output}$ ).

$$W_{i,input} = W_{i,output} \quad (5)$$

Since an element can exist in different phases, equation 5 can be rewritten as equation 6.

$$W_{i,feed} + W_{i,reducing\ gas} = W_{i,DRI} + W_{i,top\ gas} \quad (6)$$

Where  $W_{i,feed}$  is weight of element  $i$  in the feed,  $W_{i,reducing\ gas}$  is the weight of element  $i$  in the reducing gas,

$W_{i,DRI}$  is the weight of element  $i$  in DRI, and  $W_{i,top\ gas}$  is the weight of element  $i$  in the top gas.

As it can be seen in Table 1, even in one phase an element can exist in several components. So, equation 6 can be expanded to equation 7.

$$W_{feed} \times \sum_j W_{ij} X_{j\ in\ feed} + W_{reducing\ gas} \times \sum_j W_{ij} X_{j\ in\ reducing\ gas} = W_{DRI} \times \sum_j W_{ij} X_{j\ in\ DRI} + W_{top\ gas} \times \sum_j W_{ij} X_{j\ in\ top\ gas} \quad (7)$$

Where  $W_{feed}$  is weight of feed material,  $W_{reducing\ gas}$  is weight of reducing gas,  $W_{DRI}$  is weight of DRI,  $W_{top\ gas}$  is weight of top gas,  $W_{ij}$  is weight of element  $i$  in component  $j$ ,  $X_{j\ in\ feed}$  is the weight fraction of component  $j$  in feed,  $X_{j\ in\ reducing\ gas}$  is the weight fraction of component  $j$  in reducing gas,  $X_{j\ in\ DRI}$  is the weight fraction of component  $j$  in DRI, and  $X_{j\ in\ top\ gas}$  is the weight fraction of component  $j$  in top gas.

### 2.2 Energy balance

The general equation of heat balance can be written as equation 8.

$$\Delta H_{inp} + \Delta H_{out} + \Delta H_{exo} + \Delta H_{end} = 0 \quad (8)$$

Where  $\Delta H_{inp}$  is enthalpy of input components,  $\Delta H_{out}$  is enthalpy of output components,  $\Delta H_{exo}$  is enthalpy of exothermic reactions, and  $\Delta H_{end}$  is enthalpy of endothermic reactions. The heat of a reaction at temperature  $T$  can be calculated by equation 9.

$$(\Delta H_r)_T = (\Delta H_r)_{298} + \sum (H_{298}^T)_{prod} - \sum (H_{298}^T)_{react} \quad (9)$$

Where  $(\Delta H_r)_T$  is enthalpy of reaction at temperature  $T$ ,  $(H_{298}^T)_{prod}$  is enthalpy of product components by changing temperature from 298 K to  $T$ , and  $(H_{298}^T)_{react}$  is enthalpy of reactant components by changing temperature from 298 K to  $T$ .  $H_{298}^T$  can be calculated using equation 10.

$$H_{298}^T = n \int_{298}^T C_p dT \quad (10)$$

Where  $n$  is the mole number, and  $C_p$  is heat capacity or the amount of heat energy released or absorbed by a mole of the substance with the change in temperature at a constant pressure.  $C_p$  is usually calculated using equation 11 (Kubaschewski and Alcock, 1979).

$$C_p = a + bT + cT^{-2} + dT^{-1} \quad (11)$$

Where  $a$ ,  $b$ ,  $c$ , and  $d$  are thermodynamic coefficients which are different for each substance.

### 2.3 Model assumptions and considerations

Figure 2 shows the schematic flowsheet of the model. As it can be seen, the model consists of two input streams of reducing gas at 900°C and feed material which its temperature should be calculated, and output stream of top gas at 400°C and DRI at 800°C (CHATTERJEE, 2012). In the real process, temperature of shaft furnace can vary from 850°C to 950°C. So, three different temperatures of 850°C, 900°C, and 950°C have been utilized in the model.

Iron ore pellet which is currently used in direct reduction process has been selected as feed material in this study. The chemical composition of the pellet can be found in Table 2.

In the current MIDREX process, the top gas contains approximately 33-37% hydrogen and 21-24% water vapor and CO and CO<sub>2</sub> as the rest, indicating an excess of inlet hydrogen compared to the stoichiometric amount, which is attributed to the process design and considerations (Liu et al., 2014). Consequently, a ratio of 1.7 between hydrogen and water vapor ( $H_2/H_2O = 1.7$ ) has been employed in this model for the top gas. The characteristics of the top gas are detailed in Table 3.

Table 4 illustrates that the reducing gas comprises hydrogen. However, to assess the impact of water vapor on the energy consumption of the process, water vapor was incrementally added up to 10% to the system.

Assuming a 100% metallization degree, the Direct Reduced Iron (DRI) consists of metallic iron and gangue materials, as depicted in Table 5.

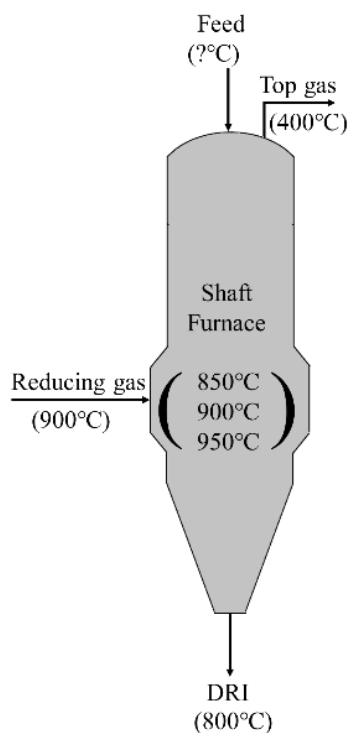


Fig. 2. Schematic flowsheet of model.

Table 2. Chemical composition of pellet (feed material).

Component	Fe <sub>2</sub> O <sub>3</sub>	Fe <sub>3</sub> O <sub>4</sub>	SiO <sub>2</sub>	CaO	MgO
Mass%	84.64	9.11	3.78	0.73	0.45
	Al <sub>2</sub> O <sub>3</sub>	TiO <sub>2</sub>	MnO		
	0.64	0.23	0.028		

Table 3. Chemical composition of top gas.

Temperature (°C)	400
H <sub>2</sub> (%)	63
H <sub>2</sub> O (%)	37

Table 4. Chemical composition of reducing gas.

Temperature (°C)	900
H <sub>2</sub> (%)	90-100
H <sub>2</sub> O (%)	0-10

Table 5. Chemical composition of DRI.

Temperature (°C)	800			
Component	Fe	SiO <sub>2</sub>	CaO	MgO
Mass%	65.79	3.78	0.73	0.45
	Al <sub>2</sub> O <sub>3</sub>	TiO <sub>2</sub>	MnO	
	0.64	0.23	0.028	

It should be noted that all calculations were carried out for one ton of feed material, and heat loss and mass loss have been neglected. Furthermore, kinetics and rate of reactions have not been considered in the calculations.

### 3. RESULTS AND DISCUSSIONS

#### 3.1 Thermodynamic studies

Figure 3 illustrates the stability diagram of the Fe-O-H system, known as the Baur-Glössner diagram. Since hydrogen reduction is endothermic, increasing the temperature broadens the stability zone of iron in the diagram. From a thermodynamic perspective, hydrogen reduction should occur at the highest possible temperature, which is also advantageous for kinetics. However, supplying the necessary energy to achieve high temperatures is economically challenging. Furthermore, the diagram shows that in the temperature range of 850-950°C, with up to over 30% water vapor in the system, regardless of kinetics, reduction to metallic iron is still possible. So, in this model 0-10% water vapor considered in the input to be sure that with the produced water vapor from the reduction reactions, it will not exceed from the stable zone of iron.

#### 3.2 Model with preheating feed materials

The heat and mass balance module of the HSC Chemistry software was utilized to model the system when heat is supplied by preheating the feed materials. Figure 4 illustrates the energy required to reduce one ton of pellets using reducing gas with varying water vapor contents. It shows that reduction with pure hydrogen demands 92 kWh at 850°C, 100 kWh at 900°C, and 110 kWh at 950°C. Adding water vapor to the reducing gas consistently decreases the required energy at all temperatures. Figure 5 presents the model results as the preheating temperatures of the feed materials with reducing

gas compositions ranging from 0 to 10% water vapor at operating temperatures of 850°C, 900°C, and 950°C. At a furnace temperature of 850°C with pure hydrogen as the reducing gas, the feed materials must be preheated to 428°C to meet the heat requirements. Raising the operating temperature to 900°C necessitates more energy, requiring the feed materials to be heated to 459°C. Although increasing the furnace temperature from 850°C to 900°C results in a 31°C rise in feed material temperature, raising the furnace temperature to 950°C requires a further increase of 37°C, bringing the feed material temperature to 496°C. This trend persists across different reducing gas compositions.

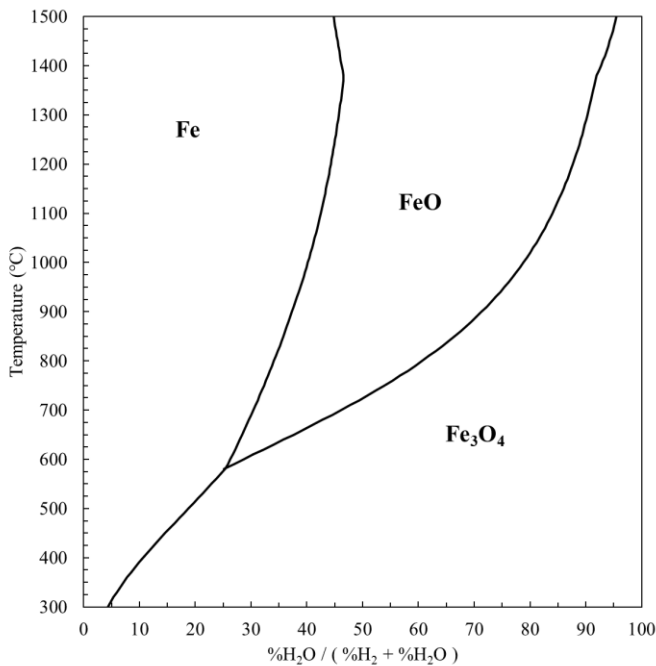


Fig. 3. Baur-Glaessner diagram for Fe-O-H system (calculated with HSC Chemistry version 10.4.1.1).

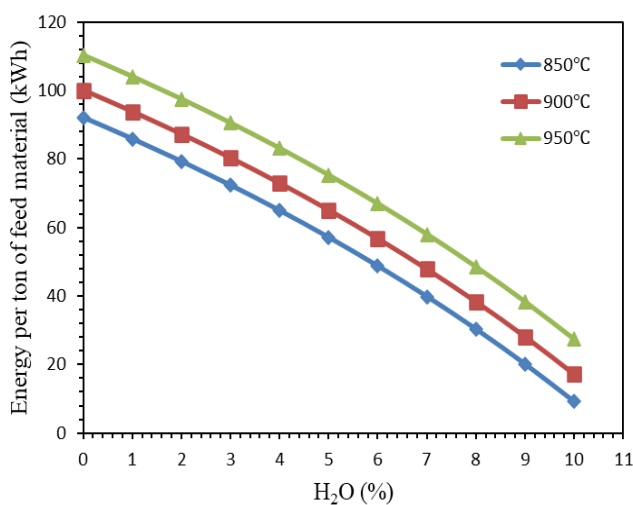


Fig. 4. The required energy to reduce one tone of pellet with different water vapor content in the reducing gas composition.

Figure 5 also shows that at all three operating temperatures, increasing the water vapor content in the reducing gas results in a decrease in the required energy and consequently lowers the preheating temperature of the pellets. To understand this phenomenon, the volume of reducing gas was monitored. Figure 6 displays the volume of reducing gas needed to reduce one ton of pellets with varying water vapor contents. It is evident that increasing the water vapor content raises the volume of reducing gas. Since the top gas composition is maintained at a constant value, as shown in Table 3, increasing water vapor content necessitates an increase in the reducing gas volume to keep the H<sub>2</sub>/H<sub>2</sub>O ratio constant as per Table 3. Consequently, because the reducing gas is warmer than the pellets, an increase in its volume supplies more heat to the system, significantly reducing the preheating temperature of the feed materials.

Furthermore, Figure 7 illustrates the heat capacity of hydrogen and water vapor at different temperatures ranging from 850°C to 950°C, according to Equation 11. It is observed that the heat capacity of water vapor is higher than that of hydrogen, and for both components, the heat capacity increases with temperature. However, the increase in heat capacity is more pronounced for water vapor than for hydrogen. This indicates that adding water vapor to the reducing gas, even while keeping the gas volume constant, contributes additional energy to the system due to the high heat capacity of water vapor.

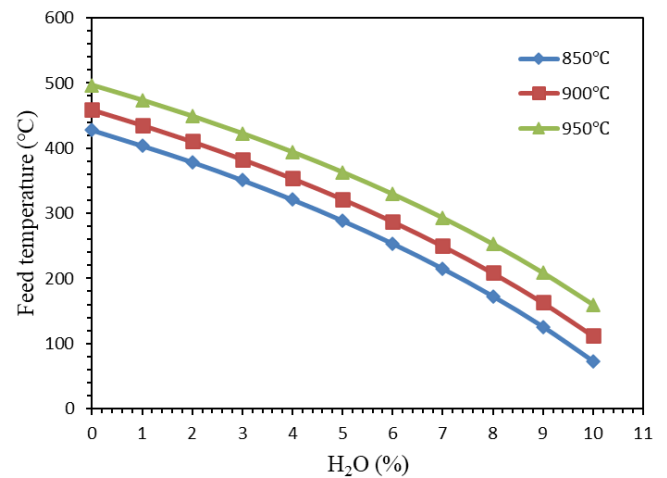


Fig. 5. Temperature of preheated feed materials with different water vapor content in the reducing gas composition.

### 3.3 Model with oxygen injection

In this model, the necessary heat for the hydrogen reduction process is generated by injecting oxygen into the furnace, facilitated through reaction 4. Figure 8 indicates that to achieve the required reduction heat using pure hydrogen at 850°C, 13.25 Nm<sup>3</sup> of oxygen must be injected. When the operating temperature is increased to 900°C, the volume of oxygen needed rises to 14.41 Nm<sup>3</sup>, and at 950°C, it further increases to 15.87 Nm<sup>3</sup>. This increase in required oxygen volume with higher temperatures underscores the endothermic nature of the

hydrogen reduction process, which demands more energy input as temperatures rise.

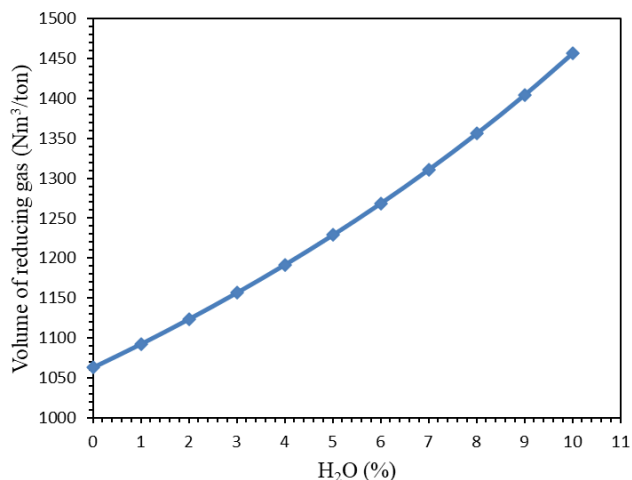


Fig. 6. Volume of reducing gas to reduce one ton of pellet with different water vapor content in the reducing gas composition.

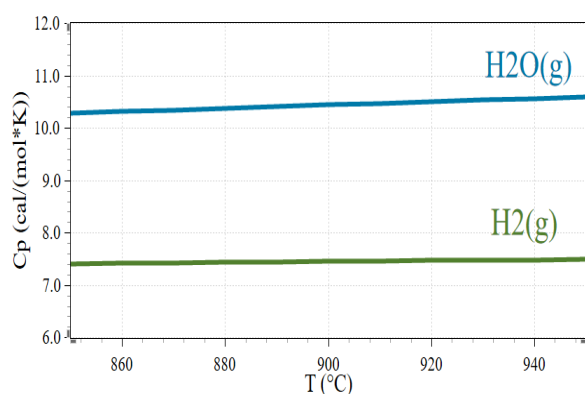


Fig. 7. Heat capacity of hydrogen and water vapor at different temperatures.

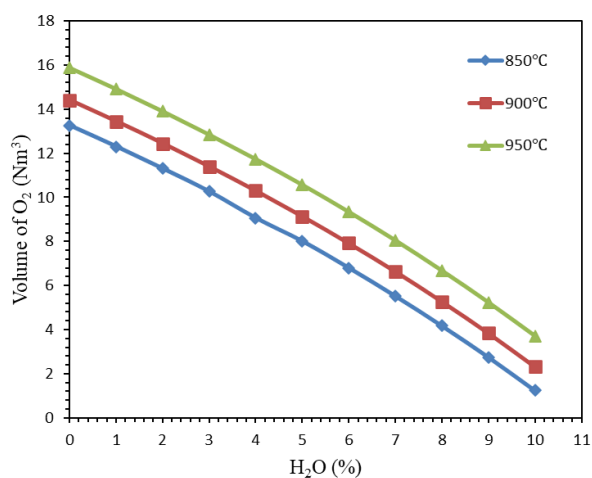


Fig. 8. Volume of injected oxygen with different water vapor content in the reducing gas composition.

Figure 9 illustrates that the volume of reducing gas increases with higher water vapor content, consistent with the earlier discussed principles. However, due to the presence of oxygen in the system, the reducing gas volume also varies with the furnace temperature. As the water vapor content rises, the necessary heat—and consequently the amount of oxygen required—decreases significantly. This reduction in oxygen demand is particularly important because it highlights the interplay between water vapor and oxygen in the system. By optimizing the water vapor content, the process can achieve substantial energy savings and enhanced efficiency.

An ironmaking plant with a capacity of 1 Mt DRI per year can be considered as a realistic example. With the same pellet and gas composition that used in this study, with the reduction temperature of 900°C and 5% H<sub>2</sub>O in the system, 65.12 GWh energy is needed for the reduction, which can be supplied by heating the feed material to 322°C or injecting  $9.14 \times 10^6$  Nm<sup>3</sup> oxygen to the system. Therefore, understanding the precise oxygen requirements at various temperatures and managing both the oxygen injection and water vapor levels are crucial for maintaining an optimal balance of heat supply, reducing overall energy consumption, and improving the economic and environmental performance of the hydrogen reduction process in ironmaking.

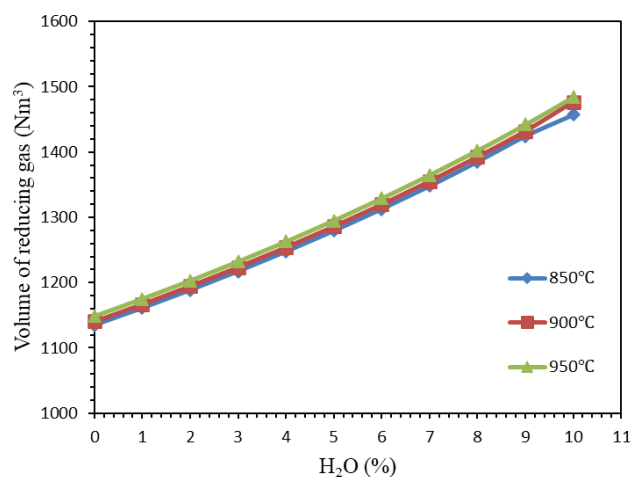


Fig. 9. Volume of reducing gas to reduce one ton of pellet with different water vapor content in the reducing gas composition and the presence of oxygen in the system.

#### 4. CONCLUSIONS

Energy consumption of hydrogen reduction process with the presence of 0 to 10% water vapor was studied in this research. Modeling has been done using the heat and mass module of HSC Chemistry software and with MIDREX condition.

This study offers an in-depth examination of the energy consumption and thermodynamic behavior of hydrogen reduction of iron ore pellets in a shaft furnace. It particularly focuses on the effects of varying water vapor content.

The study reveals that increasing the reduction temperature is thermodynamically favorable for hydrogen reduction, but it necessitates additional energy input, posing economic



challenges. The volume of oxygen required to provide the necessary heat rises with the operating temperature. For instance, 13.25 Nm<sup>3</sup> of oxygen is needed at 850°C, 14.41 Nm<sup>3</sup> at 900°C, and 15.87 Nm<sup>3</sup> at 950°C.

The study also finds that a higher water vapor content augments the volume of reducing gas to keep the top gas composition constant. However, the presence of oxygen in the system also influences the volume of reducing gas to vary with furnace temperature. An increase in water vapor content reduces the heat required and, consequently, the amount of oxygen needed.

By optimizing the water vapor content and operating temperature, the study suggests that significant improvements in energy efficiency and reduction kinetics can be achieved. This enhances the sustainability and cost-effectiveness of the process.

These findings contribute to the development of more efficient and environmentally friendly methods for iron ore reduction in the steelmaking industry. They support efforts to reduce the carbon footprint associated with steel production.

#### ACKNOWLEDGMENTS

We extend our heartfelt appreciation to the Sustainable Hydrogen - Potential for Bothnia Gulf Cluster project (20357962), funded by the European Regional Development Fund (ERDF), for its invaluable support. We also gratefully acknowledge funding from the Finnish Foundation for Technology Promotion (Tekniikan Edistämissäätiö). The financial contributions from these organizations were instrumental in making this research possible.

#### REFERENCES

- Ahmed, H., Sideris, D., and Björkman, B. (2020). Injection of H<sub>2</sub>-rich carbonaceous materials into the blast furnace: devolatilization, gasification and combustion characteristics and effect of increased H<sub>2</sub>-H<sub>2</sub>O on iron ore pellets reducibility. *Journal of Materials Research and Technology*, 9(6), 16029–16037.
- CHATTERJEE, A. (2012). *Sponge Iron Production By Direct Reduction Of Iron Oxide*. PHI Learning.
- El-Zoka, A. A., Stephenson, L. T., Kim, S. H., Gault, B., and Raabe, D. (2023). The Fate of Water in Hydrogen-Based Iron Oxide Reduction. *Advanced Science*, 10(24), 1–8.
- Fradet, Q., Kurnatowska, M., and Riedel, U. (2023). Thermochemical reduction of iron oxide powders with hydrogen: Review of selected thermal analysis studies. *Thermochimica Acta*, 726, 179552.
- Heidari, A., Niknahad, N., Iljana, M., and Fabritius, T. (2021). A review on the kinetics of iron ore reduction by hydrogen. *Materials*, 14(24), 7540.
- Kubaschewski, O., and Alcock, C. B. (1979). *Metallurgical Thermochemistry*. Pergamon Press.
- Liu, B. N., Li, Q., Zou, Z. S., and Yu, A. B. (2014). Discussion on chemical energy utilisation of reducing gas in reduction shaft furnace. *Ironmaking and Steelmaking*, 41(8), 568–574.
- Özgün, Ö., Lu, X., Ma, Y., and Raabe, D. (2023). How much hydrogen is in green steel? *Npj Materials Degradation*, 7(1), 1–5.
- Raabe, D., Klug, M. J., Ma, Y., Büyüksulu, Ö., Springer, H., and Souza Filho, I. (2023). *Hydrogen Plasma Reduction of Iron Oxides* (C. Fleuriault, J. D. Steenkamp, D. Gregurek, J. F. White, Q. G. Reynolds, P. J. Mackey, and S. A. C. Hockaday, Eds.; pp. 83–84). Springer Nature Switzerland.
- Souza Filho, I. R., Ma, Y., Kulse, M., Ponge, D., Gault, B., Springer, H., and Raabe, D. (2021a). Sustainable steel through hydrogen plasma reduction of iron ore: Process, kinetics, microstructure, chemistry. *Acta Materialia*, 213, 116971.
- Souza Filho, I. R., Ma, Y., Raabe, D., and Springer, H. (2023). Fundamentals of Green Steel Production: On the Role of Gas Pressure During Hydrogen Reduction of Iron Ores. *Jom*, 75(7), 2274–2286.
- Souza Filho, I. R., Springer, H., Ma, Y., Mahajan, A., da Silva, C. C., Kulse, M., and Raabe, D. (2022). Green steel at its crossroads: Hybrid hydrogen-based reduction of iron ores. *Journal of Cleaner Production*, 340, 130805.
- Spreitzer, D., and Schenk, J. (2019). Reduction of Iron Oxides with Hydrogen—A Review. *Steel Research International*, 90(10), 1900108.
- Sun, M., Pang, K., Barati, M., and Meng, X. (2023). Hydrogen-Based Reduction Technologies in Low-Carbon Sustainable Ironmaking and Steelmaking: A Review. *Journal of Sustainable Metallurgy*, 10, 10-25.

## Using an advanced simulation tool for successful conversion of reheating furnace to full oxyfuel operation

Esin Iplik\* Tomas Ekman\* Kristofer Bölke\*\* Otto Kankaanpää\*\*\*

\*Linde Sverige AB, Varuvägen 2, 125 30 Älvsjö, Sweden (e-mail: [esin.iplik@linde.com](mailto:esin.iplik@linde.com), [tomas.ekman@linde.com](mailto:tomas.ekman@linde.com)).

\*\*Linde Gas AB, Rättarvägen 3, 169 68 Solna, Sweden (e-mail: [kristofer.bolke@linde.com](mailto:kristofer.bolke@linde.com))

\*\*\*Ovako Imatra Oy AB, Terästehtaantie 1, 55610 Imatra, Finland (e-mail: [otto.kankaanpaa@ovako.com](mailto:otto.kankaanpaa@ovako.com))

---

**Abstract:** Oxyfuel combustion complements decarbonization efforts by reducing the energy needs in high-temperature industries. Steel reheating furnaces are good candidates for full oxyfuel operation since this can lead to up to 30% energy savings. Linde uses an in-house tool to simulate reheating furnaces for air-fuel to oxyfuel conversion. This paper follows a real customer case, starting with an airfuel simulation setup used to analyze the furnace, followed by oxyfuel simulations for burner design and energy savings estimations. These simulations lead to a successful installation of oxyfuel burners for the reheating furnace located at Ovako Imatra site. After the commissioning is completed, performance evaluation is done by comparing a reference airfuel operation period with an oxyfuel combustion period. Full oxyfuel conversion results in 27% energy savings for hot charge and high production rate periods thanks to significantly lower flue gas losses. Removing nitrogen from the oxidizer decreases the flue gas volume, reducing the total heat capacity of the off-gas stream. The savings are around 30% for cold charge and average production rate periods.

*Keywords:* oxyfuel combustion, energy savings, decarbonization, steel reheating.

---

### 1. INTRODUCTION

The decarbonization of electricity production, combined with increased electrification, has the potential to reduce the carbon footprint of low-temperature industries. However, high-temperature industries such as steel and glass production have a greater demand for heat that cannot currently be met without combustion. It is essential to develop efficient burner technologies to save energy and complement efforts to decarbonize fuel. One promising method is oxyfuel combustion, which uses industrial oxygen as an oxidizer to eliminate nitrogen from the gas mixture. This approach offers several benefits, including increased production capacity, reduced heat losses through flue gases, lower NO<sub>x</sub> emissions, and lower CO<sub>2</sub> emission levels due to fuel savings. Research on oxyfuel combustion has been conducted for the steel industry (Hu et al., 2019) and the aluminum industry (Paubel et al., 2019).

Oxyfuel combustion has some inherent benefits. At a specific energy release, it creates low volumes of flue gas that has a higher fraction of highly radiating triatomic gas molecules. Any furnace designed for airfuel combustion can be fired more efficiently using oxygen as the oxidizer. The longer residence time of the highly radiating gases allows for a very efficient heat recuperation inside the furnace's dark zone. One potential drawback of traditional oxyfuel is the risk of forming high

amounts of NO<sub>x</sub> due to the higher flame temperature. Linde has developed burners that overcome this problem by implementing semi-flameless and flameless combustion. The potential to efficiently supply more power than by using airfuel in any given furnace design also allows for the possibility to produce more if the material being heated allows for this.

Linde, with over 40 years of experience in oxyfuel combustion, is a trusted leader in the field. From studying lab-scale kinetics to converting large furnaces to 100% oxyfuel combustion, Linde has demonstrated its expertise. Reheating furnaces, being one of the most energy-demanding units in steel processing (Zhao et al., 2021), are excellent candidates for oxyfuel conversion. These furnaces, used to reheat blooms or billets to the rolling temperatures of around 1150 - 1300 °C, constitute up to 67% of the total steel energy demand (Vögele et al., 2020).

Ovako Group has been active in their decarbonization efforts since 1995 by switching reheating furnaces (Hofors, Sweden) into oxyfuel combustion as well as High-Level Lancing (HLL) of oxygen at other Ovako locations (Smedjebacken and Hällefors, Sweden). The following efforts included electrification of heat treatment furnaces and electrolyzer investments to use fossil-free hydrogen for combustion. This paper follows the conversion of Ovako Imatra's reheating furnace to full oxyfuel operation using simulations and real data for airfuel vs. oxyfuel comparisons regarding energy savings and capacity increase. These numerical simulations

allow a better understanding and an optimal design of the oxyfuel combustion system. Additionally, they predict the furnace performance with the suggested changes ensuring a more confident project execution.

Another installation of oxyfuel technology is planned at their Boxholm plant in Sweden. The Imatra plant in Finland is making a major effort to decarbonize Ovako's steel production, and the conversion of the bloom furnace is the first major leap. The converted furnace, the first large-scale reheating furnace in Finland using solely flameless oxyfuel technology, is a testament to the potential of oxyfuel conversion. This successful conversion not only enables the use of hydrogen as a fuel, but also allows Ovako's to move towards its goal of zero-carbon production.

Conversion to full oxyfuel lowers the total flue gas volume and can cause difficulties in furnace pressure control depending on the flue gas system sizing. Maintaining the furnace pressure is essential to prevent air ingress, and slight positive pressure is preferred for industrial furnaces. Therefore, some furnaces may have a minimum power limitation to generate enough flue gas volume. The flue gas system is often redesigned during the conversion, and the furnace pressure control damper is rebuilt. The target furnace of this work is a walking beam furnace used to heat blooms. During airfuel operation, it had roof-mounted burners for the top zones and lateral burners for the bottom zones. For oxyfuel operation, the roof burners are repositioned as lateral burners. The baffles between the recuperative and fired bottom zones were removed to accommodate two additional fired zones. No other major modifications were made. The furnace has an installed economizer in the stack, and it is kept in use after the oxyfuel conversion despite not using the hot combustion air anymore. Heat recovered by the economizer, as well as from cooling of walking beams, is used in district heating around the mill. The furnace pacing system stayed unchanged, and the heating control system was adapted to new zones and gas compositions.

## 2. METHODOLOGY

The simulation code used for this work is based on the blooms being evenly spaced in the furnace with no empty spaces. Even pacing of blooms is assumed for transportation through the furnace. One target condition is uniform temperature distribution within the material with less than 30 °C gradient between 10% under the top surface and the center of the bloom when it exits the furnace. Reference thermocouple measurements are taken from the furnace zones. The temperature-dependent emissivity of the furnace walls and the heated material are selected as inputs, as well as the chemical composition of the reference alloy. The model calculates the fired zone temperatures and the power used in the fired zones to achieve the material target temperatures at the selected production rate.

The furnace model is divided into zones that can be used for heating or recuperation. While heating zones are fired, recuperation zones are not. These zones are adapted to the customer furnace and correspond to the control strategy. The furnace is, then, subdivided into computational cells, as shown in Fig. 1. Each cell has two energy balances that must be fulfilled, one on the cell wall and one for the gas volume in the cell. The cell is assumed to be an ideal continuous stirred-tank

reactor, and the cell models include the radiation and convection energy fluxes from gas to wall, gas to material, wall to material, wall to surroundings, and cell to cell. The gas radiation is evaluated according to the method described in VDI-Wärmeatlas section KC (Stephan et al., 2021). The standard view factor methods compute the wall and material surface radiative interactions (Stephan et al., 2021). Convection constitutes a minor part of the heat transfer and uses the Juerges' standard forced convection equation (Mörtstedt and Hellsten, 1994). The simulation code sets up a system of differential equations based on the number of cells. The solver uses two input parameters for each cell and solves the equation system by finding the correct gas temperature or power input with wall temperatures that satisfies the set of equations with a minimum error.

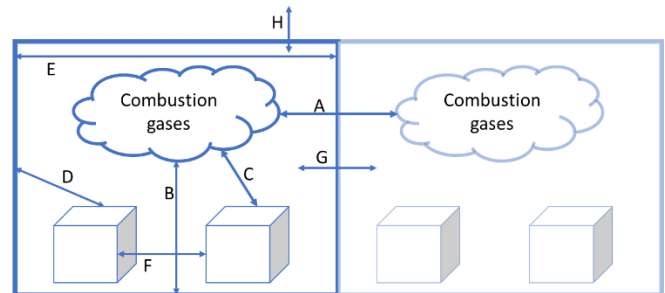


Fig. 1. Energy fluxes considered. (A: Gas to gas, B: Gas to wall, C: Gas to material, D: Wall to material, E: Wall to wall, F: Material to material, G: Cell to cell, H: Wall to surroundings.)

The material pieces are also divided into computational nodes, and the transient response of the nodes is calculated by the very efficient Thomas algorithm, modified to use non-constant material properties (Lee, 2024).

The solver generates a steady-state solution to the particular problem setup. This solution highly depends on the temperature status of all the material surfaces in the furnace. To generate this solution, the transient response of the material transported through the furnace is recalculated many times during the solution procedure. The KINSOL solver (<https://computing.llnl.gov/projects/sundials>), with Broyden root finder (Broyden, 1965), is used for the simulation.

For the Ovako Imatra case, 0.5 m furnace cells were selected, which generated 37 cells for both the top and bottom zones. This setup created 148 energy balance equations to be solved simultaneously. For computational efficiency and desired accuracy, the material node was selected to be 39 x 46 mm. The selected material dimensions were 310 x 370 x 4400 mm.

Linde and Ovako agreed on the following simulation scenarios for airfuel vs. oxyfuel comparison.

- Scenario 1: Cold charged blooms, 20 °C, 38 tph pace, ~4 hours in the furnace with target temperature of ~1250 °C
- Scenario 2: Hot charged blooms, 800 °C surface temperature, 900 °C core temperature, 75 tph pace, ~2 hours in the furnace with target temperature of 1250 °C

These simulations represent ideal production scenarios; however, the reality is usually non-ideal. The production follows the demand. Therefore, material size, characteristics,

and heating requirements change during continuous operation. It is hard to simulate the real furnace, considering the temperature, power, and production rate changes. Ideal scenarios are used to demonstrate the differences between airfuel and oxyfuel operations.

Before oxyfuel conversion, airfuel data is collected from the furnace, including fuel flow, bloom charge and discharge temperature, and production rate. Data is collected during a week and classified according to charge temperature to evaluate the energy requirements separately. The reference airfuel simulation is forced to match the reference airfuel data, and the furnace losses are calculated according to the energy balances. These losses are kept constant and used for the oxyfuel simulation to represent the furnace as close as possible to reality without expecting furnace improvements.

Data sets are described below.

- Data set 1: Cold charged blooms, below 100 °C, above 25 t/h
- Data set 2: Hot charged blooms, over 800 °C, above 55 t/h
- Data set 3: Hot charged blooms, over 800 °C, all production rates
- Data set 4: Mixed temperature charges, between 100 and 800 °C, all production rates

After the oxyfuel installation, performance evaluation data is collected for one week. Oxyfuel data is also classified as described above for comparison.

Data sets 1 and 2 are similar to the simulation scenarios; however, values differ slightly due to the very limited number of data points that satisfy the simulation conditions.

### 3. SIMULATION RESULTS

Linde creates simulations during the project evaluation period using furnace design, burner equipment, fuel selection, material heating requirements, fuel consumption, production rate, combustion air temperatures, and material grade information supplied by Ovako.

An image of the simulation interface for airfuel operation is given in Fig. 2. The fired top and bottom zones and flue gas flow direction are shown in this figure, in addition to the thermocouple locations. Material flows opposite the flue gas flow (pink arrows) direction. The furnace starts with two recuperation zones (left-hand side), followed by six fired zones (three top and three bottom). The off-gas duct is marked with red arrows, and potential air leakage is marked with blue arrows. Fired zones are shown with the yellow flame, and the thermometer indicates the location of the thermocouples.

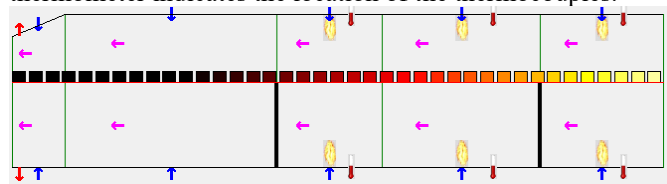


Fig. 2. Airfuel simulation setup.

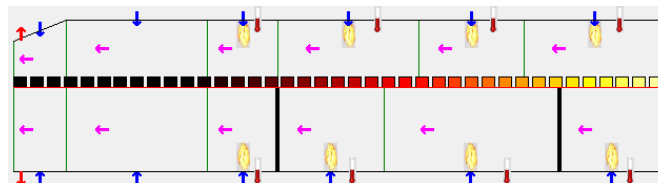


Fig. 3. Oxyfuel simulation setup.

Oxyfuel simulation presents the availability for two additional fired zones thanks to efficient recuperation of the heat in the low flue gas volume. As shown in Fig. 3, one top and one bottom fired zone are added to the furnace, shortening the recuperation zone to start heating the blooms earlier and allowing the material to distribute the heat for a longer time.

Table 1 shows the operating parameters and the resulting energy consumption. The solver is restricted by the maximum allowed wall temperature, available installed power, and the temperature homogeneity requirements of the blooms at discharge. Simulation results are reasonable with respect to the acquired furnace data.

Table 1. Airfuel simulation results as reference points.

Case	Charge T (°C)	Discharge T (°C)	Production rate (t/h)	Energy Consumption (kWh/t)
Sc. 1	20	1250	38	334
Sc. 2	800	1250	75	144

To achieve the target temperatures with oxyfuel installation, simulations are reset, and new power values are estimated using oxyfuel burners. The results of these simulations are used to calculate energy savings, as summarized in Table 2.

Table 2. Results of the oxyfuel furnace simulations.

Case	Charge T (°C)	Discharge T (°C)	Production rate (t/h)	Energy Consumption (kWh/t)
Sc. 1	20	1250	38	296
Sc. 2	800	1250	75	119

Based on the simulations, an 11% improvement is calculated for cold charging (Sc. 1), and a 17% improvement is expected for hot charging at high production rates (Sc. 2). Based on these simulations, Linde and Ovako agreed on full oxyfuel conversion of the selected furnace.

#### 4. PERFORMANCE EVALUATION

Performance comparison is based on the data obtained during the reference period prior to the oxyfuel installation and from a total production week using oxyfuel after the commissioning. Energy usage data is logged as hourly totals. For each hour, the number of tons of product discharged is evaluated to represent the production rate. The charge and discharge temperatures are averages of all the blooms extracted during the hour. Selected hours are consecutive in time to reflect continuous production and charging of blooms.

Table 3. Airfuel reference data.

Case	Charge T (°C)	Discharge T (°C)	Average Production Rate (t/h)	Energy Consumption (kWh/t)
Set1	41	1234	40	373
Set2	849	1212	62	170
Set3	835	1212	42	333
Set4	576	1220	38	366

Reference data from the airfuel operation is given in Table 3. The data shows the highest energy consumption for the cold-charged blooms, as expected, represented as Set1. The simulation results given in Table 1 and Table 2 represent a fully continuous production. On the other hand, the actual scenarios in Table 3 and Table 4 are averages of hours with higher and lower production rates, possibly including minor stops between changing production parameters. It can be expected that the actual production has a higher energy consumption than the simulation. This explains the deviation between the simulation setups and the real data for both hot and cold-charged scenarios. Furthermore, Set2 has a significantly lower average production rate than the simulation, which also contributes to higher energy consumption than the simulation.

Set3 and Set4 are not simulated but are used for performance evaluation. Set4 is a mix of higher and lower charge temperatures, and the energy consumption value lies between the cold (Set1) and the hot charge (Set3). As expected, and simulated, higher charge temperatures require lower energy consumption per ton of product to reach the target discharge temperature. Changing the production rate affects consumption significantly. The subset taken from Set3 is given as Set2, which includes only the highest production rates. The lowest energy consumption is achieved during these high throughput hours. This high efficiency has an upper limit defined by the combustion system, furnace design, charge material, and production requirements.

The target furnace is converted to full oxyfuel operation by Linde according to the simulations with the additional fired zones. Oxyfuel-fired burners are shown in Fig. 4. After the commissioning, some blooms are charged with

thermoelements installed on them to measure the temperature distribution in the material, as shown in Fig. 5.

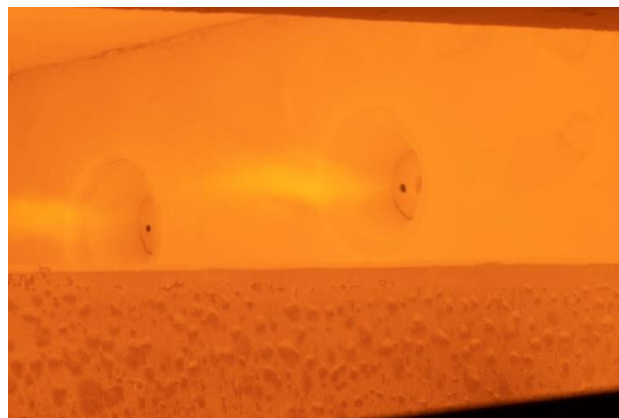


Fig. 4. Oxyfuel burners as installed with heated bloom in the foreground.



Fig. 5. Bloom with thermoelements being discharged from the furnace.

The oxyfuel performance data is presented in Table 4.

Table 4: Oxyfuel performance data.

Case	Charge T (°C)	Discharge T (°C)	Average Production Rate (t/h)	Energy Consumption (kWh/t)
Set1	14	1247	36	261
Set2	871	1236	61	123
Set3	830	1244	43	163
Set4	622	1237	33	241

Oxyfuel and airfuel follow the same trend for all the scenarios. As expected, oxyfuel energy consumption is consistently lower than airfuel. The highest energy consumption belongs to the cold-charged material (Set1). Despite a 10% lower production rate, 25 °C lower average charge temperature, and 10 °C higher discharge temperature of oxyfuel operation, energy savings of 30% are achieved for cold-charged blooms reheating. Simulation setup 1 represents a similar case to Set1. The simulation shows an improvement of only 11% in terms



of energy savings. As explained earlier, this value does not account for any furnace improvements. However, extensive commissioning maintenance carried out during the oxyfuel installation at the Imatra site should contribute to this deviation, together with the inherent ability of oxyfuel systems to quickly adapt to varying production scenarios. Oxyfuel systems are more agile than airfuel systems due to their higher heat transfer capability. Furnaces with varying production rates benefit oxyfuel combustion greatly. It shall also be noted that the simulation software is designed to be slightly conservative in its estimations of oxyfuel performance so as not to promise excessive energy savings.

A considerable 27% energy saving is achieved by oxyfuel conversion for the high production rates of the hot-charged material (Set2). For hot charging, the high flue gas temperatures penalize airfuel systems more than oxyfuel systems due to the high flue gas volume of airfuel combustion. Simulation setup 2 represents this scenario best, and it estimated an improvement of 17%, which is lower than the actual performance improvement of 27%. This deviation can be explained by the same reasons described above. Additionally, the average charge temperature of the oxyfuel Set2 is 20 °C higher than the airfuel case.

The highest energy savings are observed for the hot charge in Set3, at 51%. While the charge temperature and production rates are comparable for airfuel and oxyfuel Set3, the oxyfuel combustion data set has a 30 °C higher average discharge temperature. The possible explanation for the very high energy savings, in addition to the aforementioned furnace maintenance, is likely the higher variance in the production rates for the airfuel reference period compared to the oxyfuel period.

Set4, with the mixed charge temperatures, shows a 34% improvement after the oxyfuel conversion. Oxyfuel data has the advantage of receiving a higher-temperature charge. However, it also has a higher discharge temperature and a lower average production rate, balancing the comparison.

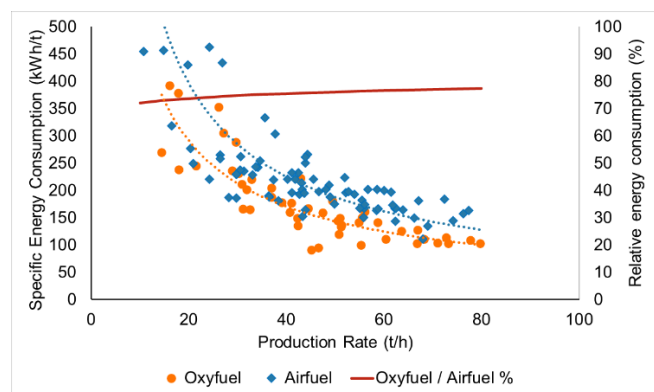


Fig. 6. Performance evaluation data for hot charge.

The hourly performance evaluation data for hot charge is plotted against the production rate for both airfuel and oxyfuel reference periods, given in Fig. 6. Percent improvement is plotted based on the trendlines. The benefit of oxyfuel slightly increases with the increasing production rate. If the airfuel-

fired furnace is run at a much higher production rate, the furnace efficiency will drop. This effect can be seen in Fig. 6. At rates above 70 t/h, all data points are found above the trendline that best fits the data set.

Based on the overall performance evaluation data, total energy savings for the reference period are estimated at 43%. It is worth mentioning that the economizer and cooling water heat recovery gains are not included in the efficiency calculations of oxyfuel operation, which translates to a further 10% savings of the total supplied energy.

## 5. CONCLUSIONS

Oxyfuel combustion is an efficient method to save energy in steel reheating furnaces. Increasing the heat transfer capacity and lowering the flue gas volume, hence the energy loss to the atmosphere, these systems have a proven track record of achieving higher performance against traditional airfuel combustion. This work shows both simulation and performance evaluation of oxyfuel and airfuel combustion at Ovako Imatra blooms reheating furnace. The simulation scenario showing a 17% improvement for high production rates of hot-charged material is supported by the real data showing a 27% improvement after complete oxyfuel conversion. Finally, 43% energy savings is achieved during the reference oxyfuel evaluation period.

## REFERENCES

- Broyden, C. G. (1965). Class of Methods for Solving Nonlinear Simultaneous Equations. *Mathematics of Computation*, 577 - 593.
- Hu, Y., Tan, C., Niska, J., Chowdhury, J. I., Balta-Ozkan, N., Varga, L., Roach, P. A., and Wang, C. (2019). Modelling and simulation of steel reheating processes under oxy-fuel combustion conditions - Technical and environmental perspectives. *Energy*, 730-743.
- Lee, W. (Retrieved July 12, 2024). *Tridiagonal matrices: Thomas Algorithm*. Retrieved from <http://www.industrial-maths.com/>
- Mörtstedt, S.-E., Hellsten, G. (1994). *Data och diagram*. Stockholm: Liber Utbildning.
- Paubel, X., Rheker, F., Juma, S., Jepson, S., Wieck, and D., Ollerton, B. (2019). Oxy-fuel Technologies for Improved Efficiency in Aluminum Scrap Melting. *Light Metals* (págs. 1165-1172). San Antonio: Springer.
- Stephan, P., Kabelac, S., Kind, M., Mewes, D., Schaber, K., and Wetzel, T. (2021). *VDI-Wärmeatlas*. Berlin: Springer.
- Vögele, S., Grajewski, M., Govorukha, K., and Rubbelke, D. (2020). Challenges for the European steel industry: Analysis, possible consequences and impacts on sustainable development. *Applied Energy*, 114633.
- Zhao, J., Ma, L., Zayed, M. E., Elsheikh, A. H., Li, W., Yan, Q., and Wang, J. (2021). Industrial reheating furnaces: A review of energy efficiency assessments, waste heat recovery potentials, heating process characteristics and perspectives for steel industry. *Process Safety and Environmental Protection*, 1209-1228.

## Computational Designing Approach for Medium Manganese Steels with Potential Better Hydrogen Embrittlement Resistance

Mahmoud Elaraby <sup>\*,\*\*</sup>, Mohammed Ali <sup>\*,\*\*</sup>, Mamdouh Eissa <sup>\*\*</sup>, Jukka Kömi <sup>\*</sup>, Pentti Karjalainen <sup>\*</sup>,  
Vahid Javaheri <sup>\*</sup>

<sup>\*</sup> *Materials and Mechanical Engineering, Centre for Advanced Steels Research, University of Oulu, P.O. Box 4200, FI-90014 Oulu, Finland (e-mail: [Mahmoud.elaraby@oulu.fi](mailto:Mahmoud.elaraby@oulu.fi), [mohammed.ali@oulu.fi](mailto:mohammed.ali@oulu.fi), [jukka.komi@oulu.fi](mailto:jukka.komi@oulu.fi), [pentti.karjalainen@oulu.fi](mailto:pentti.karjalainen@oulu.fi), [vahid.javaheri@oulu.fi](mailto:vahid.javaheri@oulu.fi)).*

<sup>\*\*</sup> *Steel Technology Department, Central Metallurgical Research and Development Institute, Helwan, 11421, Cairo, Egypt (e-mail: [mamdouh.eissa@gmail.com](mailto:mamdouh.eissa@gmail.com)).*

**Abstract:** Medium manganese steels (MMnS) are known as third-generation high-strength steels, providing an excellent balance of high strength and ductility at a lower cost than those of the second-generation steels. However, the increasing demand for steels with improved hydrogen embrittlement resistance highlights the need for the effective development of new alloys. This study explores the computational design of MMnS with a better combination of strength, ductility, and hydrogen embrittlement resistance. Mechanical properties vary mainly due to changes in chemical composition and processing routes. Computational approaches enable precise optimization of these parameters, avoiding the inefficiencies of traditional trial-and-error. Therefore, CALPHAD-based thermodynamic calculations using Thermo-Calc (TCFE12, MOBFE 7) and JMatPro 14.1 software were employed to design a novel MMnS chemistry, increasing the fraction and stability of the retained austenite and providing efficient traps for hydrogen. As a result, the optimised chemical compositions were determined to be (in wt.%): 0.35C, 9Mn, 1Si, 1Mo, 1 and 3 Al, 0.1Nb, and 0.35C, 9Mn, 1Si, 1Mo, 3Al, 0.05Nb and 0.3V. Thermo-Calc precipitation simulations identified 0.1% Nb as optimal since higher Nb contents reduce carbon in austenite, lowering its stability, and increase the size of the carbides. This Nb content results in NbC formation with an average size distribution around 1 nm, 36 nm, and a size distribution of  $1.2 \times 10^{30}$ , and  $5.4 \times 10^{27}$  respectively. 3% Al promotes the delta ferrite formation and avoids the formation of kappa carbides, and 1% Mo compromises the volume fraction of NbC, strengthening the alloy and serving as an effective hydrogen deep trapping site. 0.3% V was chosen, compromising its effects on the size distribution of VC and available C for the austenitic phase, improving its mechanical stability.

**Keywords:** Medium Manganese Steels, Hydrogen Embrittlement Resistance, Computational Design, CALPHAD-based Thermodynamic Calculations.

### 1. INTRODUCTION

Steels with high strength – high ductility, and improved hydrogen embrittlement (HE) resistance have become mandatory in the current era regarding energy applications. The first-generation steels had a relatively high strength but low ductility, which was enhanced in the second-generation steels at the expense of high-cost alloying elements. The third-generation steels, including medium manganese steels (MMnS), compromise the drawbacks of the previous two generations, having better ductility with lower costs (Sun et al., 2023). The physical and mechanical properties of MMnS vary widely depending on alloying elements as well as processing routes (Suh and Kim, 2017). MMnS can have a wide range of chemical compositions (in the following, all concentrations are in wt.%): C 0.01~0.7 (Zhao et al., 2016), Mn 3~12 (Sun et al., 2023), Al 0~10 (Zhao et al., 2016), Si 0~3 (Sun et al., 2018, 2019). Mn is the primary alloying element, and it mainly affects the phase fraction, stability, and morphology of the austenite, defining the mechanical properties and HE resistance. The critical parameter to achieve strength without

sacrificing the ductility is retained austenite; that is why it is vital to control its fraction, stability, and morphology achieved by optimal alloying elements, mainly C and Mn content, in addition to processing parameters. Carbon is an austenite stabilizer as it retards the martensitic transformation temperature and decreases the  $A_{e1}$  and  $A_{e3}$  temperatures. Increasing C content might be beneficial in stabilizing the austenitic phase, though it would deteriorate the weldability, form coarse precipitates, impair the mechanical properties, and lead to intergranular fracture. As indicated by Sun et al. (Sun et al., 2023) Mn serves as the main alloying element as it defines the deformation mechanism by controlling the stacking fault energy (SFE) of the alloys; below  $18 \text{ mJ/m}^2$ , the transformation-induced plasticity (TRIP) mechanism will be dominating, and higher than  $45 \text{ mJ/m}^2$  will activate the dislocation slip, and in the range of 18 to  $45 \text{ mJ/m}^2$ , will trigger the twinning induced plasticity (TWIP) deformation mechanism. Unlike C, Mn has low diffusivity inside austenite, achieving outstanding mechanical properties through controlling the microstructure via adequate processing parameters of intercritical annealing temperatures and soaking

times (Sun et al., 2016; Ding et al., 2020). Al stabilizes delta ferrite, which was reported to be beneficial for HE resistance for three reasons. First, Al atoms can reduce H diffusivity in austenite due to local expansion around them. Second, enriching the surface oxides prevents H permeation to the alloy. Third, delta ferrite, which Al mainly stabilizes, arrests the propagated cracks, thus retarding failure and enhancing the material toughness. It also has a critical role in controlling the deformation mechanism as it increases the SFE by 8.5 mJ/m<sup>2</sup> (Kang et al., 2012; Lehnhoff et al., 2014). Si addition like Al stabilizes delta ferrite. Wang et al. (Wang et al., 2022a) reported enhancing total elongation (TEL) from 38% to 68% after adding 0.6% Si to 0.2C-6Mn-3Al. Si inhibits cementite precipitation, thereby providing higher carbon content for austenite, enhancing its stability, and reducing dynamic recovery of ferrite, thus increasing the strain hardening rate. Microalloying elements, Mo, Nb, and V, contribute to yield strength and ultimate strength by grain refining and precipitation hardening. They also enhance hydrogen embrittlement resistance; Mo was reported to improve grain boundary cohesion (Yoo et al., 2021) and V and Nb act as active trapping sites for hydrogen (Bhadeshia, 2016).

The traditional methods of achieving required steel properties have been both time-consuming and resource-intensive, underscoring the need for computational approaches in alloy design. Despite their critical importance, studies in this area have been limited, yet the existing research highlights significant benefits. Techniques such as CALPHAD have revolutionized the development of steels by allowing for precise simulations of alloy behaviour and accurate predictions of material outcomes (Patra et al., 2021). These computational methods enhance the efficiency of customizing steel properties, reducing the need for extensive experimental testing. They enable rapid alloy prototyping and optimization, achieving an ideal balance between ductility, strength and cost, and enhancing resistance to HE (Huang et al., 2022).

Despite decades of research (Miller, 1972), industrial production of MMnS with 3–12 Mn still faces significant challenges. These challenges are primarily economic and technical, related to scaling up conventional blast furnace processes, as seen in Voestalpine's trials (Steineder et al., 2019). Alternative electric arc furnaces with argon oxygen decarburizing or vacuum oxygen decarburizing units are necessary but costly. This method also struggles with low Mn recovery and high costs for essential additives such as low-carbon ferromanganese. Processing challenges include maintaining precise control over casting to avoid defects and managing rolling processes to mitigate cracking and other deformities. Weldability issues further complicate MMnS adoption in the automotive industry, with solutions such as introducing nickel interlayers in welds proposed to enhance joint quality (DiGiovanni et al., 2021). Advancements in process optimization, cost efficiency, and microstructural control are crucial to moving MMnS toward commercial viability. Transitioning from pre-commercial stages to total industrial production may take about ten years. This process will require cooperation among steelmakers, automotive manufacturers, and researchers to create viable and cost-effective solutions for broad implementation.

The present study utilized different modules of Thermo-Calc and JMatPro to optimize the alloy's chemical composition. The aim was to increase the fraction and stability of retained austenite by optimizing annealing temperature and time, thereby improving mechanical properties and enhancing HE resistance.

## 2. METHODOLOGY

The CALPHAD methodology employs Thermo-Calc and JMatPro to design steel alloys by optimizing the effects of alloying elements such as C, Mn, Al, Si, Mo, Nb, and V. Annealing temperatures ranged from 500–800°C, with a total of 2,880 alloy compositions assessed, as detailed in Table 1. Specifically, Thermo-Calc (TCFE 12, MOBFE 7) utilized the TC and PRISMA modules. These modules facilitated the simulation of the effect of alloying on the characteristics of the austenitic phase, including the martensite start temperature *M<sub>s</sub>* and the austenite volume fraction. These parameters were studied to obtain optimal retained austenite fraction and stability as well. Moreover, the PRISMA module, which employs the Langer-Schwartz theory (Langer and Schwartz, 1980) and the Kampmann-Wagner (Wagner et al., 2001) numerical method to simulate the nucleation, growth, and coarsening of precipitates in multi-component systems, was used to simulate the precipitation behaviour (Chen et al., 2014). This includes predictions of the mean size of precipitates, volume fraction, size distribution, and the contribution of precipitation strengthening to the yield strength, considering varying alloy compositions with temperature and time. This extension of classical nucleation theory enables accurate modelling of real system precipitations based on appropriate thermodynamic and kinetic databases. Furthermore, JMatPro was utilized to investigate the alloy's mechanical and physical properties through the solidification, thermo-physical properties, and phase transformation modules, thereby analysing the diffusion of elements over time and the effect of cooling rates on phase formation. This integrative approach underscores the pivotal role of computational tools in advancing our understanding and development of high-performance steel alloys. The simulation graphs were made to study the effects of alloying elements using 0.35C-9Mn-1Si-1Mo-3Al-0.1Nb (with designated code 310) and 0.35C-9Mn-1Si-1Mo-3Al-0.05Nb-0.3V (with designated code 353) where the numbers represent Al, Nb, and V respectively.

## 3. RESULTS AND DISCUSSION

### 3.1 Carbon

Carbon is one of the main alloying elements in MMnS; it affects the volume fraction and stability of austenite as well as the formation temperature of precipitates and their fraction. The effect of carbon on the austenite volume fraction, shown in Fig. 1, can be divided into two distinct stages. In the first stage, the austenite fraction exhibits a parabolic curve with increasing temperature, indicating a progressive increase until the cementite particles are completely dissolved, as seen from Fig. 2. In the subsequent stage, the rate of increase in the austenite volume fraction follows a linear relationship with temperature, as seen in Fig. 1. Cementite formation greatly

Table 1 Range of alloy compositions (wt.%) used in the simulation for optimizing the chemical composition.

	Lower limit	Upper limit	Step
Temperature (°C)	500	800	2
C	0.1	0.4	0.1
Mn	7	11	1
Si	1 (constant)		
Al	1	3	1
Mo	0.5	2	0.5
Nb	0.05	0.2	0.05
V	0.2	0.4	0.1
Fe	balance		
Total number of alloys	2880		

affects the austenite volume fraction due to the reduction of carbon and manganese contents.

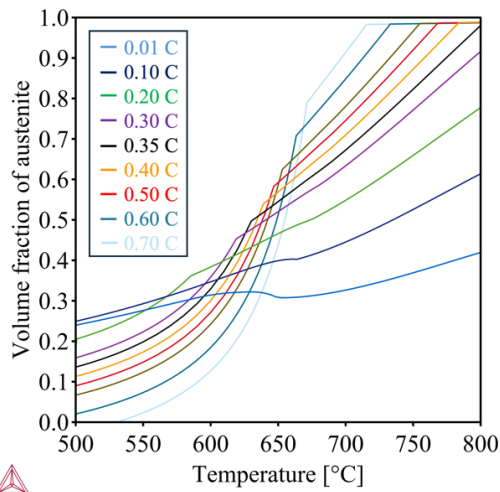


Fig. 1 Effect of carbon content on the austenite volume fraction in the temperature range of 500-800°C for alloy 310.

Increasing the carbon content to some extent not only enhances the ultimate strength due to the transformed hard martensitic phase and induces high dislocation density to the soft ferritic phase through the TRIP effect (Ennis et al., 2017) but also improves the material's ductility by increasing the mechanical stability of austenite, enabling it to accommodate more strain during the progressive work hardening thereby retarding the necking (localization of deformation) and enhancing homogeneous deformation (Zou et al., 2021). Below 0.2% C, the carbon content is relatively low, and so is the austenite volume fraction, as shown in Fig. 1; above 0.40% C, it would negatively impact the toughness due to the formation of brittle phases like cementite. At these levels, there is also a risk of reduced weldability and increased susceptibility to cracking under stress (Bhattacharya et al., 2019). A range of 0.2–0.4% C was selected to compromise the austenite volume fraction and stability. Regarding the studied effect of 0.2–0.4% C on the ultimate tensile strength and martensite start temperature  $M_s$  with annealing time that represents the austenite stability shown in Fig. 3. It is seen that 0.35% C shows a reliable austenite stability.

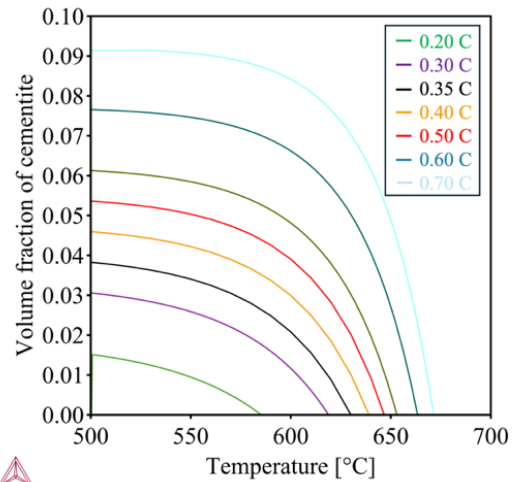


Fig. 2 Effect of carbon content on cementite's formation temperature and fraction, alloy 310.

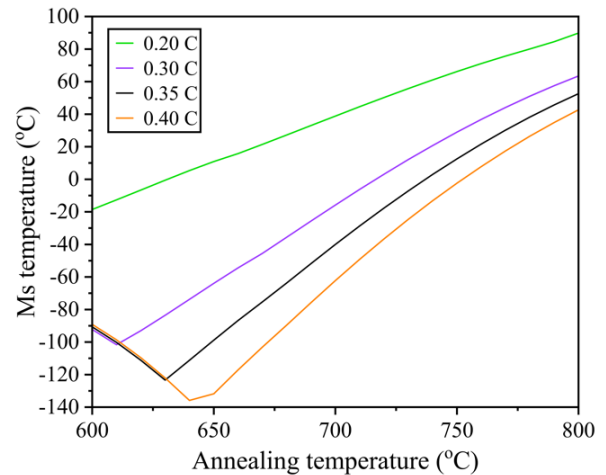


Fig. 3 Effect of carbon content on  $M_s$  temperature of intercritically annealed austenite at different annealing temperatures.

### 3.2 Manganese

Manganese is the key role element in MMnS stabilizing austenite phase and controlling the deformation mechanism through manipulating the SFE to activate both TRIP+TWIP mechanisms so that the material can accommodate higher strain levels without the deformation localization resulting in material fracture and improving both yield strength through grain refining of twinning effect and enhanced ultimate tensile strength by the TRIP effect. Previous studies have shown outstanding mechanical properties for Mn content 7–11%. (Sun et al., 2023; Zhang et al., 2024), reaching a product of strength and elongation (PSE) of 64 GPa% (Luo and Dong, 2016) for 9% Mn. Increasing Mn content will increase the fraction of the austenitic phase, reducing C content due to the dilution effect and suppressing the formation of carbides such as cementite by preferentially stabilizing carbon in the austenite phase, as shown in Figs. 4 and 5.



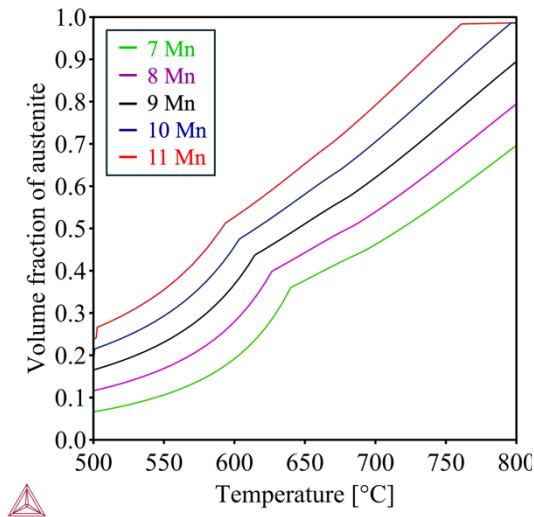


Fig. 4 Effect of Mn content on austenite volume fraction within the 500–800 °C temperature range, alloy 310.

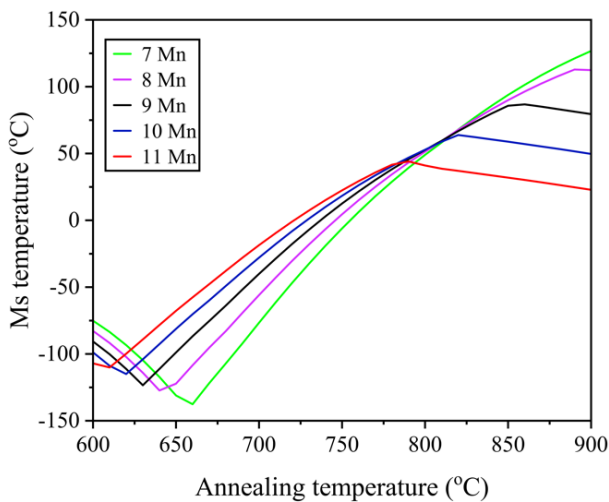


Fig. 5 Effect of Mn content on Ms temperature, alloy 310.

This suppression is beneficial for retaining ductility and toughness, as excessive cementite can embrittle the steel (Kozłowska et al., 2019). Austenite stability is affected by both C and Mn content; austenite stability is highest (lowest Ms temperature) when C content reaches its maximum value, then austenite stability increases (Ms decreases) to its minimum due to the increased carbon content resulting from cementite dissolution, then with increasing annealing temperature stability decreases (Ms increases) due to dilution effect (increased austenite volume fraction) till it is completely austenite ( $Ac_3$ ) as shown in Figs. 5 and 6. Optimizing austenite volume fraction, stability (represented by Ms), and alloying addition cost; 9% Mn was chosen for the alloys.

### 3.3 Silicon and Aluminum

Previous studies showed that 1–3% Al does not form delta ferrite, while 3–5% Al is high enough to obtain delta ferrite, but more than 5% results in kappa-carbide formation as shown in Fig. 7. Hence, it is critical to have the proper Al content for obtaining delta ferrite and avoid formation of kappa-carbide,

which would deteriorate the mechanical properties (Zhang et al., 2024). Al stabilizes ferrite and increases SFE per  $8.5 \text{ mJ/m}^2$  by 1%, which is beneficial for controlling the deformation mechanism (Sun et al., 2023), resulting in improved HE resistance through reducing the fraction of transformed martensite (Ryu et al., 2013). Al can affect the mechanical stability of austenite due to the heterogeneity of Mn content in austenite. This is explained by the slower velocity of interface movement, which decreases from  $(1.54 \times 10^{-11} \text{ to } 7.24 \times 10^{-12} \text{ m s}^{-1})$ , with the addition of ( $\sim 2\%$  Al). Consequently, there is a successive austenite to martensite transformation from regions of a low Mn content to those of a higher Mn content. This transformation improves the material strength without losing material toughness (Ye et al., 2022). Al was reported to reduce HE (Koyama et al., 2017) through two possible mechanisms: first, the strain field around Al atoms, acting as weak trapping sites, reduces the diffusion of hydrogen atoms (Han et al., 2015; Song et al., 2014), and secondly, it reduces the hydrogen permeation in the material by enriching surface oxides (Park et al., 2012).

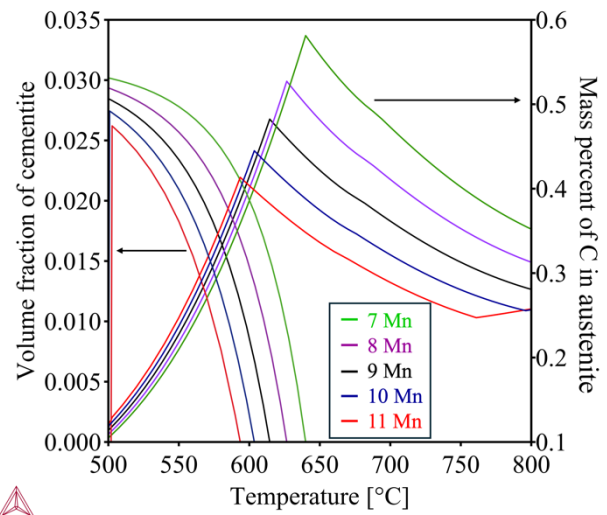


Fig. 6 Effect of Mn content on cementite fraction, formation temperature, and C in austenitic phase, alloy 310.

Two levels of Al, 1% and 3%, were chosen to study their effect on mechanical properties, the formation of delta ferrite, and the HE resistance.

Si has a similar effect to Al as it restrains cementite formation, thereby providing enough C content for austenite stabilization (Kwok and Dye, 2023). The addition of Si also reduces the dynamic recovery of ferrite, thus increasing the strain-hardening rate (Ma, 2017). It was reported that adding 0.6%Si enhanced TEL by 30% and enhanced the HE resistance as well (Wang et al., 2022b). In another study, effects of 0, 1, and 3 Si were studied in 0.2C-10Mn-3Al alloy, and it was found that 1% Si showed enhanced TEL, ultimate strength, and work hardening rate that can be attributed to the strain partitioning between austenite and ferrite. This was the main reason for the higher and uniform TEL and higher work hardening rate in 1 Si, unlike in 3 Si, where no strain partitioning was noticed. Regarding the literature survey, the alloys with 1% Si content was selected.



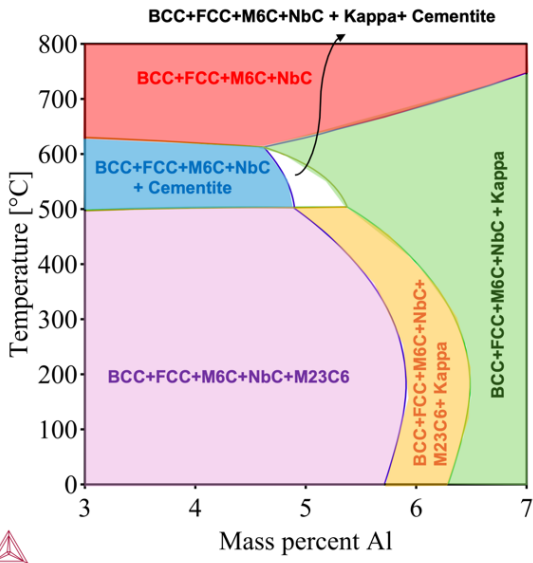


Fig. 7 Effect of Al content on kappa carbide formation, alloy 310.

3.4 Niobium and Molybdenum

Niobium exists in steel in two forms: the first and most familiar one is the carbide/carbonitride form, as Nb has a high affinity for carbon, and the second one is as a solid solution in the matrix, which can be achieved through proper processing routes (Ali et al., 2024). Both forms improve steel mechanical properties at high temperatures through the drag effect of Nb by inhibiting grain growth. At relatively low temperatures, grain growth is hindered, and hydrogen embrittlement resistance is enhanced by the pinning effect of NbC particles, which act as an active trapping site for hydrogen (Wei and Tsuzaki, 2012).

Wang et al. (Wang et al., 2010) investigated the influence of Mo on the formation and morphology of NbC and VC precipitates in steel. Their study revealed that the addition of Mo led to a reduction in the size of NbC precipitates, attributed to the segregation of Mo at the carbide/matrix interface. This segregation impedes the diffusion of Nb and V atoms from the matrix into the carbide phase and retards the diffusion kinetics of carbon and the alloying elements within the matrix. Consequently, the growth of the NbC and VC particles is significantly suppressed. The hydrogen trapping capacity of VC, NbC, and TiC was investigated through thermal desorption spectroscopy, and NbC was reported to have the largest trapping capacity (Wei and Tsuzaki, 2012). NbC was reported to have a beneficial effect on reducing the number of Σ3 boundaries, which reduces hydrogen-induced cracking (HIC) resistance (Zhang et al., 2018), and increasing the Σ11 boundaries, which increases the HIC resistance (Venegas et al., 2009). In an investigation on the effect of Nb and/or Mo addition on retained austenite volume fraction, stability, and hydrogen embrittlement resistance, the combined addition of Nb and Mo was found to result in the highest austenite volume fraction and stability (lowest Ms temperature) with the best HE resistance (Luo et al., 2022). Various empirical studies (Geng et al., 2000, 2001) using first-principles full-potential

linearized augmented plane-wave method calculations were performed to investigate the effect of elements on Fe Σ3 grain boundary cohesion, revealing that Nb and Mo have embrittling values  $\Delta E_B^A$  (eV) -1.24 and -0.96 (lower is better) respectively. The effect of Mo on Σ3 boundary was experimentally investigated (Yoo et al., 2021) to confirm the positive effect of Mo on enhancing grain cohesion and HE resistance. Precipitation of NbC was simulated using PRISMA. The results showed that 0.10 Nb is the optimal content as 0.05 and 0.2 have larger particle size and lower number density than 0.10 Nb as shown in Fig. 8. Increasing the Nb content increases the volume fraction of NbC as shown in Fig. 9. It was reported that a larger number of carbides with smaller sizes is more effective for trapping hydrogen atoms (Liu et al., 2024). While increasing Mo content has a beneficial effect on increasing the volume fraction of NbC and lowering its size, as mentioned before, it also has a negative effect on the austenite as it reduces the austenite volume fraction by reducing the available carbon, which stabilizes the austenite as illustrated in Fig. 10.

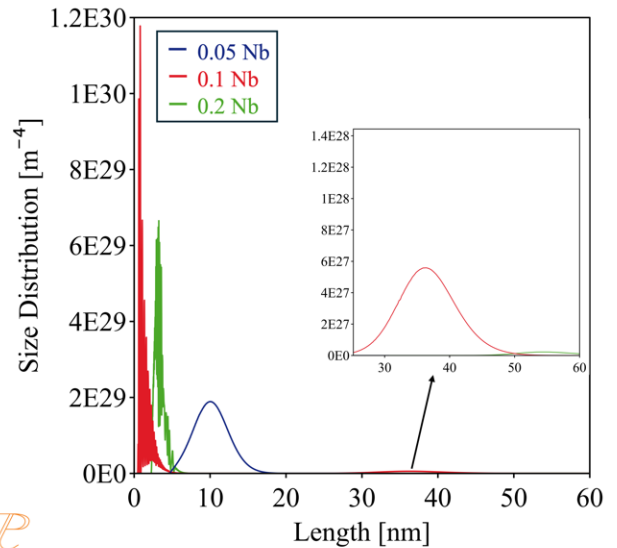


Fig. 8 Effect of Nb content on size distribution on NbC, alloy 310.

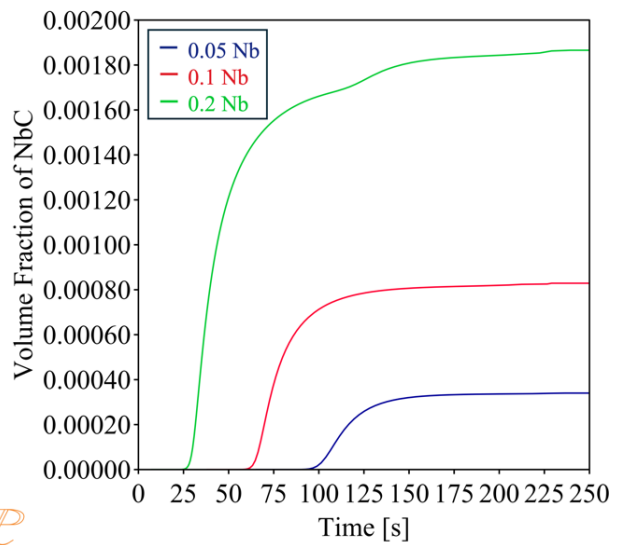


Fig. 9 Effect of Nb content on volume fraction of NbC, alloy 310.

3.5 Vanadium

The effect of V content on austenite stability through its impact on C in the austenitic phase was simulated in Fig. 11, showing that increasing V content leads to a decrease in the available C of the austenitic phase, resulting in lower stability, thus deteriorating the HE resistance. On the other hand, increasing the V content from 0.2% to 0.4% increases the size distribution of VC from  $4.8 \times 10^{30}$  to  $3.4 \times 10^{31}$ , as shown in Fig. 12, enhancing the HE resistance. Compromising the effects of V content on austenite stability and size distribution of VC, 0.3% V was chosen along with 0.05% Nb to obtain C content high enough to increase austenite stability and nano-carbide precipitates for improving both mechanical properties and HE resistance.

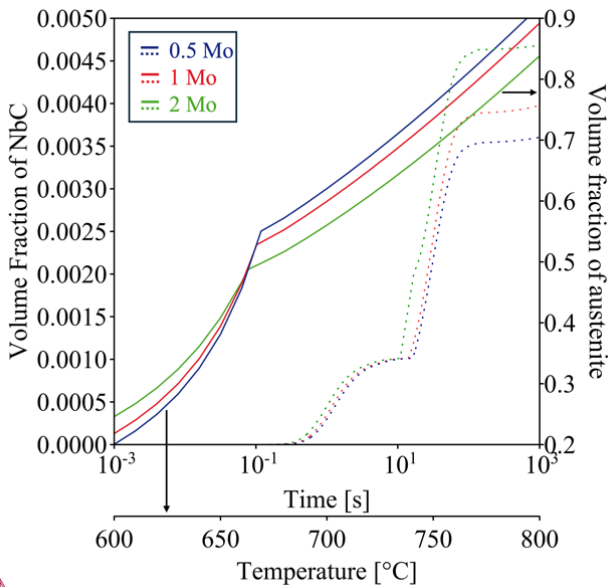


Fig. 10 Effect of Mo on precipitation of NbC and austenite volume fraction, alloy 310.

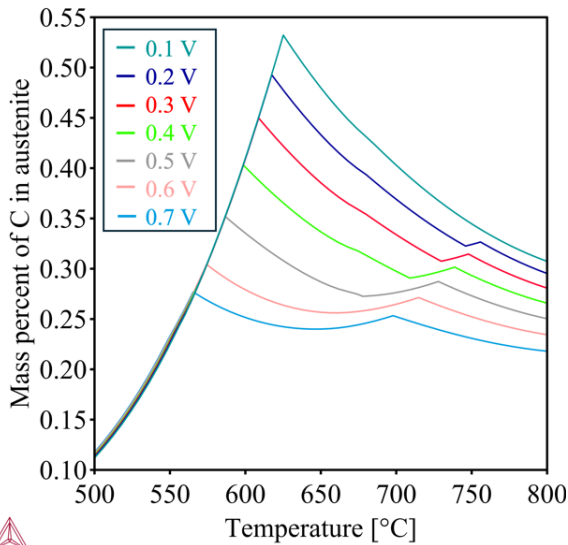


Fig. 11 Effect of V content on C concentration in austenite, alloy 353.

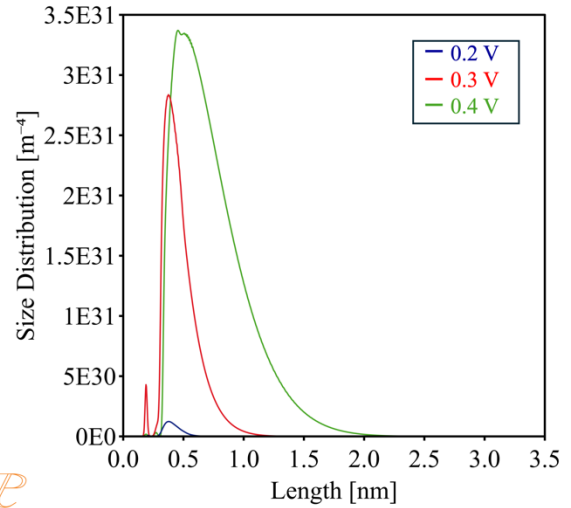


Fig. 12 Effect of V content on the size distribution of VC, alloy 353.

4. CONCLUSIONS

This study presents a comprehensive computational design approach for medium manganese steels aimed at enhancing mechanical properties and hydrogen embrittlement resistance. Through CALPHAD-based thermodynamic calculations and JMatPro simulations, we optimized the chemical composition to achieve a balance of high strength, ductility, and resistance to hydrogen embrittlement, as illustrated in the flow chart in Fig. 13.

Key findings include:

**Optimized Composition:** The optimized chemical compositions, consist of (in wt.%):

- 0.35C, 9Mn, 1Si, 1Mo, with variations of 1 and 3 Al, and 0.1 Nb.
- 0.35C, 9Mn, 1Si, 1Mo, 3 Al, 0.05 Nb and 0.3 V.

as optimal for enhancing mechanical properties and hydrogen embrittlement resistance.

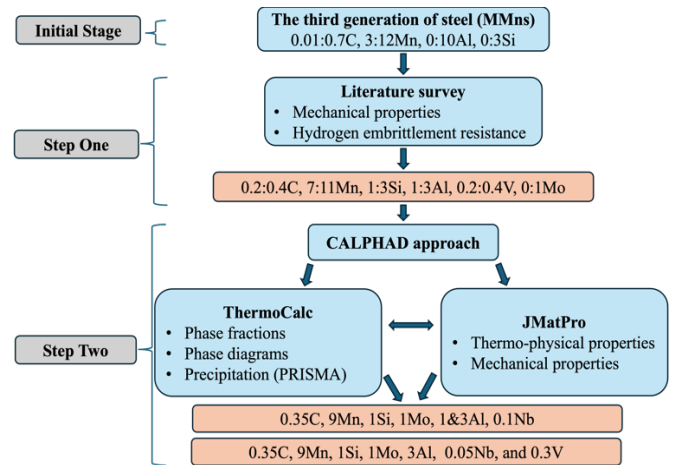


Fig. 13. Flow chart for computational designing approach for medium manganese steels with potential better hydrogen embrittlement resistance.

**Role of Alloying Elements:**

**Carbon (C):** A 0.35% C was found to provide a reliable strengthening effect and austenite stability.

**Manganese (Mn):** 9% Mn stabilizes the austenitic phase, enabling both TRIP and TWIP deformation mechanisms, which enhances ductility and strength.

**Silicon (Si) and Aluminum (Al):** Both elements help to stabilize the ferritic phase and control the deformation mechanism, with 1% Si and 1 and 3% Al being particularly effective.

**Niobium (Nb):** 0.1% Nb was optimal for forming fine NbC precipitates with high number density, which improve strength and serve as hydrogen deep trapping sites.

**Molybdenum (Mo):** 1% Mo was effective in refining the NbC precipitate size and increasing its volume fraction, further enhancing strength and hydrogen embrittlement resistance.

**Vanadium (V):** 0.3% V provides a balanced austenite stability and carbide size distribution, improving mechanical properties and hydrogen embrittlement resistance.

## ACKNOWLEDGMENTS

The authors would like to thank Jane ja Aatos Erkko Foundation and Tiina and Antti Herlin Foundation for their financial supports on Advanced Steels for Green Planet project. Mohammed Ali would like to acknowledge the financial support of Interreg Aurora, project Sustainable Hydrogen - Potential for Bothnia Gulf Cluster, project number 20357962.

## REFERENCES

- Ali, M., Alatarvas, T., and Kömi, J. (2024). Impact of niobium addition and non-metallic inclusions' characteristics on the microstructure and mechanical properties of low-carbon CrNiMnMoB ultrahigh-strength steel: A comprehensive investigation. *Journal of Materials Research and Technology*, 30, 6133–6153.
- Bhadeshia, H. K. D. H. (2016). Prevention of hydrogen embrittlement in steels. *ISIJ International*, 56(1), 24–36.
- Bhattacharya, A., Bokinala, P. K., Mitra, R., and Chakrabarti, D. (2019). Relative effect of C and Mn on strength-toughness of medium Mn steels. *Materials Science and Technology*, 35(1), 55–67.
- Chen, Q., Wu, K., Sterner, G., and Mason, P. (2014). Modeling precipitation kinetics during heat treatment with Calphad-based tools. *Journal of Materials Engineering and Performance*, 23(12), 4193–4196.
- Ennis, B. L., Jimenez-Melero, E., Atzema, E. H., Krugla, M., Azeem, M. A., Rowley, D., Daisenberger, D., Hanlon, D. N., and Lee, P. D. (2017). Metastable austenite driven work-hardening behaviour in a TRIP-assisted dual phase steel. *International Journal of Plasticity*, 88, 126–139.
- DiGiovanni, C., Ghatei Kalashami, A., Goodwin, F., Biro, E., and Zhou, N. Y. 2021. Occurrence of sub-critical heat affected zone liquid metal embrittlement in joining of advanced high strength steel. *Journal of Materials Processing Technology*, 288:116917.
- Geng, W. T., Freeman, A. J., and Olson, G. B. (2001). Influence of alloying additions on grain boundary cohesion of transition metals: First-principles determination and its phenomenological extension. *Physical Review B*, 63, 165415.
- Geng, W. T., Freeman, A. J., Wu, R., and Olson, G. B. (2000). Effect of Mo and Pd on the grain-boundary cohesion of Fe. *Physical Review B*, 62(10), 6208.
- Huang, C., Hu, C., Liu, Y., Liang, Z., and Huang, M. (2022). Recent developments and perspectives of advanced high-strength medium Mn steel: from material design to failure mechanisms. *Materials Futures*, 1(3), 032001.
- Kang, M., Woo, W., Lee, Y. K., and Seong, B. S. (2012). Neutron diffraction analysis of stacking fault energy in Fe–18Mn–2Al–0.6C twinning-induced plasticity steels. *Materials Letters*, 76, 93–95.
- Kozłowska, A., Janik, A., Radwański, K., and Grajcar, A. (2019). Microstructure evolution and mechanical stability of retained austenite in medium-Mn steel deformed at different temperatures. *Materials*, 12(18), 3042.
- Kwok, T. W. J., and Dye, D. (2023). A review of the processing, microstructure and property relationships in medium Mn steels. *International Materials Reviews*, 68(8), 1098–1134.
- Langer, J. S., and Schwartz, A. J. (1980). Kinetics of nucleation in near-critical fluids. *Physical Review A*, 21(3), 948.
- Lehnhoff, G. R., Findley, K. O., and De Cooman, B. C. (2014). The influence of silicon and aluminum alloying on the lattice parameter and stacking fault energy of austenitic steel. *Scripta Materialia*, 92, 19–22.
- Liu, P. Y., Zhang, B., Niu, R., Lu, S. L., Huang, C., Wang, M., Tian, F., Mao, Y., Li, T., Burr, P. A., Lu, H., Guo, A., Yen, H. W., Cairney, J. M., Chen, H., and Chen, Y. S. (2024). Engineering metal-carbide hydrogen traps in steels. *Nature Communications*, 15(1), 1–13.
- Luo, H. and Dong, H. (2016). CN106119493B.
- Luo, Y., Lu, H., Min, N., Li, W., and Jin, X. (2022). Effect of Mo and Nb on mechanical properties and hydrogen embrittlement of hot-rolled medium-Mn steels. *Materials Science and Engineering: A*, 844, 143108.
- Ma, Y. (2017). Medium-manganese steels processed by austenite-reverted-transformation annealing for automotive applications. *Materials Science and Technology (United Kingdom)*, 33(15), 1713–1727.

- Miller, R. L. (1972). Ultrafine-grained microstructures and mechanical properties of alloy steels. *Metallurgical Transactions*, 3(4), 905–912.
- Park, I. J., Jeong, K. H., Jung, J. G., Lee, C. S., and Lee, Y. K. (2012). The mechanism of enhanced resistance to the hydrogen delayed fracture in Al-added Fe–18Mn–0.6C twinning-induced plasticity steels. *International Journal of Hydrogen Energy*, 37(12), 9925–9932.
- Patra, A. K., Athreya, C. N., Mandal, S., Hari Kumar, K. C., and Subramanya Sarma, V. (2021). High strength-high ductility medium Mn steel obtained through CALPHAD based alloy design and thermomechanical processing. *Materials Science and Engineering: A*, 810, 140756.
- Ryu, J. H., Kim, S. K., Lee, C. S., Suh, D. W., and Bhadeshia, H. K. D. H. (2013). Effect of aluminium on hydrogen-induced fracture behaviour in austenitic Fe–Mn–C steel. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 469, 2149.
- Steineder, K., Daniel K., Schickinger, M., Thomas Hebesberger, T., and Schneider, R. (2019). Wärmebehandlungsoptionen von Medium-Mn-Stählen für den (Automobil) Leichtbau. *BHM Berg- und Hüttenmännische Monatshefte*, 164(9), 379–384.
- Suh, D. W., and Kim, S. J. (2017). Medium Mn transformation-induced plasticity steels: Recent progress and challenges. *Scripta Materialia*, 126, 63–67.
- Sun, B., Fazeli, F., Scott, C., Brodusch, N., Gauvin, R., and Yue, S. (2018). The influence of silicon additions on the deformation behavior of austenite-ferrite duplex medium manganese steels. *Acta Materialia*, 148, 249–262.
- Sun, B., Fazeli, F., Scott, C., and Yue, S. (2016). Phase transformation behavior of medium manganese steels with 3 Wt Pct aluminum and 3 Wt Pct silicon during intercritical annealing. *Metallurgical and Materials Transactions A: Physical Metallurgy and Materials Science*, 47(10), 4869–4882.
- Sun, B., Kwiatkowski da Silva, A., Wu, Y., Ma, Y., Chen, H., Scott, C., Ponge, D., and Raabe, D. (2023). Physical metallurgy of medium-Mn advanced high-strength steels. *International Materials Reviews*, 68(7), 786–824.
- Sun, B., Liu, G., He, J., Li, T., Wan, X., Dai, Z., Ponge, D., Raabe, D., Zhang, C., Godfrey, A., Miyamoto, G., Furuhashi, T., Yang, Z., van der Zwaag, S., and Chen, H. (2020). Chemical boundary engineering: A new route toward lean, ultrastrong yet ductile steels. *Science Advances*, 6(13), eaay1430.
- Sun, B., Palanisamy, D., Ponge, D., Gault, B., Fazeli, F., Scott, C., Yue, S., and Raabe, D. (2019). Revealing fracture mechanisms of medium manganese steels with and without delta-ferrite. *Acta Materialia*, 164, 683–696.
- Venegas, V., Caleyó, F., Baudin, T., Hallen, J. M., and Penelle, R. (2009). Role of microtexture in the interaction and coalescence of hydrogen-induced cracks. *Corrosion Science*, 51(5), 1140–1145.
- Wagner, R., Kampmann, R., and Voorhees, P. W. (2001). Homogeneous second-phase precipitation. *Phase Transformations in Materials*, 309–407.
- Wang, Z., Wan, Z., Chen, X., Xu, J., Zhou, Y., Huang, T., Kong, X., Ou, C., Zhang, J., and Li, J. (2022b). The H-induced fracture behavior in hot rolled medium Mn steels with and without  $\delta$ -ferrite. *Engineering Failure Analysis*, 141, 106651.
- Wang, Z., Zhu, X., and Liu, W. (2010). Influence of Mo on tempering precipitation in Nb–Mo–V microalloyed steels. *Chinese Journal of Materials Research*. 24(2) 217–222.
- Wei, F. G., and Tsuzaki, K. (2012). Hydrogen trapping phenomena in martensitic steels. In *Gaseous Hydrogen Embrittlement of Materials in Energy Technologies: The Problem, its Characterisation and Effects on Particular Alloy Classes (Vol. 2, pp. 493–525)*. Woodhead Publishing.
- Ye, Q., Yu, Y., Zhu, H., Li, J., Qiao, L., and Yan, Y. (2022). Formation of core-shell austenite as promoted by alloying solutes in hot-rolled medium Mn steel. *Journal of Materials Research and Technology*, 21, 1053–1060.
- Yoo, J., Jo, M. C., Jo, M. C., Kim, S., Kim, S. H., Oh, J., Sohn, S. S., and Lee, S. (2021). Effects of solid solution and grain-boundary segregation of Mo on hydrogen embrittlement in 32MnB5 hot-stamping steels. *Acta Materialia*, 207, 116661.
- Zhang, Y., Ye, Q., and Yan, Y. (2024). Processing, microstructure, mechanical properties, and hydrogen embrittlement of medium-Mn steels: A review. *Journal of Materials Science & Technology*, 201, 44–57.
- Zhao, C., Song, R., Zhang, L., Yang, F., and Kang, T. (2016). Effect of annealing temperature on the microstructure and tensile properties of Fe–10Mn–10Al–0.7C low-density steel. *Materials & Design*, 91, 348–360.
- Zou, Y., Ding, H., and Tang, Z. (2021). Effect of carbon content on deformation behavior and partitioning of manganese in medium-Mn steels. *Metals*, 11(4), 667.

## Physical Simulation of Heat-Affected Zones in a Weld Metal Used with 500 MPa Offshore Steel

H. Tervo\*. M. Gáspár\*\*. J. Kovács\*\*. A. Kaijalainen\*. V. Javaheri\*. J. Sainio\*\*\* and J. Kömi\*

\*Materials and Mechanical Engineering, Centre for Advanced Steel Research, University of Oulu, P.O. Box 4200, FI-90014, Oulu, Finland (e-mail: [henri.tervo@oulu.fi](mailto:henri.tervo@oulu.fi), [antti.kaijalainen@oulu.fi](mailto:antti.kaijalainen@oulu.fi), [vahid.javaheri@oulu.fi](mailto:vahid.javaheri@oulu.fi), [jukka.komi@oulu.fi](mailto:jukka.komi@oulu.fi))

\*\*Institute of Material Science and Technology, University of Miskolc, HU-3515, Miskolc-Egyetemváros, Hungary (e-mail: [marcell.gaspar@uni-miskolc.hu](mailto:marcell.gaspar@uni-miskolc.hu), [judit.kovacs@uni-miskolc.hu](mailto:judit.kovacs@uni-miskolc.hu))

\*\*\*SSAB Europe, Rautaruukintie 155, P.O. Box 93, FI-92101, Raahе, Finland (e-mail: [johannes.sainio@ssab.com](mailto:johannes.sainio@ssab.com))

**Abstract:** Offshore steels are engineered to have an outstanding combination of high strength and toughness to withstand extreme conditions needed in offshore and marine applications. Multiple welding passes may be needed when welding thick steel sections. In this case, microstructures of pre-existing passes are affected by the thermo-cycles caused by the subsequent passes. These heat-affected zones (HAZ) in the weld metal are less studied than the HAZs in the base metal next to the weld. HAZs after real welding process are relatively narrow and that way challenging to study and test reliably. Physical simulation provides an opportunity to produce different kinds of HAZs on sufficiently large area for various types of microstructural and mechanical properties characterization. Moreover, the effect of different welding methods and parameters can be easily studied by adjusting the simulation settings. Therefore, the aim of this study was to produce the coarse-grained (CGHAZ-W), intercritical (ICHAZ-W) and intercritically reheated heat-affected zones of the weld metal (ICCGHAZ-W) using physical simulation. Submerged arc welding (SAW) method was used to produce the original weld. HAZs with 2 different cooling times from 800 °C to 500 °C ( $t_{8/5} = 5$  and 30 s) were simulated utilizing Gleeble 3500 thermomechanical simulator. Microstructures were characterized using a Zeiss Sigma field emission scanning electron microscope. The results indicated that the original weld metal contained acicular ferrite nucleated on oxide inclusions, and thermal cycles induced microstructural changes in the weld metal, with each simulation variant resulted in distinctive features. Microstructures obtained by the physical simulation were supported by the numerical simulation results carried out by JMatPro software.

*Keywords:* simulation, steel, weld, microstructure, Gleeble, JMatPro, heat-affected zone

### 1. INTRODUCTION

High strength offshore steels are developed for demanding conditions where both high strength and toughness are required from the materials and structures manufactured of them. Some of the typical applications for such steels are offshore oil drilling platforms, wind power mills and ships body. For example, some parts of Valhall oil drilling platform in Norway were constructed using 500 MPa offshore steels (Willms, 2009).

As a result of thermomechanically controlled hot rolling process, the microstructure of 500 MPa offshore steel usually consists of fine-grained bainite and ferrite. However, the original microstructure is altered when steel is welded. The thermal cycles originating from the welding process cause different types of heat-affected zones (HAZ) depending on the peak temperature and cooling rate of each location in the steel. Additionally, welding of thick steel sections may require multiple passes causing additional thermal cycles on already altered microstructure as well as on the weld metal produced during previous passes.

The most important HAZs where the properties are expected to change are coarse-grained (CGHAZ), intercritical (ICHAZ) and intercritically reheated heat-affected zone (ICCGHAZ). Each of these zones is usually relatively narrow making it challenging to characterize their microstructures and mechanical properties reliably.

Physical simulation provides a way to produce microstructures imitating those of the different types of HAZs on sufficiently large area for mechanical tests and microstructural examination. It also makes it possible to study the effect of different welding methods and parameters by adjusting simulation parameters. Consequently, physical simulation is nowadays rather common way to study the HAZ microstructures and properties (Afkhami et al., 2022; Gáspár, 2019; Gáspár et al., 2019, 2015; Kovács et al., 2024; Laitinen et al., 2013; Mičian et al., 2020; Sisodia et al., 2019; Henri Tervo et al., 2020; H. Tervo et al., 2020; Tervo et al., 2021; Węglowski et al., 2013). However, there are less studies where the physical simulation has been applied on studying HAZs on weld metal caused by subsequent welding passes, referred as "reheated zones" there (Kang et al., 2018; Tezuka et al., 1995).



Therefore, the aim of this study was to simulate different types of HAZs on the weld metal for enabling further investigation of them. The microstructures obtained by the physical simulation were compared with the numerically simulated microstructures calculated by the JMatPro software.

## 2. MATERIALS AND METHODS

The studied base material is a 16 mm thick 500 MPa offshore steel that was welded by single-pass submerged arc welding (SAW) method. The filler material is ESAB OK 13.24, a Ni- and Mo-alloyed, Cu-coated wire for SAW. The flux used with the filler material is OK Flux 10.62. The chemical composition of the base material and the filler material are presented in Table 1. The following welding settings were applied: the root gap 3 mm, the edge width 4 mm, and the bevel angle 40°.

**Table 1. Chemical compositions of the studied base metal and filler metal.**

	C	Si	Mn	Cu	Cr	Mo	Ni
Base	≤ 0.14	≤ 0.6	≤ 1.7	≤ 0.55	Cr+Mo ≤ 0.65	≤ 2.00	≤ 2.00
Filler	0.07	0.18	1.3	0.06	0.05	0.2	0.78

70 × 10 × 10 mm<sup>3</sup> specimens were machined keeping the welded joint nearly in the middle with 1 mm offset from the weld centerline in order to eliminate the effect of segregation. These specimens were used for the HAZ-W simulations. Example of the specimen is shown in Fig. 1.

Thermomechanical simulator Gleeble 3500 was used to produce coarse-grained (CGHAZ-W), intercritical (ICHAZ-W), and intercritically reheated heat-affected subzone on the weld metal (ICCGHAZ-W). The time and temperature points of thermal cycles were determined based on the Rykalin-3D model (Rykalin et al, 1971) and the GSL programs were manually written. This 3D model characterizes the

temperature field generated by a moving point-like heat source on the surface of a semi-infinity body (Fig. 2). In this case 3D thermal conductivity is dominant while surface heat transfer is negligible. This model was selected considering the investigated 500 MPa steel is in the medium and heavy plate thickness (16 mm) where the Rykalin 3D model provides more precise result. Furthermore, this equation is independent from the plate thickness, therefore reduced number of variables needs to be considered. Each of the HAZ was simulated using two different cooling time from 800 °C to 500 °C ( $t_{8/5} = 5$  and 30 s) to represent the typical welding parameter variation. With regard to the selection of peak temperatures the motivation was to generate the most critical parts of the selected subzones, having the lowest toughness. Peak temperature of the CGHAZ-W simulations was 1350 °C, whereas in the ICHAZ-W simulations it was 815 °C defined by determining the  $A_{c1}$  temperature of the steel by using a dummy sample and adding 50 °C to it. The ICCGHAZ-W simulations were performed by combining the previously mentioned simulations. Simulation was based on Rykalin-3D model.

Microstructure of the simulated HAZ-Ws was characterized using field emission scanning electron microscope (FESEM, Zeiss Sigma). The acceleration voltage was 5 kV and the working distance varied approximately between 4 and 6 mm. The samples for microstructural characterization were cut from the Gleeble specimens keeping the simulation region in the middle. The samples were placed in specimen holders, grinded, polished and Nital-etched before the FESEM examination.

Hardness (HV10) was measured using a Reicherter UH250 universal macro-hardness tester according to the EN ISO 15614-1 standard.

Microstructures of the weld metal were also simulated numerically by calculating the continuous cooling transformation (CCT) and time-temperature transformation (TTT) diagrams using JMatPro (v12.2) software. The input for



Fig. 1. Specimen for the physical simulation of the heat-affected zones with the original weld metal in the middle.

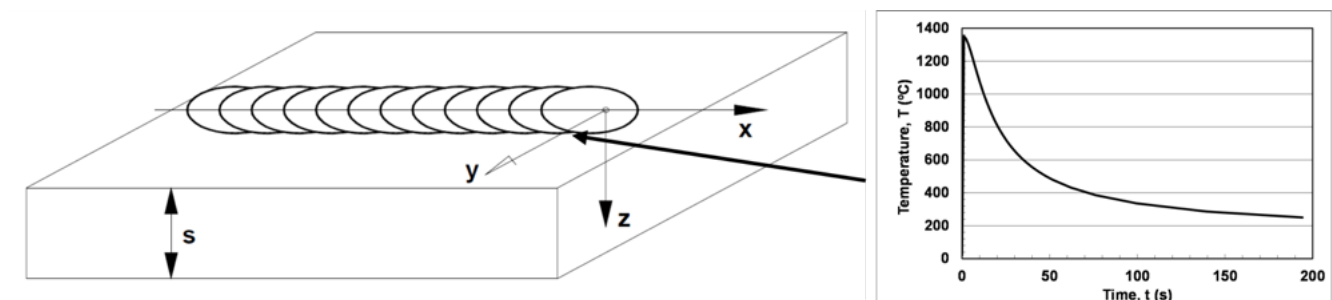


Fig. 2. CGHAZ-W thermal cycle based on the Rykalin-3D model.

the calculation is the fraction of each element present in the chemical composition of the steel in wt.%, estimated grain size and austenitization temperature. Outputs include the microstructure and hardness of the steel after various cooling rates.

### 3. RESULTS AND DISCUSSION

The simulated heat-affected zones on the weld metal as well as the original weld metal were examined by the FESEM to observe the changes in the microstructure due to the thermal cycles. The original single-pass weld metal consisted of acicular ferrite (AF). AF is a type of intergranular ferrite nucleating on certain type of non-metallic inclusions inside the prior austenite grains. AF consists of chaotically oriented needle-like grains with the width and length of approximately 1–3  $\mu\text{m}$  and 5–15  $\mu\text{m}$ , respectively. These grains form a complex interlocking structure, which efficiently prevents the fracture propagation, assisting to improve the toughness of the steel (Loder et al., 2016; Xiong et al., 2015). The original weld metal microstructure of the studied steel is presented in Fig. 3.

Intercritical heat-affected zone of the weld metal (ICHAZ-W) was simulated using the peak temperature 815  $^{\circ}\text{C}$ . The steel is partly austenitized in this temperature, so some changes in the microstructure are expected. The applied cooling times from 800  $^{\circ}\text{C}$  to 500  $^{\circ}\text{C}$  ( $t_{8/5}$ ) were 5 s and 30 s to simulate the practical heat input range of the arc welding processes. Figures 4(a) and (b) show the ICHAZ-W microstructures.

The peak temperature for the coarse-grained heat-affected zone of the weld metal (CGHAZ-W) simulation was 1350  $^{\circ}\text{C}$ . Therefore, the steel was fully austenitized. In this case, it is expected that the prior austenite grain growth occurs. The transformation microstructure depends on the cooling rate. Using high heat input, the cooling rate slows down (cooling time increases), and the resulting microstructure is less hardened than with low heat input welding. Therefore, a significant difference in the hardness values between the samples simulated with  $t_{8/5} = 5$  s and 30 s was observed. The CGHAZ-W microstructures are presented in Figs. 4(c) and (d).

Intercritically reheated coarse-grained heat-affected zone of the weld (ICCGHAZ-W) combines the thermal cycles of CGHAZ-W and ICHAZ-W, in this order. Firstly, the steel is

fully austenitized at 1350  $^{\circ}\text{C}$ , prior austenite grain size increases and the transformation microstructure becomes more or less hardened depending on the cooling rate. Secondly, the heat-affected microstructure experiences another thermal cycle, this time peaking at 815  $^{\circ}\text{C}$ , partly austenitizing the steel. As a result, coarsened prior austenite grains together with other microstructural changes are expected. The ICCGHAZ-W microstructures are shown in Figs. 4(e) and (f).

Weld metal microstructures were also simulated numerically by JMatPro software. Continuous cooling transformation (CCT) and time-temperature transformation (TTT) diagrams are presented in Figs. 5(a) and (b), respectively.

Based on the calculations from the CCT and TTT diagrams, it is evident that pearlite formation is highly unlikely, with ferrite and bainite being the predominant microstructures expected to form. Considering the cooling rates applied in this study, ferrite remains the primary microstructure. These results are in line with the obtained microstructures by the physical simulation.

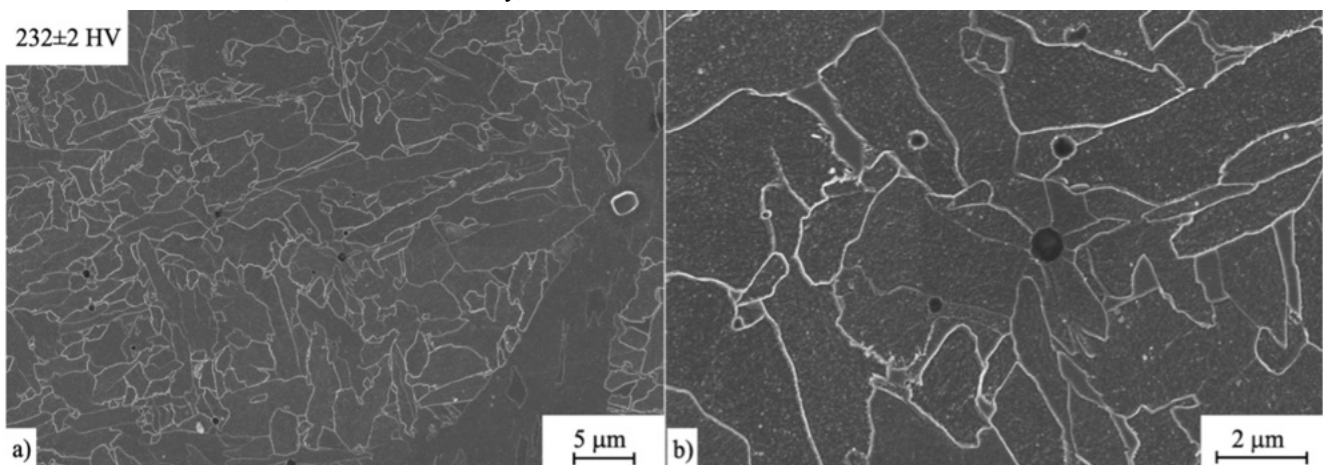


Fig. 3. FESEM images and the average hardness of the original weld metal taken with the magnification of 5000x (a) and 20000x (b). 95% interval for the mean hardness was determined using the student-t method.



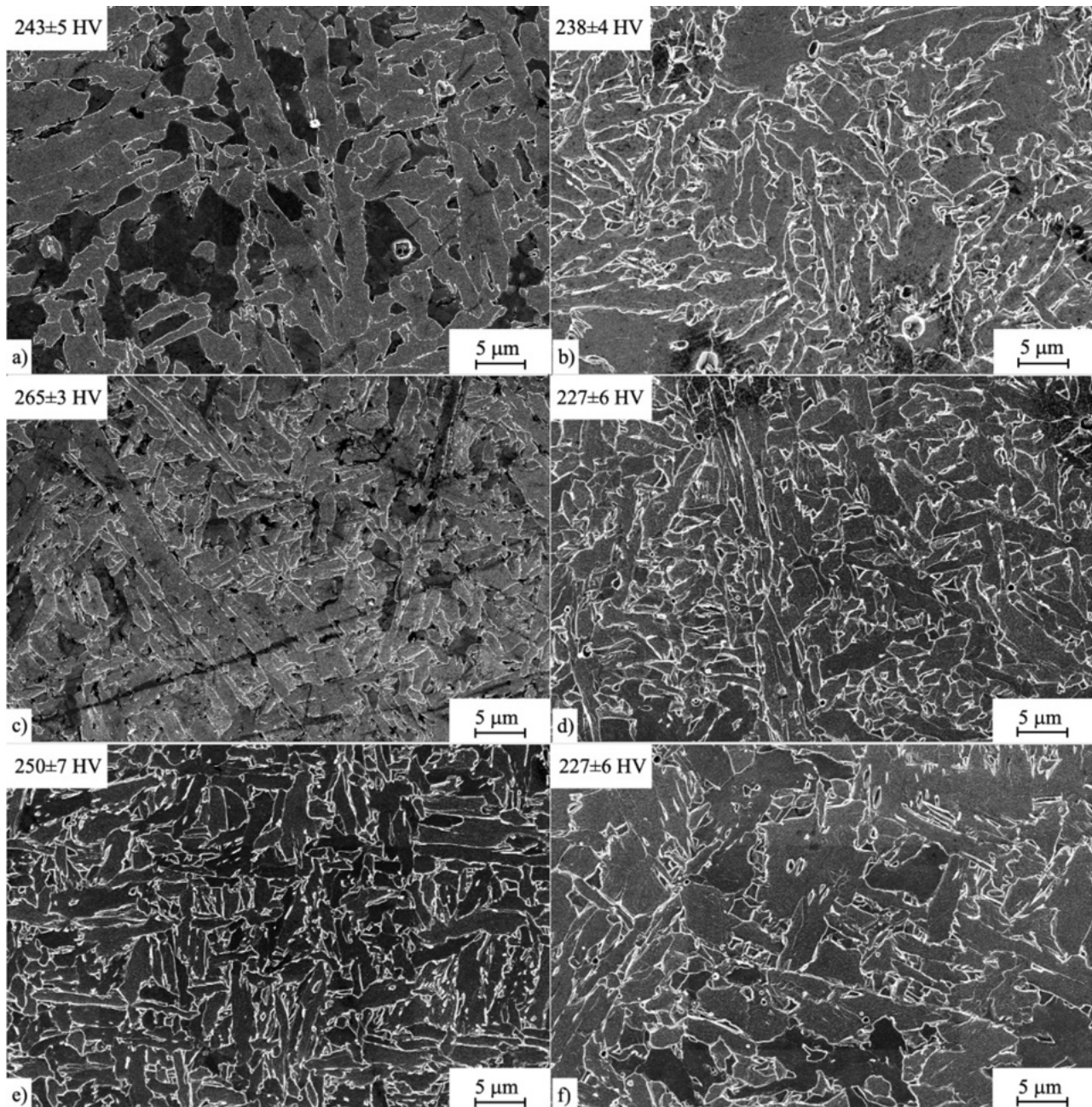


Fig. 4. FESEM images and the average hardness of the simulated ICHAZ-W (a) and (b), CGHAZ-W (c) and (d), ICCGAZ-W (e) and (f) using the  $t_{8/5} = 5$  s and 30 s, respectively, taken with the magnification of 5000x. 95% intervals for the mean hardness's were determined using the student-t method.

#### 4. CONCLUSIONS

Physical simulation was performed by Gleeble 3500 thermomechanical simulator to produce heat-affected zones on the pre-existing weld metal bonding two pieces of a 500 MPa offshore steel. The aim was to simulate the effect of multiple welding passes on the original single-pass weld microstructure. Following conclusions were made:

- Coarse-grained (CGHAZ-W), intercritical (ICHAZ-W), and intercritically reheated coarse-grained heat-affected zones (ICCGHAZ-W) were successfully simulated on the weld metal using two different

cooling times ( $t_{8/5} = 5$  and 30 s) to be able to study the microstructures in each variant.

- The original weld metal consisted of acicular ferrite nucleated on oxide inclusions.
- As a result of the thermal cycles of the physical simulation, changes in the weld metal microstructure were observed. CGHAZ-W, ICHAZ-W and ICCGAZ-W simulations all were seen to cause their characteristic features in the microstructure.
- Numerical simulation of the microstructures by JMatPro supported the results obtained by the physical simulation.

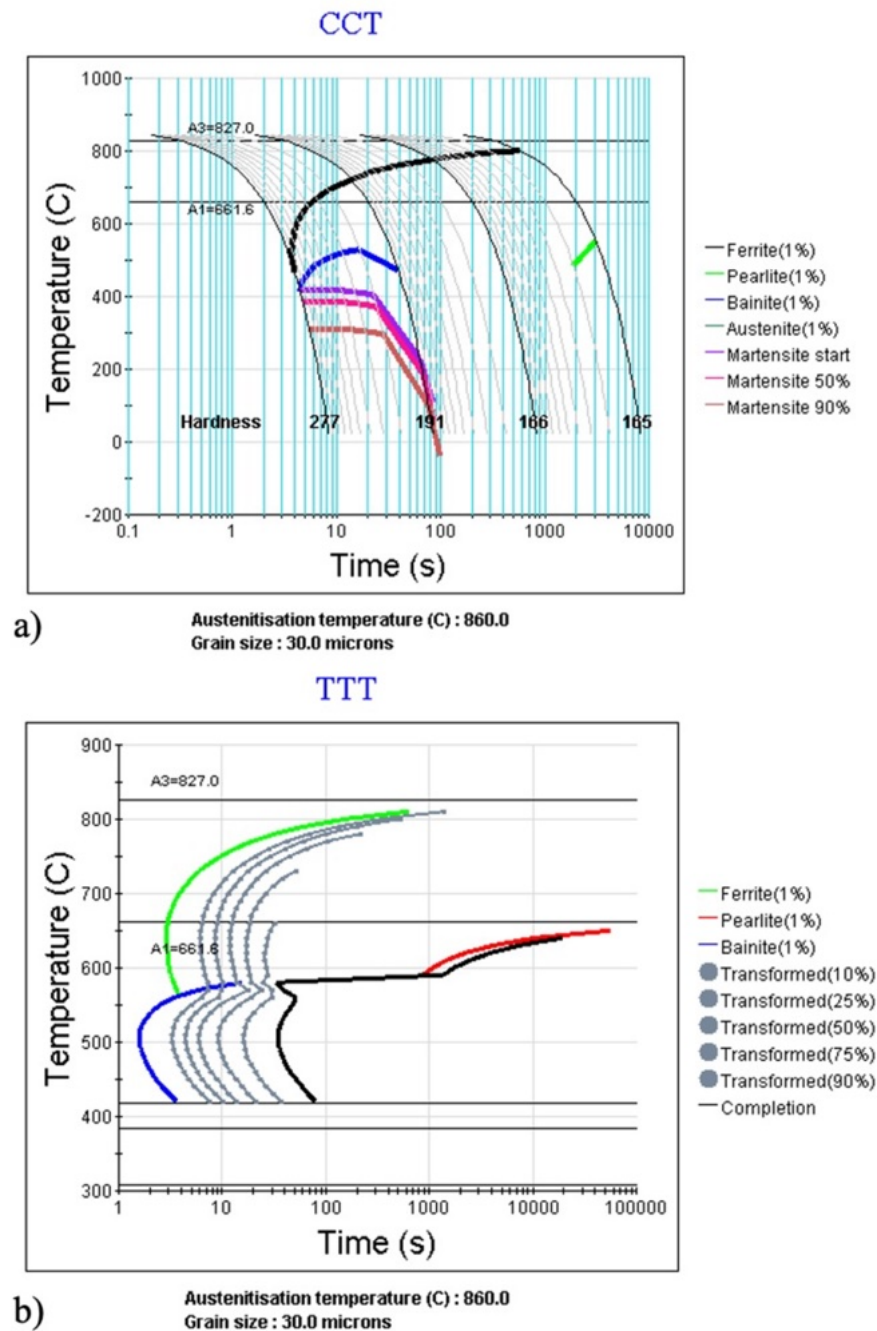


Fig. 5. Continuous cooling transformation (CCT) (a) and Time-temperature transformation (TTT) (b) diagrams of the weld metal calculated using JMatPro software.

#### ACKNOWLEDGEMENTS

The authors are grateful to the funding of the research program FOSSA II (Fossil-Free Steel Applications, Dnro 5562/31/2023) funded by Business Finland. Dr. Gáspár was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (Grant number: Bo/00643/22/6). Dr. Javaheri would also like to thank Jane ja Aatos Erkon säätiö (JAES) and Tiina ja Antti Herlinin säätiö (TAHS) for their financial supports on Advanced Steels for Green Planet project.

#### REFERENCES

- Afkhami, S., Javaheri, V., Amraei, M., Skriko, T., Piili, H., Zhao, X.-L. and Björk, T., (2022). Thermomechanical simulation of the heat-affected zones in welded ultra-high strength steels: Microstructure and mechanical properties. *Mater Des* 213, 110336. doi:10.1016/j.matdes.2021.110336
- Gáspár, M., (2019). Effect of welding heat input on simulated HAZ areas in S960QL High strength steel. *Metals* (Basel) 9, 1226. doi:10.3390/met9111226
- Gáspár, M., Balogh, A. and Sas, I., (2015). Physical simulation aided process optimisation aimed sufficient HAZ

- toughness for quenched and tempered AHSS, in: *IIW International Conference High-Strength Materials - Challenges and Applications. IIW, Helsinki, Finland.*
- Gáspár, M., Sisodia, R.P.S. and Dobosy, A., (2019). Physical simulation-based characterization of HAZ properties in steels. Part 2. Dual-phase steels. *Strength of Materials* 51, 805–815. doi:10.1007/s11223-019-00128-y
- Kang, Y., Park, G., Jeong, S. and Lee, C., (2018). Correlation between microstructure and low-temperature impact toughness of simulated reheated zones in the multi-pass weld metal of high-strength steel. *Metallurgical and Materials Transactions A* 49, 177–186. doi:10.1007/s11661-017-4384-3
- Kovács, J., Gáspár, M., Lukács, J., Tervo, H. and Kaijalainen, A., (2024). Comparative study about the results of HAZ physical simulations on different high-strength steel grades. *Welding in the World*. doi:10.1007/s40194-024-01714-8
- Laitinen, R.O., Porter, D.A., Karjalainen, L.P., Leiviskä, P. and Kömi, J., (2013). Physical simulation for evaluating heat-affected zone toughness of high and ultra-high strength steels. *Materials Science Forum* 762, 711–716. doi:10.4028/www.scientific.net/MSF.762.711
- Loder, D., Michelic, S.K. and Bernhard, C., (2016). Acicular ferrite formation and its influencing factors - a review. *Journal of Materials Science Research* 6, 24. doi:10.5539/jmsr.v6n1p24
- Mičian, M., Winczek, J., Harmaniak, D., Koňár, R., Gucwa, M. and Moravec, J., (2020). Physical simulation of individual heat-affected zones in S960MC steel. *Archives of Metallurgy and Materials* 66, 81–89. doi:10.24425/amm.2021.134762
- Sisodia, R., Gáspár, M. and Guellouh, N., (2019). HAZ characterization of automotive DP steels by physical simulation. *International Journal of Engineering and Management Sciences* 4, 478–487. doi:10.21791/IJEMS.2019.1.59.
- Tervo, H., Kaijalainen, A., Javaheri, V., Ali, M., Alatarvas, T., Mehtonen, M., Anttila, S. and Kömi, J., (2021). Comparison of impact toughness in simulated coarse-grained heat-affected zone of al-deoxidized and ti-deoxidized offshore steels. *Metals (Basel)* 11. doi:10.3390/met11111783
- Tervo, H., Kaijalainen, A., Javaheri, V., Kolli, S., Alatarvas, T., Anttila, S. and Kömi, J., (2020). Characterization of coarse-grained heat-affected zones in al and ti-deoxidized offshore steels. *Metals (Basel)* 10, 1–18. doi:10.3390/met10081096
- Tervo, H., Kaijalainen, A., Pallaspuro, S., Anttila, S., Mehtonen, S., Porter, D. and Kömi, J., (2020). Low-temperature toughness properties of 500 MPa offshore steels and their simulated coarse-grained heat-affected zones. *Materials Science and Engineering: A* 773. doi:10.1016/j.msea.2019.138719
- Rykalin, N.N. and Nikolaev, A.V., (1971). Welding arc heat flow. *Welding in the World* 9 3-4, 112-133.
- Tezuka, N., Shiga, C., Yamaguchi, T., Bosansky, J., Yasuda, K. and Kataoka, Y., (1995). Toughness degradation mechanism for reheated Mo-Ti-B bearing weld metal. *ISIJ International* 35, 1232–1238. doi:10.2355/isijinternational.35.1232
- Węglowski, M.S., Zeman, M. and Lomozik, M., (2013). Physical simulation of weldability of Weldox 1300 steel. *Materials Science Forum* 762, 551–555. doi:10.4028/www.scientific.net/MSF.762.551
- Willms, R., (2009). High strength steel for steel constructions. *Nordic Steel Construction Conference. Malmö, Sweden* 597–604.
- Xiong, Z., Liu, S., Wang, X., Shang, C., Li, X. and Misra, R.D.K., (2015). The contribution of intragranular acicular ferrite microstructural constituent on impact toughness and impeding crack initiation and propagation in the heat-affected zone (HAZ) of low-carbon steels. *Materials Science and Engineering: A* 636, 117–123. doi:10.1016/j.msea.2015.03.090



# Comparison of ML and ASM models for effluent nutrient estimation in the Hias Process<sup>\*</sup>

Tiina M. Komulainen<sup>\*</sup> Malik Baqeri<sup>\*</sup>  
Katrine Marsteng Jansen<sup>\*\*</sup> Arvind Keprate<sup>\*</sup>

<sup>\*</sup> *GrønnMet - Green Energy Lab, Department of Mechanical, Electrical and Chemical Engineering, Oslo Metropolitan University, Norway*  
(e-mail corresponding author: [tiina.komulainen@oslomet.no](mailto:tiina.komulainen@oslomet.no)).

<sup>\*\*</sup> *Hias IKS, Norway*

**Abstract:** The aim of this article is to develop and compare machine learning (ML) methods with activated sludge models (ASM) for estimation of effluent nutrients in the Hias Process. The Hias Process is a novel moving bed bioreactor with enhanced biological phosphorus removal and simultaneous nitrification and denitrification (MBBR-EBPR-SND). As the main energy cost of the nutrient removal process is aeration, it is necessary to design of energy-efficient control strategies that ensure compliance with legal requirements for nutrient removal in real-time while optimizing the aeration rates. The first step in control strategy design is development of models that represent the main process dynamics.

The case study data set of four months was collected from a 192 000 PE municipal MBBR process at Hias water resource recovery facility in Norway. The Hias Process consists of three anaerobic and seven aerobic zones, where biomass carriers flow continuously submersed in the used water and remove over 90 % of the phosphorus. The online measurements include used water flowrate, aeration rates, dissolved oxygen, suspended solids, and soluble nutrients  $PO_4$ , COD,  $NO_2$  and  $NO_3$ . Reduced ASM model, support vector regression (SVR) and long short-term memory neural network (LSTM), with and without dynamic time-delay, were developed to predict the effluent  $PO_4$  in the Hias process. The model prediction accuracies were compared using correlation coefficients and trend figures. The SVR model with fine gaussian kernel gave best results with strong R index of 0.9. The LSTM model reached a sufficient R index of 0.6 and the reduced ASM2d model a weak R index of 0.2. Including the dynamic time-delay improved the model accuracy. The machine learning models with dynamic time-delay will be developed further for energy-efficient control strategy development.

*Keywords:* water resource recovery facility; activated sludge model; support vector regression; long short-term memory network.

## 1. INTRODUCTION

As the water industry is responsible for approximately one percent of the total energy consumption in the European union, new legal requirements for energy neutrality are underway European Commission (2021). Therefore, development of energy-efficient control strategies is essential to minimize the energy consumption and to meet the strict nutrient recovery requirements at water resource recovery facilities (WRRF). The Hias Process is a novel, compact biological nutrient removal process that consists of a continuous-flow moving bed bioreactor with enhanced biological phosphorus removal and simultaneous nitrification and denitrification (MBBR-EBPR-SND) Rudi et al. (2019). As the main energy cost of the Hias Process is aeration, energy-efficient control strategies need to be developed to optimize the aeration rates and to ensure

compliance with legal requirements for nutrient removal in real-time.

Development of energy-efficient control strategies requires models that capture the main dynamic behaviour of the nutrient removal phenomena. First principles models such as the ASM2d model Henze et al. (1999) and its simplified version the reduced ASM2d model Nair et al. (2019) can be used as development and testing environment for control strategies. However, it is time-consuming and sometimes unfeasible to develop such models due to scarce instrumentation. Hence, machine learning models have gained high research interest in the water industry, for example for applications such as virtual/soft sensors Paepae et al. (2021).

In our previous work, data-driven models were developed to estimate the effluent nutrients in the Hias Process Neruo (2023), Komulainen et al. (2023). Virtual sensors were developed for estimation of  $PO_4$  and COD in the Hias Process using additional electrical conductivity (EC)

<sup>\*</sup> RFF Innlandet, Norway, is gratefully acknowledged for funding the PACBAL project (nr 337727).

measurements Komulainen et al. (2024). In this work we will use the virtual sensor estimating  $PO_4$  at inlet, develop reduced ASM2d models, and refine two best performing machine learning models from Baqeri (2024), Support Vector Regression (SVR) and Long-Short Term Memory (LSTM). In this study we answer the following research questions: Can the reduced ASM2d model, SVR model and LSTM model follow the dynamic trends of the effluent  $PO_4$  data? Which model gives the highest prediction accuracy?

## 2. MATERIALS AND METHODS

### 2.1 Software

Matlab software package version R2023a was used in the work. The simulation method was ode23s with automatic settings for the time step and error tolerance.

### 2.2 The Hias Process and instrumentation

The Hias Process is a biological nutrient removal process at a 192 000 PE municipal water resource recovery facility in Hamar, Norway. The Hias Process with instrumentation is illustrated in Fig. 1. The clarified used water (influent) and the recirculated biofilm carriers (from zone 10 via a conveyor belt) enter the anaerobic zone. The Hias Process consists of three anaerobic and seven aerobic zones, where biomass carriers flow continuously submerged in the used water and remove over 90 % of the phosphorus. The three anaerobic basins are mixed to ensure sufficient distribution of biofilm carriers in the water. Aeration in the following seven basins ensures sufficient dissolved oxygen concentrations for aerobic nutrient removal. Used water and submerged biofilm carriers float through the process with gravity.

The Hias Process instrumentation includes continuous online measurements of flowrate, temperature, aeration, dissolved oxygen, suspended solids, and nutrient compositions of  $PO_4$ , COD,  $NO_2$  and  $NO_3$ . Additional online measurements of electrical conductivity at inlet and in zone 3 were installed during the PACBAL research project 2022-2023. Soluble COD,  $NO_2$  and  $NO_3$  are measured continuously at inlet and zone 7. Suspended solids  $SS$  are measured at zone 10 and in the effluent after the disc filter. Effluent  $PO_4$  is measured using an online-analyzer with 10 minutes sampling time. The Hias Process online measurements utilized in this study are listed in Table 1. Hias laboratory assesses nutrient composition of  $PO_4$  and soluble COD at inlet, zones 3,4,7,10, and outlet, and  $NH_3$  at inlet and outlet from daily grab samples five days a week. Hence, there are 5 samples from laboratory and 1008 samples of online data for each variable per week.

### 2.3 Data collection and pre-processing

The Industrial IoT platform KYB, developed by Digitread Connect, was used for uploading and standardizing operational data from SCADA system of municipal Hias water resource recovery facility at Hamar, Norway. The online data set was collected in .csv format and the laboratory data set in .xlsx format.

Table 1. Online measurements.

Symbol	Description	Unit
$F$	Water flowrate inlet	$m^3/h$
$T$	Temperature inlet	$^{\circ}C$
$CODIN$	COD inlet	$g/m^3$
$NOIN$	$NO_2$ and $NO_3$ inlet	$g/m^3$
$ECIN$	El.conductivity inlet	$mS/cm$
$FO_i$	Aeration rate zones 4,5,6,7,8,9,10	$m^3/h$
$DO_i$	Dissolved oxygen zones 4,5,6,8,9	$m^3/h$
$PO$	$PO_4$ effluent	$g/m^3$

The inter-quartiles method was chosen for outlier removal. The outliers are identified as measurements more than 1.5 inter-quartile range above the upper quartile (75 percent) or below the lower quartile (25 percent). Missing and removed values were replaced with previous feasible values. Prior to ML modeling, the data is normalized to zero mean and standard deviation of one.

### 2.4 Modeling methods

The aim of this study is to develop and compare modeling methods that enable energy-efficient control strategy design for nutrient removal in the Hias Process. International Water Association has led work in developing Activated Sludge Models (ASM) that represent biological nutrient removal Henze et al. (1999). The ASM models were reduced and developed for a sequential MBBR pilot plant by Nair et al. (2019). In this study, these models were to be adapted for the continuous large-scale operation of the municipal Hias WRRF.

In the literature, Long-Short Term Memory (LSTM) neural network and Support Vector Regression (SVR) have gained lots of attention, and hence these two data-driven modeling methods are applied in this study. LSTM is a recurrent neural network suitable for modeling dependencies in and forecasting of sequential or time-series data. LSTM architecture includes memory cells and gates that regulate the flow of sequential data. These gates can learn which data in a sequence is important to keep or discard, enabling the network to maintain a longer context of information as described in Hochreiter and Schmidhuber (1997).

SVR is frequently used to predict relationship between continuous input and output variables. SVR minimizes error between the model prediction and the data by fitting a hyperplane in a high-dimensional space of the input variables and the output variables. The kernel trick converts the dataset to higher dimensions by combining the features using for example linear, quadratic, cubic, or Gaussian functions as described in Cortes and Vapnik (1995).

### 2.5 Model comparison

The Hias Process effluent  $PO_4$  measurements were used as the output variable for the models. Both  $R^2$  index and correlation coefficient R were used to compare the modeling accuracy of the different methods. The model prediction of the real data points is weak for index between 0-0.3, moderate for index between 0.3-0.5, sufficient for index between 0.5-0.7, and strong for index between 0.7-1.

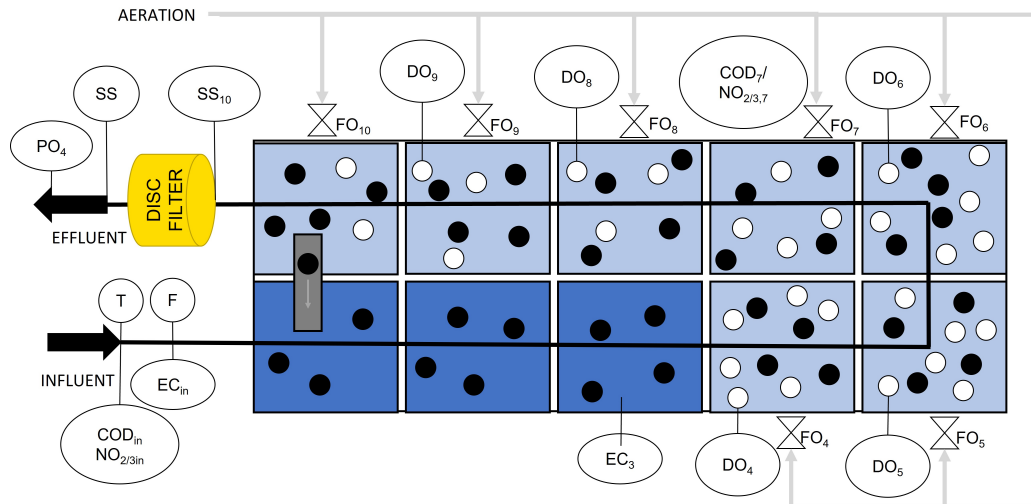


Fig. 1. The Hias Process with instrumentation.

### 3. RESULTS

#### 3.1 Data selection and pre-processing

The online data and laboratory data were collected for a period of 21.3.-31.7.2023. During period of 1.4.-15.5.2023 many online measurements were missing, hence, the online data set is for period 15.5.2023-31.7.2023. Period of 21.3.-31.7.2023 is used for the laboratory data set.

The outliers in the data set were using inter-quartiles method. The missing values were filled in with previous feasible values. Prior to the machine learning model development, the online data was normalized.

#### 3.2 Dynamic time delay

The time delay through the ten zones of the Hias Process has a mean of 6 hours (36 samples), a standard deviation of 1 hour (6 samples), a minimum of 4.8 hours (29 samples) and a maximum of 15 hours (90 samples). Hence, in this study the effect of dynamic time delay on model accuracy was analysed. For the reduced ASM2d models, the varying time delay  $t_d$  is calculated for a lumped volume  $V$  combining three real process zones ( $3 \cdot 215m^3$ ) and continuous measurement of used water flowrate  $F(t)$  in  $[m^3/h]$  according to Equation 1:

$$t_d(t) = V/F(t) = 645m^3/F(t) \quad (1)$$

As the sharp variations and small oscillations in the time delay can cause numerical challenges in simulation, the time-delay was smoothed with a moving-mean approach. Time windows of 6 samples (1h), 12 samples (2h) and 18 samples (3h) were plotted against the calculated time-delay. Varying time-delay with moving mean of 12 samples (2h) was chosen for this study as it removes the fast oscillations that are present in moving mean of 6 samples, but follows the main trends more closely than moving mean of 18 samples.

#### 3.3 Virtual measurements of inlet $PO_4$ and $NH_3$

As the online data set does not include online measurements of inlet  $PO_4$  and  $NH_3$ , these were estimated from

the laboratory data and online measurements of electrical conductivity and COD. In Komulainen et al. (2024) virtual sensors were developed for estimation of  $PO_4$  at the Hias Process inlet using additional electrical conductivity (EC) measurement. Based on 32 unique laboratory data points, a linear regression was fitted between  $EC_{IN}$ ,  $PO_{IN}$  and  $COD_{IN}$ , given in Equation 2. The parameter values were  $c_1 = 0.3488$  and  $c_2 = 0.6138$  with strong modeling accuracy of  $R^2 = 0.86$ .

$$EC_{IN} = c_1 PO_{IN} + c_2 COD_{IN} \quad (2)$$

In this work we developed a simple estimation of  $NH$  at the inlet. Based on 28 unique laboratory data points, a linear regression was fitted between  $NH_{IN}$  and  $COD_{IN}$ , given in Equation 3. The parameter values were  $c_3 = 0.1211$  with strong modeling accuracy of  $R^2 = 0.97$ .

$$NH_{IN} = c_3 COD_{IN} \quad (3)$$

#### 3.4 Reduced ASM2d model development

The activated sludge models (ASM) describe the dynamic changes in the nutrient concentrations and dissolved oxygen in the process zones. The simplified ASM2d models developed for a pilot scale batch-MBBR process by Nair et al. (2019) were further modified to fit the available measurements in the continuous large-scale municipal WRRF process that is the subject of this study. Significant simplifications are necessary to match the model variables with the available online measurements in the Hias process. The simplified models included nutrient uptake and release by the microbes in the biomass carriers. The following assumptions were applied:

- Phosphate  $PO_4$ , soluble organic substrate  $COD$ , ammonia  $NH_3$ , sum of nitrate  $NO_2$  and nitrite  $NO_3$ , and dissolved oxygen  $DO_2$  are the components in the simplified models.
- Ready biodegradable substrate ( $S_F$ ) and volatile fatty acids ( $S_A$ ) presented in Nair et al. (2019) are lumped together as soluble substrate ( $S_S$ ). In this study, interpreted as the measured variable, soluble  $COD$ .

- Particulate biodegradable components ( $X_S$ ) are omitted due to missing online measurement of influent total suspended solids.
- The biomass variables of stored poly-phosphate ( $X_{PP}$ ) and stored organic compounds COD ( $X_{PHA}$ ) are omitted to simplify the equations.
- Temperature effect is neglected, as it varies very little between days. It increases slowly from 8.8 °C to 13.6 °C during the four months period.
- Plug flow is assumed for water movement between zones. The time delay is dynamically time-delayed  $t_d$ .
- To match the model variables and real measurements, the process volume is divided to three parts: one lumped anaerobic volume, where V1 includes zones 1-3, two lumped aerobic volumes where V2 includes zones 4-6 and V3 includes zones 7-9 as illustrated in Figure 2. Zone 10, representing 10% of total process volume is omitted to simplify the calculations.

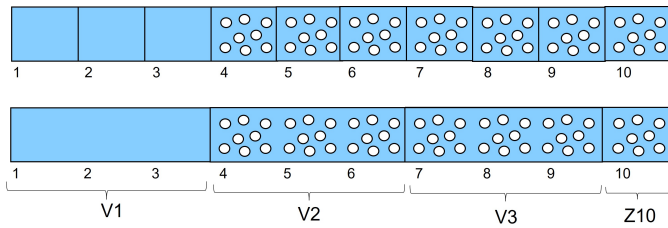


Fig. 2. Above: The Hias Process with 10 zones. Below: Modeling approach with 3 lumped zones.

**Anaerobic volume V1** In anaerobic zones  $i$ , the biomass consumes soluble  $COD$  and the component balance follows Equation 4.

$$\frac{dCOD_i(t)}{dt} = \frac{F(t)}{V} (COD_{i-1}(t - t_d) - COD_i(t)) - r_3 \frac{COD_i(t)}{K_A + COD_i(t)} \quad (4)$$

Simultaneously, the biomass releases  $PO$  to water with stoichiometric relation  $Y_{PO}$  to  $COD$  uptake following Equation 5.

$$\frac{dPO_i(t)}{dt} = \frac{F(t)}{V} (PO_{i-1}(t - t_d) - PO_i(t)) + Y_{PO} r_3 \frac{COD_i(t)}{K_A + COD_i(t)} \quad (5)$$

It is assumed that concentrations of ammonia  $dNH(t)/dt = 0$ , nitrate/nitrite  $dNO(t)/dt = 0$  and dissolved oxygen  $dDO(t)/dt = 0$  remain unchanged through the anaerobic zone.

**Aerobic volumes V2 and V3** In aerobic zones  $j$ , the biomass takes up soluble  $COD$  and consume oxygen  $DO$ , the component balance follows Equation 6.

$$\frac{dCOD_j(t)}{dt} = \frac{F(t)}{V} (COD_{j-1}(t - 3t_d) - COD_j(t)) - r_1 \frac{COD_j(t)}{K_S + COD_j(t)} \frac{DO_j(t)}{K_O + DO_j(t)} \quad (6)$$

In aerobic zones, the biomass takes up phosphate  $PO$  and consume oxygen  $DO$ , the component balance follows Equation 7.

$$\frac{dPO_j(t)}{dt} = \frac{F(t)}{V} (PO_{j-1}(t - 3t_d) - PO_j(t)) - r_4 \frac{PO_j(t)}{K_{PS} + PO_j(t)} \frac{DO_j(t)}{K_O + DO_j(t)} \quad (7)$$

In aerobic zones, the biomass convert ammonia  $NH$  to nitrite and nitrate  $NO$ , and consume oxygen  $DO$ . The component balance follows Equation 8.

$$\frac{dNH_j(t)}{dt} = \frac{F(t)}{V} (NH_{j-1}(t - 3t_d) - NH_j(t)) - r_5 \frac{NH_j(t)}{K_{NH} + NH_j(t)} \frac{DO_j(t)}{K_{OAOB} + DO_j(t)} \quad (8)$$

Simultaneously in deeper layers of biofilm, biomass converts nitrite and nitrate  $NO$  into nitrogen gas. The component balance follows Equation 9.

$$\frac{dNO_j(t)}{dt} = \frac{F(t)}{V} (NO_{j-1}(t - 3t_d) - NO_j(t)) + r_5 \frac{NH_j(t)}{K_{NH} + NH_j(t)} \frac{DO_j(t)}{K_{OAOB} + DO_j(t)} - r_6 \frac{NO_j(t)}{K_{NO} + NO_j(t)} \frac{K_O(t)}{K_O + DO_j(t)} \quad (9)$$

In aerobic zones, the dissolved oxygen  $DO$  component balance consists of mass transfer in and out of the zone, mass transfer from aeration  $FO$ , the biomass consuming oxygen for nutrient uptake. The oxygen component balance follows Equation 10.

$$\frac{dDO_j(t)}{dt} = \frac{F(t)}{V} (DO_{j-1}(t - 3t_d) - DO_j(t)) + K_L \frac{FO(t)}{V} (DO_{max}^*(t) - DO_j(t)) - r_1 \frac{COD_j(t)}{K_S + COD_j(t)} \frac{DO_j(t)}{K_O + DO_j(t)} - Y_{PA} r_4 \frac{PO_j(t)}{K_{PS} + PO_j(t)} \frac{DO_j(t)}{K_O + DO_j(t)} - Y_{NH} r_5 \frac{NH_j(t)}{K_{NH} + NH_j(t)} \frac{DO_j(t)}{K_{OAOB} + DO_j(t)} \quad (10)$$

The saturation coefficients, stoichiometric constants and rate constants from Nair et al. (2019) were used in this work, presented in Table 2. The initial conditions for simulation models are calculated for aggregated laboratory and online data set, presented in Table 3. The models were implemented in Matlab and Simulink. The parameters from Nair et al. (2019) did not give a reasonable fit to Hias Process data, hence, the reaction rates  $r$  and stoichiometric constant  $Y_{PO}$  were optimized further using the Hias data set. The parameters were fitted to the initial steady state data 3 by setting the ordinary differential

equations 4 - 10 to zero. The tuned reaction rates  $r$  and constant  $Y_{PO}$  together with correlation coefficient  $R$  for model fitness are presented in Table 4.

Table 2. Reduced ASM2d model parameter values for saturation coefficients  $K$ , rate constants  $r$ , and stoichiometric constants  $Y$ .

Symbol	Description	Value	Unit
$K_A$	COD anaerobic	2.20	$gCOD/m^3$
$K_S$	COD aerobic	0.11	$gCOD/m^3$
$K_{PS}$	PO aerobic	0.2	$gP/m^3$
$K_O$	DO aerobic	2.96	$gO_2/m^3$
$K_{OAOB}$	DO aerobic nitrifiers	1.57	$gO_2/m^3$
$K_{NH}$	NH aerobic	1	$gN/m^3$
$K_{NO}$	NO aerobic	1.02	$gN/m^3$
$K_L$	Aeration	38	$gN/m^3$
$r_3$	COD anaerobic	112	$gCOD/(m^3h)$
$r_1$	COD aerobic	20.8	$gCOD/(m^3h)$
$r_4$	PO aerobic	171	$gP/(m^3h)$
$r_5$	NH aerobic	17.9	$gN/(m^3h)$
$r_6$	NOX aerobic	12.9	$gN/(m^3h)$
$Y_{PO}$	$P/COD$ anaerobic P release	0.577	$gP/gCOD$
$Y_{PA}$	$DO/P$ aerobic P storage	1.496	$gO_2/gP$
$Y_{NH}$	$DO/NH$ aerobic nitrification	4.32	$gO_2/gNH$

Table 3. Initial conditions for nutrients, flowrates and temperature dated 12.05.2023 at 12:00. Values marked with \* are assumed values.

Nutrient	IN	V1	V2	V3	unit
$PO$	5.56	36.4	3.83	0.14	$gP/m^3$
$COD$	450	202	72	67.5	$gCOD/m^3$
$NH$	53	53*	49*	45	$gN/m^3$
$NOX$	3.63	3.63*	0.698	0.5*	$gN/m^3$
$DO$	0*	0*	8.75	5.92	$gO_2/m^3$
$F$	417.6				$m^3/h$
$T$	8.3				$^{\circ}C$
$FOV2$			9571		$m^3/h$
$FOV3$				2897	$m^3/h$

Table 4. Tuning of of the reduced ASM2d model parameters and resulting correlation coefficient  $R$ .

Param.	Nair	Tuned
$K_L$	38	5.80 & 0.18
$r_3$	112	162
$r_1$	20.8	113 & 4.34
$r_4$	171	20.8 & 6.09
$r_5$	17.9	3.11 & 0.31
$r_6$	12.9	72.7 & 48.4
$Y_{PO}$	0.577	0.125
$R$	-0.0079	0.0735

### 3.5 Machine learning model results

The input variables were first dynamically time-delayed using Simulink block "variable time delay". The variables at inlet, including flow ( $F$ ), temperature ( $T$ ), sum of nitrates and nitrites ( $NOXIN$ ), soluble organic matter

( $CODIN$ ) and estimated phosphorus ( $POIN$ ) in were delayed using dynamic time delay  $Td3$ . The total aeration rate in the lumped zone V2 ( $FOV2$ ) was delayed using dynamic time delay  $Td2$ , and the total aeration rate in lumped zone V3 ( $FOV3$ ) was delayed using dynamic time delay  $Td1$ . The dynamically time-delayed input variables were collected from Simulink. Both input variables and the output variables were normalized and the data set was divided into two, 50% training and 50% testing.

**Support Vector Regression** Different regression models were fitted to the dynamically time-delayed training data in Matlab Regression Learner toolbox. The models were compared using the dynamically time-delayed test data set. The results are presented in Table 5 with  $R^2$  model accuracy index obtained from the toolbox. The best results were achieved with support vector regression and gaussian process regression with  $R^2$  values up to 0.9 indicating excellent model fitness. In comparison, linear regression and support vector regression with linear kernel function results in  $R^2$  values around 0.4 indicating insufficient model fitness. The models without time delay were also developed, but these resulted in poorer model accuracy as given in Table 5. On average the model accuracy reduction was 5% for the best performing models, SVR with fine gaussian kernel function and gaussian process. The support vector regression model with fine gaussian kernel function was chosen further for comparison with other models.

Table 5. Regression model type,  $R^2$  model accuracy for train and test data sets with and without (w/o) dynamic time-delay.

Time delay Model	with train	with test	w/o train	w/o test
SVR fine gaussian	0.87	0.86	0.80	0.83
SVR medium gaussian	0.62	0.59	0.58	0.56
SVR cubic	0.54	0.50	0.52	0.48
SVR quadratic	0.50	0.47	0.46	0.43
SVR linear	0.39	0.35	0.40	0.38
Linear regression	0.42	0.38	0.42	0.39
Gaussian process quadratic	0.92	0.92	0.86	0.87

**Long-short term memory network** The LSTM neural network was implemented in Matlab. The network architecture included: a sequence input layer with 7 features (inputs), 3 LSTM layers with different number of nodes, fully connected layer with 1 response (output) and a regression layer. The initial plant options of the model were chosen as follows: optimizer-Adam, MaxEpochs-40, Mini Batch Size of 1008, Sequence Length of 144, Gradient Threshold of 1, Initial Learn Rate of 0.001, Learn Rate Schedule piecewise, Learn Rate Drop Factor of 0.001, Learn Rate Drop Period of 10. The LSTM model architecture was tested with 3 to 5 LSTM layers with 7-14-28-14-7 nodes. Increasing the depth of the network from 3 LSTM layers to 4 and to 5 LSTM layers decreased the model fit. Increasing the depth of an LSTM network, such as here, from 3 to 5 layers, can lead to issues like overfitting, where the model learns noise and details specific to the training data but fails to generalize. Thus explaining, why the output prediction of



the 5 layer LSTM gave a limited response in the trend figure.

Further, the architecture with 3 LSTM layers was used for testing of the number of nodes. Increasing the number of nodes from 7-14-7 to 14-28-14 and further to 28-56-28 improved the model fitness both numerically and visually in the trend figure. However, increasing the size to 56-112-56 and 112-224-112 decreased the modeling accuracy. This is because as the model complexity increases, each additional unit of complexity (in this case, more nodes) contributes less to capturing useful, generalizable patterns, and more to fitting random fluctuations in the data. Batch and sequence size were chosen using the daily and weekly patterns of the combined domestic and industrial used water flowrate and nutrient composition. The measurement interval was 6 samples per hour. The batch size was chosen to represent the weekly pattern including 1008 samples and the sequence size was chosen to represent the daily pattern including 144 samples. The batch size was changed from 1008 down to 144 and sequence size was changed from 144 down to 72 which did not improve the results. Decreasing the sequence size further to 144-36 showed limiting the predicted output values in the trend figure. Increasing the learning rate from 0.001 to 0.1 decreased the model accuracy and predicted values in the trend figures were limited over 0. Decreasing the learning rate from 0.001 to 0.0001 decreased the model accuracy. Role of dynamically time-delayed input variables is significant. The model accuracy decreases from 0.71 to 0.49 for training data set and from 0.61 to 0.46 for test data set when input variables are not delayed. This can indicate that the LSTM model with the current architecture does not manage to compensate for the time delays.

Table 6. LSTM parameters: Nodes per layer, Mini Batch Size MBS, Sequence Length SL, Initial Learning Rate ILR, correlation coefficients for train and test  $R_{train}$  and  $R_{test}$ . B bad fit to trend data.

Nodes	MBS	SL	ILR	$R_{train}$	$R_{test}$
7-14-7	1008	144	0.001	0.61	0.52
7-14-28-14-7	1008	144	0.001	0.57B	0.48B
7-14-14-7	1008	144	0.001	0.61	0.51
14-28-14	1008	144	0.001	0.65	0.56
56-112-56	1008	144	0.001	0.66	0.48
112-224-112	1008	144	0.001	0.61	0.49
28-56-28	1008	144	0.001	0.71	0.61
28-56-28	1008	144	0.01	0.62 B	0.43 B
28-56-28	1008	144	0.0001	0.61	0.58
28-56-28	144	72	0.001	0.67	0.49B
28-56-28	144	36	0.001	0.68	0.47B
28-56-28 w/o Td	1008	144	0.001	0.49	0.46

### 3.6 Summary of results

The models are compared visually for training and testing data sets in Figs. 3 and 4, and using the correlation coefficient between the real data and the model predictions in Table 7. The SVR model with highest

number of parameters and very little modeling efforts (1 hour) provided excellent modeling accuracy. The LSTM model development and tuning required 10 hours of tuning and provided satisfactory modeling accuracy. The reduced ASM2d model development took over 100 hours of work and resulted in inadequate modeling accuracy. While the SVR model predictions very accurately follow the dynamic trends in the data, the LSTM model gives conservative predictions for a limited output range. SVR often provides more accurate predictions than LSTM when working with small datasets due to its ability to find a hyperplane in a high-dimensional space that best fits the data, minimizing overfitting. LSTMs, which are good in capturing long-range dependencies within large datasets, may struggle with overfitting and underperformance in scenarios with limited data due to their complex architectures. Thus, SVR is typically more suitable and reliable for small-scale data modeling where the primary goal is generalization over capturing sequential patterns.

Likewise, the reduced ASM2d model suffers from large oscillations related to the flowrate term in the mass-balance equation of the nutrient removal. When time-delayed data was used, both ML models improved prediction accuracy, the LSTM for training data 22% and testing data 15%, the SVR for training data 2% and testing data 0%. The ML model prediction accuracy increased when time-delayed data was used.

Table 7. Time consumption and correlation coefficient R for different models.

Model type	Development time [h]	R Train	R Test
SVR fine gaussian	1	0.94	0.95
SVR f.g. w/o Td	1	0.92	0.95
LSTM	10	0.71	0.61
LSTM w/o Td	10	0.49	0.46
rASM2d	100+	0.07	0.07

## 4. CONCLUSIONS AND FURTHER WORK

Answers to the research questions conclude the results of this study: Can the reduced ASM2d model, SVR model and LSTM model follow the dynamic trends of the effluent  $PO_4$  data? The support vector regression with time delayed variables was very accurate in matching the dynamic trends in the data. The LSTM model had a sufficient fit. The time-delayed variables increased the ML model accuracy. The reduced ASM2d models require more tuning and development to be able to match the dynamic trends in the data. Which model gave the highest prediction accuracy? The SVR model with fine gaussian kernel function and dynamic time-delay gave the best modeling results.

Further work is suggested on control strategy design. The next step is comparison of the unit step responses of the reduced ASM2d model and ML models. If the dynamic responses are similar, the SVR model with highest model accuracy should be used further. If the ML model step responses are not similar to the reduced ASM2d model, a method of online-adaptation of the reduced ASM2d model parameters is suggested to make the model more accurate and suitable for control studies.

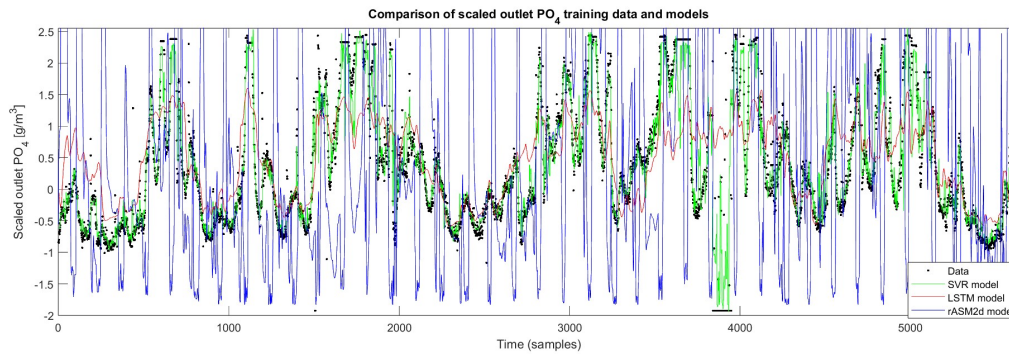


Fig. 3. Training data set. Outlet  $PO_4$  data in black dots, and model predictions with SVR in green, LSTM in red and rASM2d in blue. Time in samples.

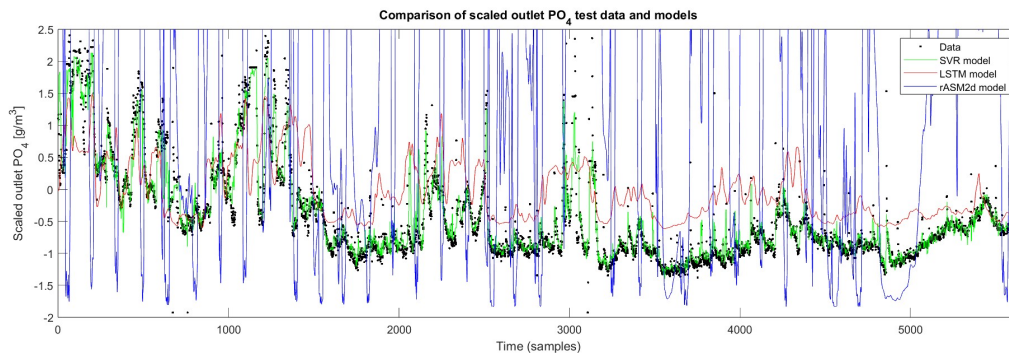


Fig. 4. Testing data set. Outlet  $PO_4$  data in black dots, and model predictions with SVR in green, LSTM in red and rASM2d in blue. Time in samples.

## 5. CONTRIBUTIONS

Gjermund Sørensen, Katrine Marsteng Jansen (Hias IKS WRRF) and Torgeir Saltnes (Hias How2O) provided valuable insights in the Hias Process. Tobias Korten and Katrine Marsteng Jansen from (Hias IKS), and Axel Tveiten Bech and Ola Solli Grønningsæter (Digitread Connect) collected the data. Professor Tiina Komulainen developed the reduced ASM2d models, input-output design, final development and testing of the ML models, and prepared the article draft. Master student Malik Baqeri developed the data pre-processing methods and initial development of the LSTM and SVR models. Katrine Marsteng Jansen provided facilitated data collection and provided feedback on the research work and manuscript. Professor Arvind Keprate supervised development of the ML models.

## REFERENCES

- Baqeri, M. (2024). *Machine Learning and Control for Phosphorous Removal*. Master thesis, Oslo Metropolitan University. URL <https://hdl.handle.net/11250/3163016>.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273–297. doi:10.1007/BF00994018.
- European Commission (2021). Eu strategy on energy system integration. URL [https://energy.ec.europa.eu/topics/energy-system-integration/eu-strategy-energy-system-integration\\_en](https://energy.ec.europa.eu/topics/energy-system-integration/eu-strategy-energy-system-integration_en).
- Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M.C., v.R. Marais, G., and Loosdrecht, M.C.V. (1999). Activated sludge model no.2d, asm2d. *Water Science Technology*, 39(1), 165–182.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Komulainen, T.M., Baqeri, A.M., Nermo, E., Keprate, A., Saltnes, T., Jansen, K.M., and Korostynska, O. (2023). Estimation of effluent nutrients in municipal mbbr process. In *Proceedings of The 64th International Conference of Scandinavian Simulation Society*. Linköping University Electronic Press, Västerås, Sverige 26-27.9.2023. doi:10.3384/ecp200037.
- Komulainen, T.M., Baqeri, M., Jansen, K.M., Saltnes, T., Bech, A.T., and Korostynska, O. (2024). Virtual sensors for the hias process. *Water Practice and Technology*, wpt2024176. doi:10.2166/wpt.2024.176.
- Nair, A.M., Fanta, A., Haugen, F.A., and Ratnaweera, H. (2019). Implementing and extended kalman filter for estimating nutrient composition in a sequential batch mbbr pilot plant. *Water Science Technology*, 80(2), 317–328. doi:10.2166/wst.2019.272.
- Nermo, E. (2023). *MPC of nutrient removal in wastewater treatment process*. Oslo Metropolitan University. URL <https://hdl.handle.net/11250/3101078>.
- Paepae, T., Bokoro, P.N., and Kyamakya, K. (2021). From fully physical to virtual sensing for water quality assessment: A comprehensive review of the relevant state-of-the-art. *Sensors*, 21(21). doi:10.3390/s21216971.
- Rudi, K., Goa, I.A., Saltnes, T., Sørensen, G., Angell, I.L., and Eikås, S. (2019). Microbial ecological processes in mbbr biofilms for biological phosphorus removal from wastewater. *Water Science Technology*, 79(8), 1467–1473. doi:10.2166/wst.2019.149.

## Anaerobic digestion of biosolid pyrolysis liquid and hydrolyzed sludge - simulation with extended ADM1 model

Thea Lucia Indrebø\*, Gudny Øyre Flatabø, Wenche Hennie Bergland, Gamunu Samarakoon Arachchige

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Kjølnes ring 56, 3918 Porsgrunn, Norway

\*Corresponding author (e-mail: [Thea.L.Indrebo@usn.no](mailto:Thea.L.Indrebo@usn.no))

**Abstract:** Pyrolysis of biosolids aims to reduce solid volumes and improve energy recovery; however, the pyrolysis liquid (PL) is a by-product that has no good direct application. One idea is to link pyrolysis and anaerobic digestion (AD), in which PL can be valorized for methane production. PL contains various compounds that potentially threaten the stability of AD. This study, therefore, aims to extend the current Anaerobic Digestion Model No.1 (ADM1) and evaluate the influence of phenol, furfural, 5-hydroxymethylfurfural (5-HMF), styrene, and ammonia from PL on AD. Two lab-scale AD reactors were simulated and compared with experimental data: one fed with hydrolyzed sludge and the other fed with an additional stream of PL. The simulation accurately predicts hydrolyzed sludge as substrate, while the simulation of the reactor co-digesting hydrolyzed sludge and PL overestimates methane production. Ammonia, phenol, and styrene were identified as the most significant inhibitors. However, based on the overestimation of methane production, it is clear that the PL has more inhibitors present than those implemented in the model. Simulations further showed that an additional stream of PL increased methane production by 4.3%, even with significant inhibition by the compounds.

**Keywords:** *Pyrolysis Liquid, Anaerobic Digestion, ADM1, Inhibition, Phenol, Ammonia, Styrene, Hydrolyzed Sludge, Biosolids*

### 1. INTRODUCTION

There is a need for waste management systems which address the world's growing population's energy demand and treat the enormous amount of waste produced in both an efficient and sustainable way (Tayibi et al., 2021). In recent years, coupling pyrolysis and anaerobic digestion (AD) as a waste management system has gained attention due to its possibility for energy recovery and economic value (Feng and Lin, 2017). Pyrolysis is a thermochemical process where organic matter is heated in the absence of oxygen to yield biochar and pyrolysis gas, where pyrolysis liquid (PL) is a by-product. Meanwhile, AD is the biochemical process in which organics biologically degrade to biogas, a mixture of methane and carbon dioxide. Hydrolysis is a rate-limiting step in the AD process, where organic matter is solubilized. Pretreatment methods such as the thermal hydrolysis process (THP) improve hydrolysis with the additional benefits of pathogens removal and enhanced digestate dewaterability (Han et al., 2017).

Pyrolysis coupled with AD offers numerous synergies such as increased resource use by feeding biosolid-PL (PL from dried digestate), biomethanation of pyrolysis gas, and biochar for inhibition control and increased methane production (Tayibi et al., 2021). PL is a by-product of pyrolysis with no direct application and is a complex mixture of more than 400 organics and inorganics (Giwa et al., 2019). Valorization of PL might be possible with AD, but the compounds in PL, such as phenols, furans derivatives, styrene, and ammonia, can inhibit and pose a threat to AD stability (Seyedi et al., 2019). There

have been a few attempts to add PL to AD with an increase in methane production at low PL loadings (Hubner and Mumme, 2015; Seyedi et al., 2020).

Anaerobic Digestion Model nr.1 (ADM1) is a valuable and cost-effective simulation model for predicting an AD system's robustness and efficiency (Batstone et al., 2002). A simulation allows one to anticipate challenges such as inhibition before they arise in the system and predict the outcomes and implications of substrates and substrate combinations. Some previous simulation studies have focused on adding PL to AD. Raya et al. (2021) presented a simulation focusing on how phenol, furfural, and 5-hydroxymethylfurfural (5-HMF) from aqueous-PL from softwood affected AD. Flatabø and Bergland (2022) simulated a full-scale reactor co-digesting sludge from THP, biosolids-PL, and pyrolysis gas. Some studies have focused on the simulation of furfural inhibition from steam explosion pulping wastewater (Li et al., 2023) and phenol simulation from olive mill waste in AD (Fezzani and Ben Cheikh, 2009).

Flatabø and Bergland (2022) only simulated ammonia toxicity, one of many inhibitors in the PL, but neglected the influence of other compounds. This study aims to extend the ADM1 model to predict inhibition from multiple compounds in the PL and evaluate the effects of PL in AD. The objective of this study is to (i) establish an extended model with inhibiting compounds in PL co-digested with sludge from THP and compare the model data to experimental data, (ii) investigate

which compounds contribute to inhibition, and (iii) evaluate the effect of biosolid-PL in AD.

## 2. METHODOLOGY

### 2.1 Anaerobic Digestion System

Substrate consisting of sewage sludge and food waste (60:40 v/v) was collected after treatment in a THP at 155°C with a retention time of 20 minutes. The substrate is termed hydrolyzed sludge (HS) and is the substrate for two lab-scale mesophilic (38°C) semi-continuously stirred tank reactors (semi-CSTRs). Inoculum was sampled from an industrial mesophilic CSTR fed by the same substrate. One AD bioreactor was used as a control and fed HS termed "HS-reactor", while the other was fed HS and PL termed "HSPL-reactor". The PL was produced using the Biogreen® technology by VOW ASA, where dry pelletized biosolids from the industrial CSTR digestate were pyrolyzed at 600°C. More details about the pyrolysis process and PL sampling can be found in Flatabø et al. (2023).

### 2.2 Analytical Characteristics

Total COD (tCOD), soluble COD (sCOD), pH, alkalinity, total ammonia nitrogen (TAN), and volatile fatty acids (VFAs) were analyzed as previously described by Bergland et al. (2015). TAN, pH, and VFAs were analyzed in addition to elemental analysis for the PL; check Flatabø and Bergland (2022) for details. The concentration of phenols in the PL was determined with Supelco Phenol-Test Art.100856 after the sample was filtered (0.45 µm).

### 2.3 Base Case Scenario and Data Collection

The simulations of the lab-scale reactors were implemented in ADM1 in aquasim version 2.1. Both reactors were fed with the same amount of HS, but one was fed with an additional stream of PL. Digestate measurements and characterization were done 1-2 times a week. Both reactors started with only HS as substrate, while the start of the simulation on day 0 was when the HSPL-reactor received PL. More details on the experimental part of the setup can be found in Flatabø et al. (2024). The HS reactor was simulated for 232 days (stopped early due to technical issues). The HSPL-reactor was simulated for 437 days, which was the entire period during which the reactor was fed PL.

### 2.4 Hydrolyzed Sludge (HS) Composition

Average experimental data and literature data were used to simulate both reactors. Both reactors experienced variations in inflow and concentration, and experimental effluent data was used to evaluate the accuracy of the model. The HSPL-reactor had a slightly higher inflow due to the addition of PL, resulting in a lower hydraulic retention time (HRT) than the HS-reactor.

The COD concentration of HS was, on average, 101.1 kg tCOD/m<sup>3</sup> and 21.6 kg sCOD/m<sup>3</sup> at a TS of 5.5-10.8%; a more detailed composition is given in Tab. 1. VFAs were additional inputs and were based on experimental data. Sugars denotes the concentration of n-caproic acid, isocaproic acid and heptanoic acid. Amounts of soluble inert were estimated from effluent data (36.6% of sCOD), while the total inert was

estimated on lab data (COD reduction) and data from the industrial plant (average yield), which on average estimates that 25% of tCOD is inert. For the HSPL-reactor, the total inerts were adjusted to 30% of tCOD after 248 days because of changes in the substrate. Protein composition was set to 9 % of tCOD and lipids to 28% of tCOD based on data from Flatabø and Bergland (2022). Carbohydrates were used to achieve the summed total concentration of COD. The inorganic carbon was based on experimental data. To account for the high pH in the lab experiment, the cations were adjusted in the model by adding 0.15 M in the HS-reactor and 0.2 M in the HSPL-reactor to reach the targeted pH.

**Table 1: Average input data for the hydrolyzed sludge (HS) with a tCOD of 101.1 kg COD/m<sup>3</sup> and sCOD of 21.6 kg COD/m<sup>3</sup>. Lipids, carbohydrates, and proteins were based on data from Flatabø and Bergland (2022), while the other input data were estimated on experimental data.**

Input	Unit	Value
Lipids	kg COD/m <sup>3</sup>	28.6
Carbohydrates	kg COD/m <sup>3</sup>	16.6
Protein	kg COD/m <sup>3</sup>	9.1
Sugars	kg COD/m <sup>3</sup>	0.44
Acetic acid	kg COD/m <sup>3</sup>	1.35
Propionic acid	kg COD/m <sup>3</sup>	0.43
Butyric acid	kg COD/m <sup>3</sup>	0.85
Valeric acid	kg COD/m <sup>3</sup>	1.1
Inert soluble	kg COD/m <sup>3</sup>	7.9
Inert particulate	kg COD/m <sup>3</sup>	26.8
TAN	kmol N/m <sup>3</sup>	0.09
Inorganic carbon	kmol HCO <sub>3</sub> <sup>-</sup> /m <sup>3</sup>	0.03

### 2.4 Pyrolysis Liquid (PL) Composition

The PL composition was obtained from experimental data and literature (see Tab. 2). tCOD and sCOD add up to 355 kg/m<sup>3</sup> and 164 kg/m<sup>3</sup>, respectively. 60% of tCOD is totally inert, in accordance with what Flatabø and Bergland (2022) estimated. The concentration of phenols, 5-HMF, furfural, and styrene are assumed to be soluble in the liquid phase and account for a part of the inert concentration. The 5-HMF and furfural concentrations were found in Hubner and Mumme (2015). Meanwhile, the styrene concentration (0.06 wt.%) is found in Seyedi et al. (2019).

### 2.5 Parameters in the modified ADM1

The hydrolysis constant,  $k_{hyd}$ , was determined to be 1 d<sup>-1</sup> for carbohydrates, lipids, and proteins for HS, which is in accordance with Flatabø and Bergland (2022) and Souza et al., 2013b). For the PL, the hydrolysis constant was determined to be 0.3 d<sup>-1</sup>, as described by Flatabø and Bergland (2022). The disintegration constant was kept at 0.5 d<sup>-1</sup> in

accordance with Montecchio et al. (2017) and Flatabø and Bergland (2022).

**Table 2: Input data for biosolid pyrolysis liquid obtained at a process temperature of 600 °C. The concentrations are 355 kg tCOD/m<sup>3</sup> and 164 kg sCOD/m<sup>3</sup>.**

Input	Unit	Value
Lipids	kg COD/m <sup>3</sup>	37.4 <sup>a</sup>
Carbohydrates	kg COD/m <sup>3</sup>	9.53 <sup>a</sup>
Protein	kg COD/m <sup>3</sup>	7.66 <sup>a</sup>
Sugars	kg COD/m <sup>3</sup>	18.1 <sup>a</sup>
Acetic acid	kg COD/m <sup>3</sup>	40.6 <sup>a</sup>
Propionic acid	kg COD/m <sup>3</sup>	9.45 <sup>a</sup>
Butyric acid	kg COD/m <sup>3</sup>	10.3 <sup>a</sup>
Valeric acid	kg COD/m <sup>3</sup>	8.23 <sup>a</sup>
Inert soluble	kg COD/m <sup>3</sup>	64.8 <sup>a</sup>
Inert particulate	kg COD/m <sup>3</sup>	136 <sup>a</sup>
Phenols	kg COD/m <sup>3</sup>	10.2 <sup>b</sup>
HMF	kg COD/m <sup>3</sup>	0.003 <sup>c</sup>
Furfural	kg COD/m <sup>3</sup>	0.37 <sup>5</sup>
Styrene	kg COD/m <sup>3</sup>	1.84 <sup>d</sup>
TAN	kmol N/m <sup>3</sup>	1.09 <sup>a</sup>
Inorganic carbon	kmol HCO <sub>3</sub> <sup>-</sup> /m <sup>3</sup>	0.313 <sup>b</sup>

Phenol, furfural, 5-HMF, and styrene were extended in the ADM1 to describe the inhibition of PL in AD by using a non-competitive form of inhibition control. Kinetic growth parameters of the phenol in AD were collected from Fezzani and Ben Cheikh (2009), where it is assumed that phenols degrade to hydrogen and benzoate. Inhibition data of phenol was taken from Raya et al. (2021). Phenol and benzoate are included in the charge balance equation to determine pH, as Fezzani and Ben Cheikh (2009) described. Kinetic parameters and inhibition data for furfural degradation are taken from Raya et al. (2021) and Li et al. (2023), where the latter made an extended ADM1 model which took account of the intermediate product furoic acid. Furoic acid is less inhibitory than its precursor, and therefore, inhibition from furoic acid was neglected. Data for growth kinetics and inhibition of 5-HMF in AD was collected from Raya et al. (2021) and B. Liu et al. (2017). For styrene, the inhibition constant,  $K_i$ , on anaerobic mix culture was 145 mg/L (Araya et al., 2000). Anaerobes release 4 units acetic acid and 4 units of hydrogen

<sup>a</sup> Flatabø and Bergland (2022)

<sup>b</sup> Experimental data

for each unit of styrene; thus, the yield of each compound is calculated by (1) and (2).

$$f_{ac,styr} = (1 - Y_{styr}) \cdot \frac{thCOD_{ac} \cdot 4}{thCOD_{ac} \cdot 4 + thCOD_{h2} \cdot 4} \quad (1)$$

$$f_{h2,styr} = (1 - Y_{styr}) \cdot \frac{thCOD_{h2} \cdot 4}{thCOD_{ac} \cdot 4 + thCOD_{h2} \cdot 4} \quad (2)$$

Growth kinetics of styrene degradation under anaerobic conditions is scarce; however, some microbes have been identified to degrade styrene, such as *Pseudomonas sp. E-934846*, which can survive in both anaerobic and aerobic environments (Arnold et al., 1997). Growth kinetics of *Pseudomonas sp. E-934846* on styrene degradation in Gąszczak et al. (2012), where  $k_m$  and  $K_S$  were measured to be 3.96 d<sup>-1</sup> and 0.018 kg COD<sub>s</sub>/m<sup>3</sup>.  $Y$  and  $K_{dec}$  are assumed to be 0.01 and 0.02.

Tab. 3 summarizes the kinetics and values of phenol, benzoate, styrene, 5-HMF, furfural, and furoic acid degradation. For ammonia, default data from ADM1 are used, where  $K_i$  is 0.018 M.

### 3. RESULTS AND DISCUSSION

#### 3.1 Hydrolyzed Sludge Reactor

The experimental and simulated methane productions are shown in Fig. 1.

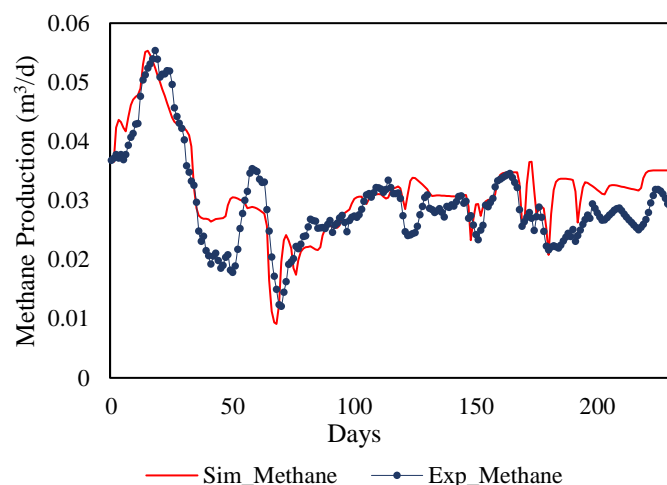


Fig. 1. Experimental (blue dots) and simulated (red line) methane production of reactor only treating hydrolyzed sludge (HS).

Experimental methane production is based on an average COD reduction of 1 week. The model follows experimental values well but overestimates methane production from day 37 to 56 due to changes in inflow in the experiment. After those days, the model predicts methane production well until day 181. The reason for this overestimation after day 181 is unclear. However, according to experimental data, the methane yield

<sup>c</sup> Hubner and Mumme (2015)

<sup>d</sup> Seyedi et al. (2019)



showed a decreasing trend, which is due to the substrate variations that were not accounted for in the model.

The methane partial pressure in the biogas is 63.3% based on simulated values, the same as the previously reported data from Flatabø and Bergland (2022). The pH was slightly lower in the simulation (7.6) than in the experimental results, an average of 7.73. Moreover, the inhibition and inorganic nitrogen concentration are shown in Fig. 2. Simulated TAN and free ammonia (NH<sub>3</sub>) are well predicted compared to experimental data; besides the first 40 days, the simulated values are slightly overestimated. Overestimations of TAN can be due to variations in the substrate composition or the fact that the model is not robust enough to compute accurate data for the first days (Souza et al., 2013b). The main contributor to inhibition is NH<sub>3</sub>, which is 0.3 on average on a scale where 0 is full inhibition and 1 is no inhibition. In the initial days, NH<sub>3</sub> was higher in experimental data compared to

simulated data. Thus, the simulated methane production should probably be more inhibited by NH<sub>3</sub> than what the simulation shows.

### 3.2 Hydrolyzed Sludge and Pyrolysis Liquid Reactor

Simulated and experimental data (based on an average COD reduction for one week) for methane production and PL loading are plotted in Fig. 3. Methane production was underestimated from day 24 to 34 and was overestimated in the next 30 days when the reactor was fed with a reduced flow rate due to operational issues. The simulated values predict the experimental values much better after 266 days, where the PL loading is less than 0.01 L/d (0.14 kg tCOD/m<sup>3</sup>/d) and with long HRTs of 48.6 days compared to the 16.2 and 32.4 days in other periods.

**Table 3: Summary of kinetic parameters of phenol, benzoate, styrene, 5-HMF, furfural, and furoic acid degradation. COD<sub>s</sub> and COD<sub>x</sub> denoted substrate COD and biomass COD, respectively. f indicates the fraction of the compound that converts to another compound.**

Parameter	Description	Unit	Phenol	Benzoate	Furfural	Furoic acid	5-HMF	Styrene	Value
C	Carbon content	kmole/kg COD	0.0268 <sup>a</sup>	0.0324 <sup>a</sup>	0.0284 <sup>a</sup>	0.0347 <sup>a</sup>	0.0312 <sup>a</sup>	0.025 <sup>a</sup>	
k <sub>m</sub>	Maximum uptake rate	d <sup>-1</sup>	15 <sup>b</sup>	8 <sup>b</sup>	20.53 <sup>d</sup>	3.71 <sup>d</sup>	10 <sup>e</sup>	3.96 <sup>h</sup>	
K <sub>s</sub>	Half saturation constant for uptake	kg COD <sub>s</sub> /m <sup>3</sup>	30 <sup>b</sup>	15.5 <sup>b</sup>	9.59 <sup>d</sup>	18.24 <sup>d</sup>	10 <sup>e</sup>	0.018 <sup>h</sup>	
Y	Yield of biomass uptake	kg COD <sub>x</sub> /kg COD <sub>s</sub>	0.01 <sup>b</sup>	0.013 <sup>b</sup>	0.08 <sup>g</sup>	0.08 <sup>a</sup>	0.1 <sup>e</sup>	0.01 <sup>a</sup>	
K <sub>i</sub>	Inhibition on methanogens		1.12 <sup>c</sup>		2.47 <sup>d</sup>		2.05 <sup>c</sup>	0.45 <sup>f</sup>	
K <sub>dec</sub>	Biomass decay rate		0.02 <sup>b</sup>		0.02 <sup>c,g</sup>	0.02 <sup>a</sup>	0.01 <sup>e</sup>	0.02 <sup>a</sup>	
K <sub>i_bnz_h2</sub>	Inhibition on benzoate degraders			9.5·10 <sup>-5</sup> <sup>c</sup>					
f <sub_bnz_phe< sub=""></sub_bnz_phe<>									0.87 <sup>b</sup>
f <sub>h2_phe</sub>									0.13 <sup>b</sup>
f <sub>ac_bnz</sub>									0.51 <sup>b</sup>
f <sub>h2_bnz</sub>									0.49 <sup>b</sup>
f <sub>ac_fua</sub>									0.82 <sup>d</sup>
f <sub>h2_fua</sub>									0.10 <sup>d</sup>
f <sub>ac_HMF</sub>									0.88 <sup>e</sup>
f <sub>h2_HMF</sub>									0.12 <sup>e</sup>

<sup>a</sup> Calculated/Estimated

<sup>c</sup> Raya et al. (2021)

<sup>e</sup> B. Liu et al. (2017)

<sup>g</sup> Brune et al. (1983)

<sup>b</sup> Fezzani and Ben Cheikh (2009)

<sup>d</sup> Li et al. (2023)

<sup>f</sup> Araya et al. (2000)

<sup>h</sup> Gąszczak et al. (2012)

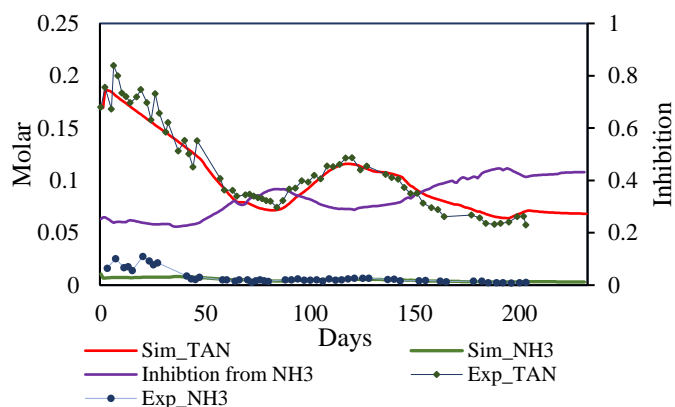


Fig. 2. Simulated total ammonia nitrogen (TAN) (red line) and free ammonia,  $\text{NH}_3$ , (green line) are plotted against experimental data (dots) for the HS-reactor. The inhibition of  $\text{NH}_3$  is also plotted, (purple line) where 0 = full inhibition and 1 = no inhibition.

Moreover, the model had difficulty when there were sharp changes in inflow. Simulated methane concentration is, on average, 67.3% in the last 100 days, close to the experimental methane concentration of 67.8% analyzed in that period.

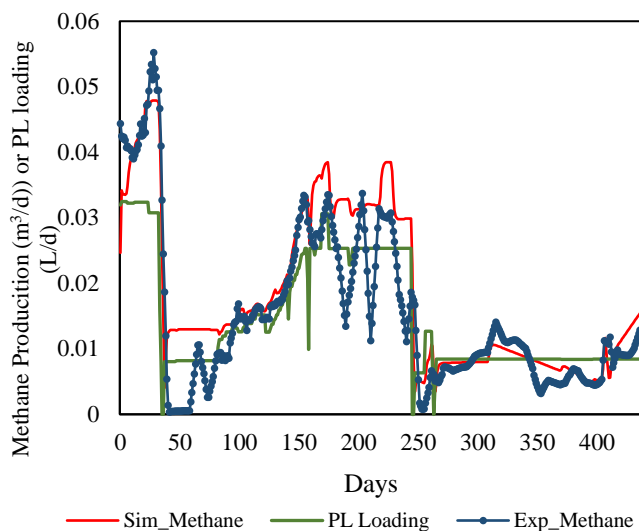


Fig. 3. Simulated (red line) and experimental methane production (black dots) of reactor treating hydrolyzed sludge (HS) and biosolid pyrolysis liquid (PL). The PL loading (green line) is added to visualize substrate load changes during the experiment.

Inorganic nitrogen and inhibitions of the HSPL-reactor are plotted in Fig. 4, where the main contributor to inhibition is  $\text{NH}_3$ . TAN is underestimated in the model until day 65; however, it is comparable to the experimental value until day 350, when the reactor effluent had an increase in TAN concentration. This effect might be due to substrate variations, such as more proteins entering the reactor or accumulation inside the reactor. Inhibition from  $\text{NH}_3$  was 0.14, 0.85 for phenol, and 0.93 for styrene, while furfural and HMF had no significant contribution (0 – 0.1%). Also, in this case,  $\text{NH}_3$  is underestimated in the first days, so inhibition was lower during that period.  $\text{NH}_3$  is correlated to pH, where ammonium and  $\text{NH}_3$  are in equilibrium, and an increase in pH shifts the equilibrium toward  $\text{NH}_3$ . pH was simulated to be 7.85, which is slightly higher than the average experimental pH (7.7).

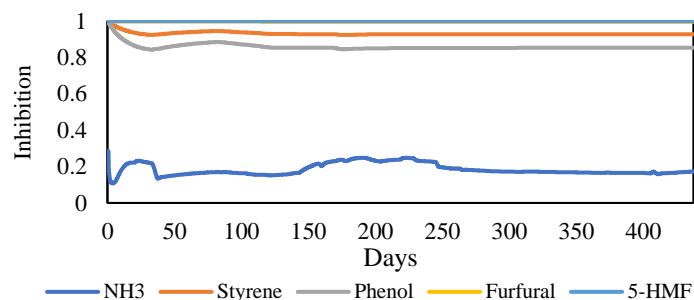
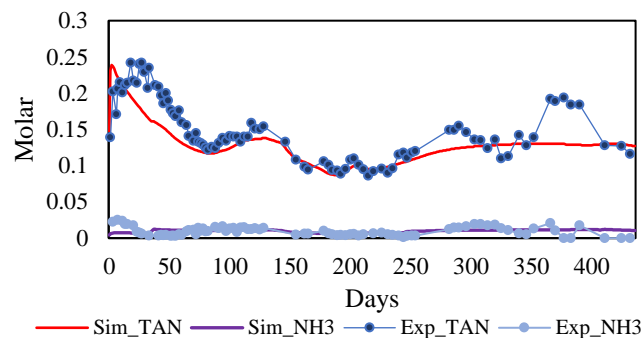


Fig. 4. (Top figure) Simulated total ammonia nitrogen (TAN) (red) and free ammonia ( $\text{NH}_3$ ) (purple) are plotted against experimental data (dots) for the HSPL-reactor. (Bottom figure) The inhibition of  $\text{NH}_3$ , styrene, phenol, furfural, and 5-HMF are plotted where 0=full inhibition and 1=no inhibition.

### 3.3 Discussion

Previous studies simulating digesters fed with HS in ADM1 showed good fits with experimental data, but VFAs were the only parameter challenging to simulate (Donoso-Bravo et al., 2020; Flatabø and Bergland, 2022). Similar problems occurred in this simulation, where the experimental acetic acid concentration was 170% higher than the simulated values.

Another study has shown that effluent COD has been accurately predicted while methane production has been overestimated (Souza et al., 2013a). The current study indicates that simulated methane production was accurately predicted for HS-reactors with an average 10.1% overestimation and sCOD concentrations (not shown) in the effluent was 3.7% overestimated) while the tCOD concentration (not shown) in the effluent was 12.6% underestimated. For the HSPL-reactor, the methane production in the period after 266 days is 14.8% higher, but with significant deviations in the simulation, the simulated tCOD in the effluent (not shown) had an underestimation of 30.4%, and sCOD (not shown) had an underestimation of 26.5%. A previous study of HS in continuous systems with an over/underestimation of 15% was acceptable due to operational variability (Souza et al., 2013b). This study, therefore, shows that the simulation predicts methane production with acceptable accuracy for HS. For the simulation of HS-PL, methane production was better predicted at the end of the period but was not well predicted at high loadings. Souza et al. (2013a) previously reported that methane production deviates in response to low HRTs and consequent load changes. In this study, simulated methane production was not well predicted at low (16.2 days) HRTs,

the first 34 days and between 158 and 245 days, as seen in Fig. 5. However, at high HRT (low mass loadings), the simulation predicts the methane production accurately.

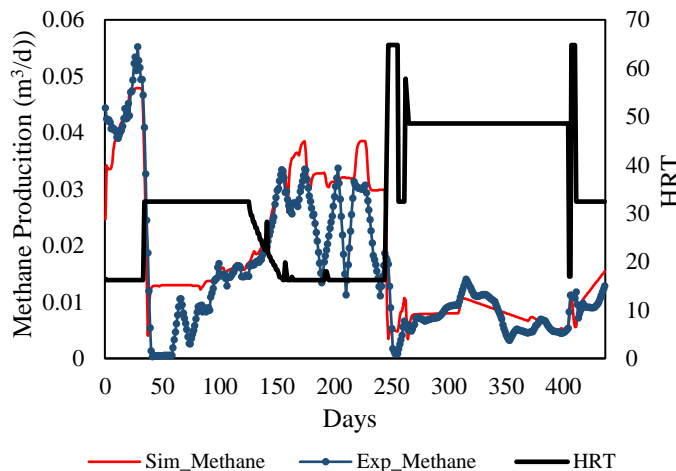


Fig. 5. Simulated (red) and experimental (blue) methane production compared to hydraulic retention time (HRT) (black line).

Considering the operating variability, the simulation's overestimation of 10.1% methane production for the HS-reactor is acceptable. The difference between simulated and experimental values is significantly higher for the HSPL-reactor. Methane production is still overestimated in the stable period with high HRT and low PL loading. The reason might be that the PL consists of more inhibitors than those accounted for in the modelling or that the inhibiting compounds have higher concentrations than the literature has suggested.

### 3.4 Inhibition

Phenols are found in agricultural and industrial wastes and are known to inhibit biodegradation, which makes biological processes difficult to occur (Hernandez and Edyvean, 2008). In AD, it has been indicated that the phenolic compounds interfere with the chain of reactions prior to methanogenesis. Therefore, hydrolysis and acetogenesis are inhibited more than methanogenesis (Hernandez and Edyvean, 2008). Literature suggests that there is some uncertainty regarding phenol methanization where data varies considerably, and some also report that methanization was not achieved (Hernandez and Edyvean, 2008). Another concern is the inhibition constant, which has varying data from 1.12 to 5 kg COD/m<sup>3</sup> (Blum and Speece, 1991; Hernandez and Edyvean, 2008; Raya et al., 2021). Therefore, there is some uncertainty regarding the inhibition constant ( $K_i$ ) and phenols degradation to methane. Styrene in typical AD reactors usually comes from synthetic monomers for plastic production. Fractions of styrene will volatilize, but some will be present in the liquid phase, which can inhibit microbes (Araya et al., 2000). Previous studies have reported styrene in PL (Kessas et al., 2021; Z. Liu et al., 2017); however, the current study's concentration is unknown, and data from the literature was used. Furfural and 5-HMF inhibition were not significant, which was expected because these inhibitors are mainly found in lignocellulosic material or low-temperature pyrolysis (Leng et al., 2021). Phenols and

inorganic nitrogen were expected to be high because of the lignin and protein content (Leng et al., 2021).

TAN (ammonium + NH<sub>3</sub>) is regarded as a nutrient for microbes, but too high concentrations (over 1700 mg/L) can lead to reactor failure (Yenigün and Demirel, 2013). However, microbes can acclimate to high TAN-concentrations when gradually exposed to higher loading over a long period. NH<sub>3</sub> is a potent inhibitor and is in equilibrium with ammonium which increases with a rise in either temperature or pH. According to a previous study of HS-PL, the inhibition from NH<sub>3</sub> was 0.19 at an organic loading rate PL of 0.41 kg tCOD/m<sup>3</sup>/d (Flatabø and Bergland, 2022). In this study, the average inhibition (0.14) is slightly higher than what Flatabø and Bergland (2022) reported but is in the range (0.126-0.268). However, the inhibition of NH<sub>3</sub> is correlated to the inflow of TAN, PL loading, HRT, and pH. Regarding pH, it was relatively stable during the entire period which means that pH did not significantly shift the equilibrium between ammonium and NH<sub>3</sub>. Therefore, this is not considered a major contributor to this simulation compared to the actual TAN concentration loaded. From day 143 to 246, there is less inhibition (see Fig. 4), which can be explained by the lower TAN inflow (average 0.067 M) and the low HRT. However, in the initial period, there was a high TAN (0.1464 M) with low HRT and less inhibition. Based on those findings, it can be seen that low HRT is beneficial for reducing the inhibition of NH<sub>3</sub>. A higher HRT allows more proteins to degrade to TAN. Another possibility is that a low HRT restrains the contact time between microbes and compounds, making the inhibitor less potent. Moreover, the lower HRT gives a more unstable digestion.

Moreover, this article does not take into account the adsorption mechanisms of the different compounds onto the sludge. Phenols can adsorb onto the sludge with a saturation level of 800-1600 mg/L (Hernandez and Edyvean, 2008). Also, microbial adaptation can reduce the effect inhibitors have on methane production over time, which is not taken into account in the model (Donoso-Bravo et al., 2022). Microbes will be selected in continuous reactors based on their adaptability to the substrate, making microbial adaptation an essential factor in the model. The inhibition data also differ from a batch reactor to a CSTR; for this study, the data is not calibrated to a semi-continuous system. Previously, Li et al. (2023) found that  $K_i$  was 2.47 kg COD/m<sup>3</sup> for batch and 6.05 kg COD/m<sup>3</sup> for a continuous system where some microbial kinetics were also changed. This suggests that inhibition and kinetic parameters could be calibrated for a better fit in the model.

### 3.5 Effect of Pyrolysis Liquid addition

The effects of PL addition were evaluated on the methane production of the simulated reactor with HS-PL and the reactor with the same inflow of HS but without PL (see Fig. 6). The difference is noticed as the additional methane production from PL. Based on the results, PL has a positive contribution even though it brings several inhibitors. However, at high PL loadings, it looks like the increased inhibition may lead to reduced methane production.

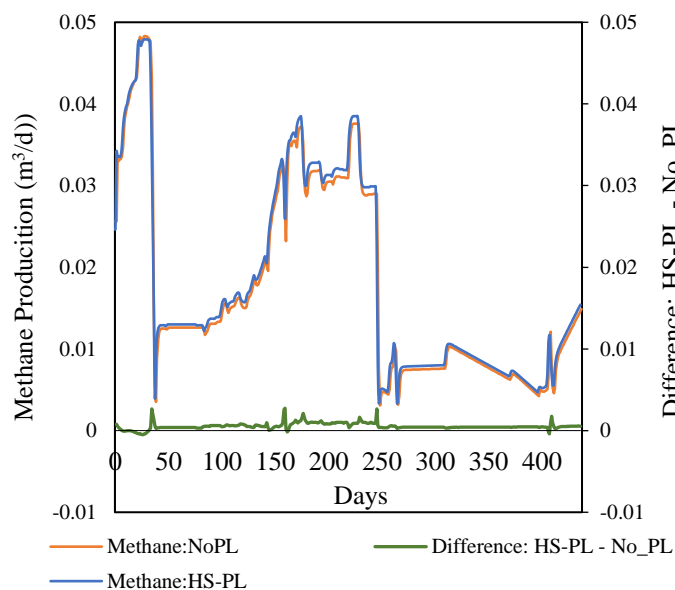


Fig. 6. Simulation of methane production of hydrolyzed sludge with (blue line) or without PL (orange line). The green line shows the difference in methane production between HS with or without PL.

Tab. 4 summarizes the most important parameters and how PL affected them. According to the simulation, PL enhances inhibition but also increases methane production. The methane partial pressure in the biogas and the reactor pH also increased slightly.

Table 4: Average data on PL's influence on AD.

Parameter	PL's influence on AD	St. dev (%)
Methane production	+4.3%	3.0
Inhibition	+31.2%	7.2
pH	+0.8 %	0.9
Methane partial pressure	+1.3%	0.6

### 3. CONCLUSIONS

This study aimed to investigate the effect of PL in co-digestion with HS in an extended ADM1 simulation. The results showed that the model predicted AD of HS with a 10.1% overestimation and that the model did not respond well to sharp load changes. For the HS-PL simulation, the model overpredicts methane production significantly, and the results suggest that there are more inhibitors in the PL than those implemented in the model. According to simulation results, ammonia, phenol, and styrene contributed to most inhibition and increased inhibition by 31%. However, due to the additional COD in the PL, there was additional methane production of 4.3%. Results showed that PL has a positive impact at low loadings. For future work, it is recommended that more studies be conducted on the inhibitor and implementation of microbial adaptation in modelling.

### ACKNOWLEDGEMENTS

The authors would like to thank the NORPART project for financing the expenses for presenting this study at the SIMS

EUROSIM 2024. This study was funded by the University of South-Eastern Norway (USN). Lindum AS and VOW ASA provided substrate and PL for the study.

### REFERENCES

- Araya, P., Chamy, R., Mota, M., and Alves, M. (2000). Biodegradability and toxicity of styrene in the anaerobic digestion process. *Biotechnology letters*, 22, 1477-1481.
- Arnold, M., Reittu, A., Von Wright, A., Martikainen, P. J., and Suihko, M. L. (1997). Bacterial degradation of styrene in waste gases using a peat filter. *Applied Microbiology and Biotechnology*, 48(6), 738-744. doi:10.1007/s002530051126
- Batstone, D. J., Keller, J., Angelidaki, I., Kalyuzhnyi, S. V., Pavlostathis, S. G., Rozzi, A., Sanders, W. T. M., Siegrist, H., and Vavilin, V. A. (2002). The IWA Anaerobic Digestion Model No 1 (ADM1). *Water Science and Technology*, 45(10), 65-73. doi:10.2166/wst.2002.0292
- Bergland, W. H., Dinamarca, C., Toradzadegan, M., Nordgård, A. S. R., Bakke, I., and Bakke, R. (2015). High rate manure supernatant digestion. *Water Research*, 76, 1-9. doi:10.1016/j.watres.2015.02.051
- Blum, D. J. W., and Speece, R. (1991). A Database of Chemical Toxicity to Environmental Bacteria and Its Use in Interspecies Comparisons and Correlations. *Research Journal of the Water Pollution Control Federation*, 63. doi:10.2307/25043983
- Brune, G., Schoberth, S. M., and Sahn, H. (1983). Growth of a Strictly Anaerobic Bacterium on Furfural (2-Furaldehyde). *Appl Environ Microbiol*, 46(5), 1187-1192. doi:10.1128/AEM.46.5.1187-1192.1983
- Donoso-Bravo, A., Olivares, D., Lesty, Y., and Bossche, H. V. (2020). Exploitation of the ADM1 in a XXI century wastewater resource recovery facility (WRRF): The case of codigestion and thermal hydrolysis. *Water Research*, 175, 115654. doi:10.1016/j.watres.2020.115654
- Donoso-Bravo, A., Sadino-Riquelme, M. C., Valdebenito-Rolack, E., Paulet, D., Gómez, D., and Hansen, F. (2022). Comprehensive ADM1 Extensions to Tackle Some Operational and Metabolic Aspects in Anaerobic Digestion. *Microorganisms*, 10(5), 948. doi:10.3390/microorganisms10050948
- Feng, Q. J., and Lin, Y. Q. (2017). Integrated processes of anaerobic digestion and pyrolysis for higher bioenergy recovery from lignocellulosic biomass: A brief review. *Renewable and Sustainable Energy Reviews*, 77, 1272-1287. doi:10.1016/j.rser.2017.03.022
- Fezzani, B., and Ben Cheikh, R. (2009). Extension of the anaerobic digestion model No. 1 (ADM1) to include phenol compounds biodegradation processes for simulating the anaerobic co-digestion of olive mill wastes at mesophilic temperature. *Journal of Hazardous Materials*, 172(2-3), 1430-1438. doi:10.1016/j.jhazmat.2009.08.017

- Flatabø, G. Ø., and Bergland, W. H. (2022, 2022). Anaerobic Co-Digestion of Products from Biosolids Pyrolysis – Implementation in ADM1.
- Flatabø, G. Ø., Cornelissen, G., Carlsson, P., Nilsen, P. J., Tapasvi, D., Bergland, W. H., and Sørmo, E. (2023). Industrially relevant pyrolysis of diverse contaminated organic wastes: Gas compositions and emissions to air. *Journal of cleaner production*, 423, 138777. doi:10.1016/j.jclepro.2023.138777
- Flatabø, G. Ø., Indrebø, T. S., Svennevik, O. K., Ahmed, B., Nilsen, P. J., and Bergland, W. H. (2024). Anaerobic Co-Digestion of Sewage Sludge and its Pyrolysis Condensate: Implications for Methane Production and Filtrate Water Quality. *Heliyon*. doi:10.2139/ssrn.4829930 (Accepted)
- Gąszczak, A., Bartelmus, G., and Greń, I. (2012). Kinetics of styrene biodegradation by *Pseudomonas* sp. E-93486. *Applied Microbiology and Biotechnology*, 93(2), 565-573.
- Giwa, A. S., Xu, H., Chang, F. M., Zhang, X. Y., Ali, N., Yuan, J., and Wang, K. J. (2019). Pyrolysis coupled anaerobic digestion process for food waste and recalcitrant residues: Fundamentals, challenges, and considerations. *Energy Science and Engineering*, 7(6), 2250-2264. doi:10.1002/ese3.503
- Han, D., Lee, C.-Y., Chang, S. W., and Kim, D.-J. (2017). Enhanced methane production and wastewater sludge stabilization of a continuous full scale thermal pretreatment and thermophilic anaerobic digestion. *Bioresource Technology*, 245, 1162-1167. doi:10.1016/j.biortech.2017.08.108
- Hernandez, J. E., and Edyvean, R. G. J. (2008). Inhibition of biogas production and biodegradability by substituted phenolic compounds in anaerobic sludge. *Journal of Hazardous Materials*, 160(1), 20-28. doi:10.1016/j.jhazmat.2008.02.075
- Hubner, T., and Mumme, J. (2015). Integration of pyrolysis and anaerobic digestion - Use of aqueous liquor from digestate pyrolysis for biogas production. *Bioresource Technology*, 183, 86-92. doi:10.1016/j.biortech.2015.02.037
- Kessas, S. A., Esteves, T., and Hemati, M. (2021). Products Distribution During Sewage Sludge Pyrolysis in a Sand and Olivine Fluidized Bed Reactor: Comparison with Woody Waste. *Waste and Biomass Valorization*, 12(6), 3459-3484. doi:10.1007/s12649-020-01209-9
- Leng, L. J., Yang, L. H., Chen, J., Hu, Y. B., Li, H. L., Li, H., Jiang, S. J., Peng, H. Y., Yuan, X. Z., and Huang, H. J. (2021). Valorization of the aqueous phase produced from wet and dry thermochemical processing biomass: A review. *Journal of cleaner production*, 294, Article 126238. doi:10.1016/j.jclepro.2021.126238
- Li, P., Wei, X., Wang, M., Liu, D., Liu, J., Pei, Z., Shi, F., Wang, S., Zuo, X., Li, D., Yu, H., Zhang, N., Yu, Q., and Luo, Y. (2023). Simulation of anaerobic co-digestion of steam explosion pulping wastewater with cattle manure: Focusing on degradation and inhibition of furfural. *Bioresource Technology*, 380, 129086. doi:10.1016/j.biortech.2023.129086
- Liu, B., Ngo, V. A., Terashima, M., and Yasui, H. (2017). Anaerobic treatment of hydrothermally solubilised sugarcane bagasse and its kinetic modelling. *Bioresource Technology*, 234, 253-263. doi:10.1016/j.biortech.2017.03.024
- Liu, Z., McNamara, P., and Zitomer, D. (2017). Autocatalytic Pyrolysis of Wastewater Biosolids for Product Upgrading. *Environmental Science and Technology*, 51(17), 9808-9816. doi:10.1021/acs.est.7b02913
- Raya, D., Ghimire, N., Flatabø, G. Ø., and Bergland, W. H. (2021). Anaerobic Digestion of Aqueous Pyrolysis Liquid in ADM1. doi:10.3384/ecp21185458
- Seyedi, S., Venkiteshwaran, K., Benn, N., and Zitomer, D. (2020). Inhibition during Anaerobic Co-Digestion of Aqueous Pyrolysis Liquid from Wastewater Solids and Synthetic Primary Sludge. *Sustainability*, 12(8), Article 3441. doi:10.3390/su12083441
- Seyedi, S., Venkiteshwaran, K., and Zitomer, D. (2019). Toxicity of Various Pyrolysis Liquids From Biosolids on Methane Production Yield. *Frontiers in Energy Research*, 7, Article 5. doi:10.3389/fenrg.2019.00005
- Souza, T. S. O., Carvajal, A., Donoso-Bravo, A., Peña, M., and Fdz-Polanco, F. (2013a). ADM1 calibration using BMP tests for modeling the effect of autohydrolysis pretreatment on the performance of continuous sludge digesters. *Water Research*, 47(9), 3244-3254. doi:10.1016/j.watres.2013.03.041
- Souza, T. S. O., Ferreira, L. C., Sapkaite, I., Pérez-Elvira, S. I., and Fdz-Polanco, F. (2013b). Thermal pretreatment and hydraulic retention time in continuous digesters fed with sewage sludge: Assessment using the ADM1. *Bioresource Technology*, 148, 317-324. doi:10.1016/j.biortech.2013.08.161
- Tayibi, S., Monlau, F., Marias, F., Cazaudehore, G., Fayoud, N.-E., Oukarroum, A., Zeroual, Y., and Barakat, A. (2021). Coupling anaerobic digestion and pyrolysis processes for maximizing energy recovery and soil preservation according to the circular economy concept. *Journal of Environmental Management*, 279, 111632. doi:10.1016/j.jenvman.2020.111632
- Yenigün, O., and Demirel, B. (2013). Ammonia inhibition in anaerobic digestion: A review. *Process Biochemistry*, 48(5), 901-911. doi:10.1016/j.procbio.2013.04.012



# Green infrastructure for resilient urban design: the mapping and management of green roofs in Oslo

Albert Likang Hu\* Joanna Maria Badach\*,\*\*  
Arnab Chaudhuri\*

\* *Department of Built Environment- Energy, OsloMet  
Oslo Metropolitan University, Pilestredet 35, 0166 Oslo, Norway*

\*\* *Faculty of Architecture, Gdańsk University of Technology  
11/12 Narutowicza Street 80-233 Gdańsk, Poland*

---

**Abstract:** Achieving “Climate-Neutral and Smart Cities” is now very high on the agenda and the city of Oslo has set an even more ambitious goal of becoming a zero-emission city. However, the promotion of more compact development may lead to some negative effects such as the entrapment of polluted air, wind tunnel effects or urban heat islands. Green infrastructure (GI) can be used as a mitigation measure, bringing many benefits such as improving air quality, regulating thermal environment, reducing energy consumption, managing storm water, or promoting urban biodiversity. In this work, we aim to map the existing green roof infrastructure in Oslo and develop an evidence-base strategy for its further development, and enhance the understanding and supplement the existing policies developed by the local authorities. Interviews with stakeholders revealed the practical challenges such as structural limitations, high installation and maintenance costs, and regulatory compliance issues. However, they also recognized the significant environmental advantages that highlight the importance of green roofs in urban sustainability strategies. Geographical information system (GIS) tools are used to identify the potential areas for further green roof implementation, taking into account the spatial, morphological and environmental conditions. 91 Priority green roof areas (PRIOGRAs) and 13 Potential green roof areas (PGRAs) in Oslo are identified as the most suitable for green roof installations after applying filters like roof surface area greater than 250 m<sup>2</sup>, and dominating roof area and slope criteria, exclusion of cultural heritage buildings and existing green roofs, tree density per person deficit, and building age. 2044 roofs can be considered as suitable without the criteria of building age. These findings will potentially help providing actionable insights for policymakers, urban planners, and the research community.

*Keywords:* Nature-based solutions, green infrastructure, green roofs, climate resilience, UHI, water retention, GIS tools.

---

## 1. INTRODUCTION

Modern cities today face unprecedented challenges in achieving sustainability goals due to the rapid growth of urban populations worldwide. Experts from United Nations estimate that by 2041, more than 6 billion people on earth will be living in urban areas (Affairs and Social, 2019). The higher population densities, increased pollution, loss of green spaces, extensive use of heat-absorbing materials like concrete and asphalt and strained energy consumption for heating and cooling are the primary causes of heightened Urban Heat Island (UHI) effect, where temperatures in urban areas become higher than surrounding non-urban areas. This, in turn, is associated with heat-stress-related public health issues and contributes to changes in the local climate as well as global warming effect (Deilami et al., 2018).

Nature-based solutions are recognized as the key response to the challenges posed by the UHI effect, offering a multifaceted approach to cooling urban landscape, at the

same time bringing many other environmental benefits. According to a recent review, local green infrastructure can reduce local peak surface temperatures even by several degrees (Wong et al., 2021). However, due to the lack of open space in many urban areas, sometimes it is challenging to implement the necessary greening solutions to achieve such effects. For this reason, green roofs emerge as a compelling solution that utilize existing roof areas to address the shortage of ground-level green space. Green roofs can mitigate UHI by up to 3 degrees, compensating for the lack of green vegetation in cities, through surface water evaporation, evapotranspiration and decreased albedo effect (Bianchini and Hewage, 2012; Jamei et al., 2021).

Apart from their role in reducing UHI, several other important ecosystem services provided by urban green roofs are also noted in literature, related to energy, pollution or water management. To begin with, they contribute to air pollution control, enhancing local air quality, although this process is more effective when combined with green walls

and green screens, especially in terms of reducing pedestrian level pollution concentration (Viecco et al., 2021). Another important aspect is the improvement of biodiversity, especially when proper design solutions, leading to rich habitat conditions, are implemented (Wang et al., 2022). In addition, existing buildings with retrofitted green roofs can benefit from improved thermal performance, especially during winter months in cold climates, leading to energy savings (Berardi et al., 2014). Finally, in terms of enhancement towards stormwater management, different roof types provide different water retention capacity, and extensive green roofs can reduce over 50% of the potential water runoff from single buildings (Mentens et al., 2006). More recent findings suggest that more technologically advanced blue-green roofs can capture between 70% and 97% of rainfall water during extreme precipitation events, which is considerably higher than that of conventional green roofs (Busker et al., 2022).

The environmental potential of green roofs varies depending on the local conditions. For example, their reported cooling potential is considerably higher in dry climates than in hot-humid climates (Jamei et al., 2021). The same applies to different urban form typologies—the higher cooling potential is captured in compact high-density urban areas than in mid-rise and low-rise neighbourhoods (Zuo et al., 2022). Similarly, the reduction in building energy consumption varies across different climates (Bevilacqua, 2021). The suitability of buildings for green roof installation is influenced by several local factors, including roof size, load capacity, building age, and design considerations. The findings from various studies suggest that a comprehensive evaluation of all of these factors is necessary to optimize the benefits of green roofs in specific urban contexts. It is evident that there is need for more accurate and comprehensive studies of green roof suitability that take into account required ecosystem services and local conditions.

The climate in Nordic countries has been largely unaffected by cooling needs in the summer. However, recent extreme weather events, like the 2012 Copenhagen cloud-burst, which resulted in costs of approximately EUR 1.6 billion (Evaluation, 2012), the Norwegian Hans or the heavy snow during last winter in Oslo, underscore the urgency for adopting more resilient urban planning and design. These recent events call for more focus on green infrastructure strategies in the Scandinavian context, including the more widespread implementation of urban green roofs, in particular for more effective water management (Nordh and Olafsson, 2021). The reported retention of stormwater is even up to 58% of the annual precipitation in the more warm and dry locations (Amorim et al., 2021). Implementing green roofs in these climatic conditions, however, comes with many risks and limitations, related to e.g. to low vegetation survival rate and cover (Lönnqvist et al., 2021).

In Norway, there are no explicit national legislation and regulations dedicated to green roofs. They are often implemented for aesthetic purposes, but there is an urgent need to consider actual spatial and infrastructural conditions and limitations more carefully to leverage their potential environmental benefits. Several parameters of green roof adoption and implementation fall under the broader leg-

islative frameworks related to urban planning, building codes and environment protection and biodiversity conservation. The national standards NS-3840 and NS-3845 set guidelines on green roofs and the Blue green factor (BGF) calculation method (NS-3845, 2020). According to the NS3840 standard, the biggest driver for normalization of green roofs in Norway is its stormwater retention capability, rather than insulation, cooling or green space provision. In addition, Byggforskserien includes technical recommendations regarding Sedum roof and Terraces with plants on load-bearing concrete decks (Byggforsk, 2009, 2013).

While the existing legal and regulatory framework provides avenues for potential integration of green roofs into urban landscapes, there is limited guidance on specific technical aspects on these parameters related to green roofs construction and adoption, thus the on-going efforts to refine and expand relevant national regulations at the local level is needed. Oslo first had green roofs as part of their objectives in their municipal plan strategy towards 2030 to strengthen Oslo's blue-green character (Oslo-Municipality, 2015). An action plan for green roofs and facades towards 2030 was approved in 2023 with three focus areas; Learning, Sharing and Incentives, accompanied by 11 measures (Planning and Building Agency, 2024). According to Helene Egeland, Climate Leader at the Planning and Building Agency (PBA), half of the buildings in Oslo that are larger than 250 square meters have the potential suitability for green roofs (Planning and Building Agency, 2024) but more careful analysis is needed to validate this assumption.

This work aims to assess and enhance the effectiveness of green infrastructure, particularly green roofs, planning in Oslo, towards achieving climate resilience and combating the impacts of extreme weather events. By analysing the existing green roofs and developing evidence-based approach for further development strategies, the study seeks to supplement existing policies and provide actionable insights for policymakers, urban planners, and the research community. Through interviews with key stakeholders and the use of Geographical Information System (GIS) tools, the research identifies suitable locations for more green roofs, prioritising locations with more favourable spatial and environmental conditions.

## 2. METHODOLOGY

### 2.1 Interviews

Semi-structured interviews conducted with experts working with green roof designing, researching, and planning which provided practical experience and knowledge that is relevant to Oslo's GI strategy. These interviews aimed at gathering nuanced insights into the practical experiences, challenges, and opportunities associated with green roof implementation, both for Oslo, Norway, and other comparative Nordic contexts. The selection criteria for interviewees focused on professionals with significant contributions and experiences to GI, ensuring a diverse range of perspectives. The interview process was structured into three parts, with a predetermined set of 9 questions to facilitate in-depth discussion on key topics such as their

own background, green roof design considerations and regulation guidelines, and the future trend of green roof development. Ethical considerations, including informed consent and confidentiality in accordance with the guidelines of the Norwegian Agency for Shared Services in Education and Research (Sikt), were strictly adhered to during the entire interview process. Five stakeholders were selected and interviewed to gain insights into the implementation of green roofs in Oslo and its benefits and challenges, the related maintenance and costs, and limitations were also emphasized during the in-depth interviews. The panel of interviewees included experts from academia, public sector and industries as follows: i) Athenna Grindaker (AFRY Norway, Landscape Architect, Private sector), ii) Bent Chrisitan (Water and sewage agency, Chief Engineer, Municipal planning), iii) David Barton (Norwegian Institute for Natural Research, Senior Research Scientist, Nature research), iv) David V. Brasfield (Norwegian Association for GI, Chairperson, Policy Advocacy), and v) Tore Mauseth (Planning and Building Agency, Environmental Consultant, Municipal planning). Interview insights strengthen the empirical basis of the study and ensures that the conclusions drawn provide insight into the theoretical and practical considerations of green roof design and implementation.

## 2.2 ArcGIS mapping strategy

The latest aerial photo mapping of existing green roofs area in Oslo was conducted by Oslo PBE in 2017, accessing the mapped geodata layer can give a realistic picture on the status of green roofs in Oslo. Within the developed zone of Oslo, the city's first green roofs strategy proposal identified 14 million m<sup>2</sup> of existing flat roof space can potentially be suited for green roofs retrofitting. Oslo's current built zone has 47% green space cover, with 60 m<sup>2</sup> of regulated green space per inhabitant. Based on surveys conducted by Oslo PBA, there is not strong enough evidence to estimate the extent of existing green roofs in square meters without incurring a large margin of error (Planning and Building Agency, 2022a). The goal of measuring green roofs implementation were therefore identified as on the number of roofs suitable for greening, while the results of total roof area was also presented. Oslo Municipality locates in Eastern Norway on the Oslo fjord. The city is the capital and the most populated city of Norway with a city area of 454 km<sup>2</sup>. Oslo's population is projected to grow by more than 100,000 people by 2050, reaching just under 813,000 people. As of January 1, 2024, the city of Oslo had 717710 inhabitants (see (SSB, 2024)).

Oslo presents a unique urban environment for exploring sustainable urban planning solutions. Oslo's climate is classified as Dfb under the Köppen system, characterized by warm summers and cold winters, a classification often referred to as a Humid Continental Climate. This climate type is relevant for the study of green roofs, as it encompasses high temperatures variations and change in precipitation patterns, thereby influencing the design, functionality, and benefits of GI. The Marka peri-urban forest greenbelt in Oslo serves as an ecological corridor connecting various green spaces and habitats. We aim to assess the suitability of green roofs across the city, with a particular emphasis on developed areas. The methodology

for ArcGIS mapping leverages Oslo's publicly accessible geodata repositories, which include detailed urban planning records, highway traffic volume, main drainage lines, and other important environmental data. The geodata layers related to building attributes, roof attributes and urban environments are obtained from Geonorge, Oslo municipality and (Riksantikvaren, 2024). These geodata layers were interpolated to analyze green roofs' spatial distribution and environmental impacts. Each step of the ArcGIS workflow is designed to support the paper's objectives by providing a systematic approach to analyzing the suitability of roofs areas for green roof implementation. The workflow steps are i) Data collection, ii) Data processing, iii) Spatial analysis, iv) Visualisation, v) Output creation and vi) Sharing. This showcases the analytical process which is suitable for reproducibility and validity of the research findings.

## 2.3 The criteria for green roof suitability, urban parameters and tree density per person mapping

As identified in the literature review, green roofs are suitable for roofs that are large enough and flat enough, as well as those with good enough structural integrity. Based on the available geodata layers, the following criteria were chosen to represent the spatial suitability for green roofs (see Table 1). There is also no reliable data on the bearing load capacity of mapped roofs in Oslo, and many other factors like building technique and roofing material could also come into play to determine the actual load capacity of roof surfaces. Buildings built after 01/07/2017 must follow (TEK17, 2017) which specifies the requirement of 20-25 cm of insulation in the wall and 30-35 cm in the ceiling. These are generally considered to be enough to withstand the weight of extensive sedum roofs in addition to snow load, thus the consideration for load capacity is simplified by assuming only buildings built after when TEK17 come into effect can be easily retrofitted for green roof implementation.

This study initially considered a comprehensive set of urban environmental parameters to assess their correlation with the mapping and management of green roofs in Oslo. These parameters included population density, tree density, highway traffic, and public drainage lines. The focus on tree density in this study was chosen due to its strong relevance to various urban environmental parameters and its comprehensive benefits for urban resilience. While other forms of GI, such as grass fields, parks, and green facades, also contribute to urban resilience, trees play a critical role in improving air quality, providing shade and green area, and regulating local urban temperature, thereby mitigating the UHI effect. Additionally, trees enhance urban greenery and contribute to residents' well-being. The definition of forest areas in Norway are defined as area with at least 6 trees per hectare that are or can grow 5 meters high, and these should be evenly distributed over the area. Proximity to greenery can be difficult to measure qualitatively without quantifiable metrics. Incorporation of tree density layer can potentially provide a more reliable and objective assessment of urban greenery quality.

Based on the point layer on nature info on trees, there are totally 177055 trees in Oslo, divided by the total

Table 1. Spatial analysis of green roofs suitability

Parameters	PRIOGRA (Priority green roof area)	PGRA (Potential green roof area)	URA (Unsuitable roof area)
Roof area	Over 250 m <sup>2</sup>		
Dominating roof angle	Under 5°	Between 5° and 30°	Over 30°
Building constraints	Non-SEFRAK buildings		
Building age	After 01/07/2017		

population of Oslo (717710), this resulted the average tree density per person in Oslo to be 0.25 tree per person. The population density of Oslo was distributed onto a 250 by 250 meter grid layer (see (SSB, 2023)) This grid layer was then imported into ArcGIS as the base for further analysis. By combining this population density layer with points layer data on tree distribution, the study aimed to highlight areas where tree density per person was the lowest as an indicator for the need for green area compensation, providing a new perspective on urban greenery and resilience. To achieve this, the nature info of all trees in Oslo was retrieved and mapped onto the same grid. For each grid cell, the total number of trees was divided by the population within that grid cell, creating a new layer that showed tree density per person. The high correlation of ES deficits with population density is a pattern observed in other cities as reported in (Langemeyer et al., 2020). Previous studies have also explored the presence of three trees within a 15-meter distance of a building as an indicator of access to greenery in urban environments (see (Venter et al., 2020)). Incorporating population density into tree density within the grid allows for a more detailed model of tree replacement. This layer considers the population density factor, which provides a more nuanced understanding of urban resilience by emphasizing areas where green roofs could have the most significant impact.

### 3. RESULTS AND DISCUSSION

#### 3.1 Interview insights

The outcome of the interviews essentially reveals the common themes concerning the environmental, social, and economic aspects of green roofs. Interviewees unanimously agreed on the benefits of green roofs, including mitigating public health issues linked to rising urban temperatures, improving air quality, reducing carbon footprints, and enhancing biodiversity. These findings align well with the previous literature reviews. However, despite of Oslo's public green roof strategy, stakeholders still encounter constraints in both the qualitative and quantitative implementation of green roofs. Regarding the spatial suitability of roof areas, *Bent Chrisitan* mentioned that even small roof areas contribute to urban sustainability. *Athena Grindaker* addressed the pressure to compensate for lack of green surfaces on the ground with green roofs at first. However, according to her, budgetary constraints often lead to the elimination of green roofs during later construction phases completely. The critical role of green roofs in effective water management, especially during extreme weather events, was emphasized. The three-step strategy for storm-water management as well as BGF were mentioned as some prevailing guidelines that requires both stringent enforcement and updates in terms of Scandinavian climates. In this aspect, *David Barton* suggests

their work (see (Barton et al., 2021)) on storm-water fees as financial mechanisms in supporting climate readiness through local storm-water management that could potentially motivate property owners to adopt GI solutions like green roofs and other GIs towards enhancing urban resilience.

Apart from environmental benefits, adoption of green roofs have increased land values for certain residential projects. The interviews also revealed the common concerns such as the cost and maintenance of green roofs, and the need for clearer guidelines and supportive policies for their adoption in both new constructions and retrofits. On the other hand, for social side, enhanced aesthetics, improved livability, and recreational spaces were highlighted as the primary benefits. Note that long-term cost savings through energy efficiency and potential increases in property values can be realized as economic benefits. However, there is a need to balance public and private costs, as often residents bear the installation and maintenance costs without realizing potential added land value as mentioned by *David V. Brasfield*.

Furthermore, added costs for retrofitting and unprofessional conduct by builders could also be of concern. For example, air-tightness should be checked via pressure tests for a building before any installation of green roof. As mentioned by *Bent Chrisitan*, it may become very uneconomical and difficult to deal with if roof leakages are discovered after the installation of a green roof. Careful consideration of keeping the prescribed temperature under control during transportation and storage is necessary to avoid irreversible damages (leading dead plant layers) of green roof layers. Interviewees mentioned, stakeholders often come in too late in the building process and thereby incurs additional costs. According to *Tore Mauseth*, "There currently lacks a sufficient method or system for determine and accounting for the benefits for green roofs for not only builders and users of the building but also the public goods comes with it, the architectural implications of green roof accounting has been underestimated." Raising public awareness and developing incentive programs are identified as key strategies to encourage the adoption of green roofs. *David V. Brasfield* also highlighted the need for biophilic design aspects in urban environments for the broader perspective on improving citizens health and well-being.

Another concern is the competition for roof space with other renewable solutions such as solar panels. Evidently the combination of green roofs and solar panels improves the environmental performance. Several interviewees mentioned about the same successful pilot projects in Oslo, those may be considered as well documented examples regarding this. Although these projects are valuable, there

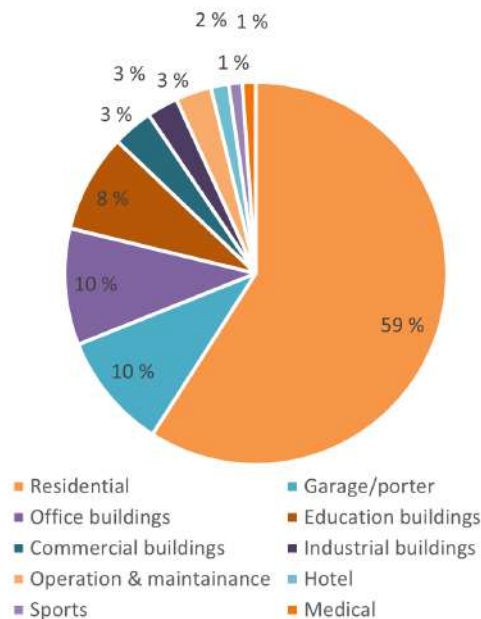


Fig 1. Existing green roofs by building type.

may be a lack of diversity in green roof implementations to consider those as benchmarks.

### 3.2 ArcGIS mapping and analysis

From an aerial photo mapping done by Oslo municipality PBA in 2017, there were 957 green roofs in Oslo with the majority being turf roofs or sedum roofs, many of the 400 green roofs in Vestre Aker district, as well as around 70 roofs from the folk museum on Bygdøy are traditional turf roofs in this mapping. Of the existing green roofs mapped in 2017, 270 were over 250 m<sup>2</sup> with 49% of those being categorized for low utilization for utilizing less than 30% of available roof area. When it comes to building types, 59% of the green roofs were implemented on residential buildings and building blocks (see Fig. 1).

The adoption of green roofs in Oslo has shown promising growth trends over the past decade. Notably, from 2013 to 2017, the city experienced an increase of 75 new green roofs annually, (Planning and Building Agency, 2022b). Continuing from the period from 2017 to today (2024) has likely witnessed another batch of growth in green roof installations. While specific annual growth data from this period are pending, it is reasonable to assume a positive annual development trend over the past seven years well beyond the initial numbers mentioned here. Interviews with stakeholders also revealed that some successful pilot projects like Vega Scene was not captured in our mapping or analysis results, indicating that green roof projects has been proactively implemented in Oslo without considering spatial suitability or ES needs.

In the initial phase of our analysis, we examined the distribution of roof areas across Oslo. Using the dataset of 189,601 roofs, the distribution of roof area size and suitable roof area percentage for all mapped buildings were studied. As depicted, a significant majority of the roofs have an area of less than 250 square meters, this indicates that while

there is a substantial number of larger roofs, they represent a smaller proportion (17077 or 9.01%) of the total number of roofs in Oslo. Consequently, this highlights potential challenges in implementing widespread green roof installations, as smaller roofs may have limited capacity to support green roofs both structurally and economically. Around 46% of the roofs have 100 percent of their existing roof area available for green roofs retrofitting, although this is highly unlikely due to various structural and practical constraints. Additionally, it is important to recognize that this estimation includes a significant margin of error, as there is insufficient evidence to fully support the suitability of all these roofs for green roof installation.

The spatial distribution of potential green roofs in Oslo was analyzed using a series of filter parameters and steps to determine the suitability of roofs for green roof installation. The suitability analysis identified a limited number of roofs in Oslo that fulfill all criteria for potential green roof installations. The filtering process effectively narrowed down the pool of potential roofs, emphasizing the critical parameters that influence suitability. These results, as summarized in Table 2, offer a comprehensive overview of the spatial distribution of potential green roofs in Oslo by outline the area most amenable to green roof installations and the inherent challenges associated with their implementation. Figure 2 illustrates the process of filtering down roofs in Oslo to identify Priority green roof area (PRIOGRA) and Potential green roof area (PGRA). The figures depict the sequential results by filtering criteria applied to determine roof suitability, with PRIOGRA shaded in dark green and PGRA in light green. This series of figures visually demonstrates the step-by-step process of narrowing down the suitable roofs, providing a clear understanding of the criteria and their impact on the spatial distribution of potential green roof areas in Oslo.

The method involved in analyzing tree density per person in 250m × 250m grids across Oslo were done by combining two ArcGIS layers through spatial joining and intersect and arcade code filtering. The results, as depicted in Fig. 3, highlight the areas with the greatest need for GI. The city's average tree density per person is 0.25 tree per person. By categorizing the grids and applying a threshold of less than 0.1 tree per person, the analysis filtered out 90% of the grids with higher tree densities than 0.1, identifying grids with the highest ES needs, where the environmental benefits are the greatest. Figure 3 illustrates the tree density per person across Oslo, depicted by 250m × 250m grids. This analysis categorizes grids based on the number of trees per person, providing a clear spatial representation of urban areas with varying levels of tree density. The categorization into grids with less than 0.1 tree per person was chosen as the critically low tree density area since 90% of the grids in Oslo have less than 1 tree per person. The symbology was achieved through arcade coding filtering the tree density per person per grid. This threshold was used to filter out 90% of the grids, thereby identifying the areas with the highest ES needs, where green roof installations would have the most significant impact. The results of the spatial suitability analysis for potential green roofs in Oslo provide a comprehensive understanding of the area most amenable to green roof installations. The evaluation of these results



Table 2. Summary of results for suitable green roof area

	PRIOGRA	PGRA (Percentage of total roofs)	Total
Filtered after roof area	27485 (14.5%)	73047 (38.5%)	100532 (53%)
Filter after roof dominating area and slope criteria	6278 (3.3%)	5383 (2.8%)	11661 (6.1%)
Filter after cultural heritage buildings and existing green roofs	6033 (3.2%)	1208 (0.6%)	7241 (3.8%)
Filter after tree density per person deficit	1208 (0.6%)	836 (0.4%)	2044 (1%)
Filter after building age	91 (0.05%)	13 (0.007%)	104 (0.057%)

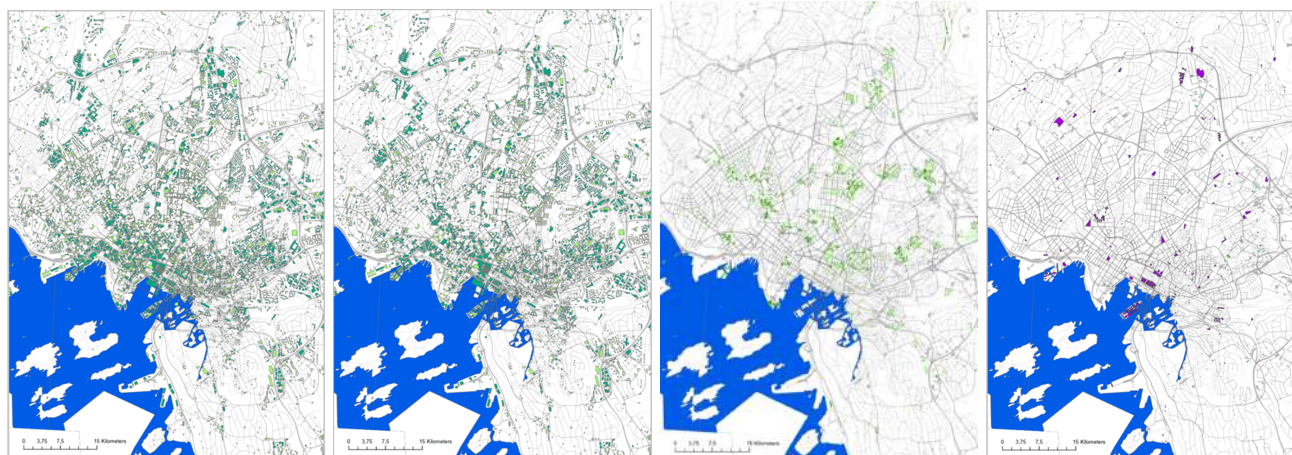


Fig. 2. Spatial distribution of PRIOGRA and PGRA around ring roads in Oslo, PRIOGRA in dark green and PGRA in light green, from left to right: filtered by area and slope requirement, excluding cultural heritage buildings and existing green roofs, tree density per person deficit, and the overall filtered roofs together within existing green roofs in purple.

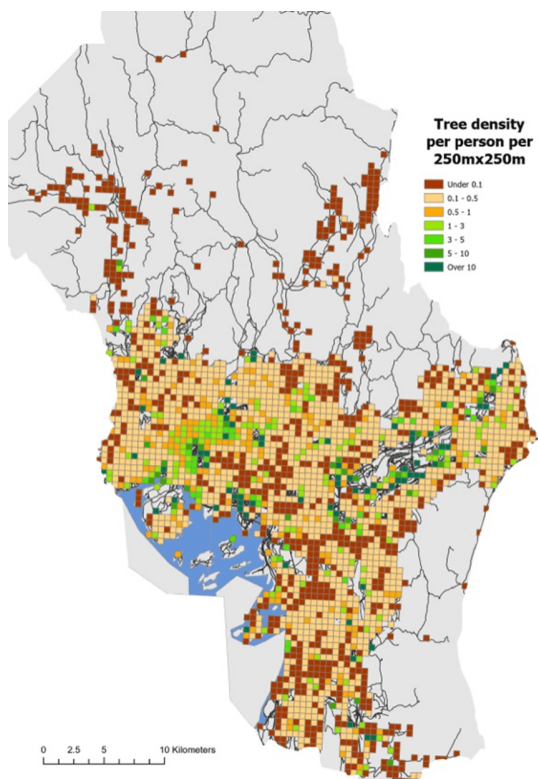


Fig. 3. Tree density per person by 250m x 250m grid.

can be broken down into several key observations and implications:

The initial filtering based on roof surface area greater than 250 m<sup>2</sup> and slope criteria identified 27485 roofs (14.5%) as suitable candidates for green roofs. This number indicates a significant potential for green roof installations in terms of available space. However, the subsequent reduction to 6278 PRIOGRA (3.3%) and 5383 PGRA (2.8%) after applying the slope criteria indicates the importance of considering structural feasibility. Roofs with inappropriate slopes are not conducive to effective green roof installations, highlighting a critical constraint.

The exclusion of cultural heritage buildings and existing green roofs refined the pool of suitable roofs to 6033 PRIOGRA (3.2%) and 1208 PGRA (0.6%). This step was necessary because including historical buildings would pose significant challenges, such as low utilization of roof area, extra costs, and complex installation procedures. This step ensured the exclusion of historical buildings and avoided redundant placement of new green roofs over existing ones. Additionally, many non-historical buildings are either too small in roof area or have sloped roof styles, which already are difficult for green roof installation. The relatively small reduction in the number of suitable roofs suggests that most cultural heritage buildings were already excluded by the roof area and slope criteria, indicating that these three factors together are effective in filtering out unsuitable candidates for green roof installation.

Considering tree density per person provided a more focused approach to identifying areas with the highest ES needs. The significant reduction to 1208 PRIOGRA (0.6%) and 836 PGRA (0.4%) after applying this criterion reveals the limited number of roofs in areas with critically low tree

density. This criterion is crucial for maximizing the environmental benefits of green roofs, targeting urban areas most in need of GI. The significant reduction in suitable roofs when applying the tree density per person criterion highlights the disparity in greenery across Oslo especially when consider population density. Areas with low tree density per person are typically urban zones with limited green spaces, higher population density, and greater environmental stresses such as poor air quality and higher temperatures due to the UHI effect. Implementing green roofs in these areas can provide greater environmental benefits where they are most needed.

The final criterion, focusing on buildings constructed after July 1, 2017, further narrowed the pool to 91 PRIOGRA (0.05%) and 13 PGRA (0.007%). This significant reduction highlights the challenge of retrofitting older buildings with green roofs. Newer buildings, compliant with newer standards, offer better structural support for green roofs, but the small number indicates that recent construction alone cannot meet the need for wider scale green roofs adoption. While identifying 2044 roofs (1208 PRIOGRA and 836 PGRA) in the earlier steps of the analysis seemed promising, the final reduction to just 104 roofs (91 PRIOGRA and 13 PGRA) was unexpected. This drastic narrowing down indicates that these 104 roofs represent the highest priority areas for green roof installations. These roofs should be able to enjoy the maximum benefits that green roofs provide, making them the ideal candidates for Oslo municipality.

Figure 4 illustrates the distribution of the final 104 roofs (91 PRIOGRA and 13 PGRA) by building type. This categorization helps evaluate the potential and focus areas for green roof installations. The figure reveals that 88 out of the 104 selected roofs are residential. This indicates that residential buildings are likely the primary candidates for retrofitting, but there is a need to encourage other building types to consider green roof installations as well. Ensuring that these residential buildings can structurally support green roofs remains a challenge. While these roofs meet the intended criteria, detailed structural assessments are required to confirm their suitability. The cost of retrofitting and maintaining green roofs on residential buildings may vary significantly based on their size and usage. Securing funding and incentives will be crucial for successful implementation. Ensuring all selected roofs comply with local regulations and building codes is essential. This includes obtaining necessary permits and adhering to any specific guidelines for green roof installations. Encouraging other building types (e.g., commercial, industrial) to adopt green roofs is necessary to diversify and optimize the environmental benefits of green roof installations across different sectors. Figure 4 also depicts the potential utilization rate of green roofs relative to the total roof area for the final 104 selected roofs. This analysis was done by calculating the sum of roof area under 5° and between 5° and 30° as a ratio to the total available roof area, this can help evaluate the efficiency and effectiveness of the selected roofs. Potential high utilization rates for green roofs on these selected roofs means higher chances of maximizing environmental benefits. Smaller roofs may face difficulties in achieving optimal utilization. Implementing green roofs on these selected roofs, especially those with irregular shapes or

smaller sizes, poses technical challenges. Customized solutions may be required to address these issues effectively. This includes selecting appropriate plant species and designing systems that support long-term ecological balance.

The findings of this study align with previous research that emphasizes the potential of green roofs to mitigate UHI, improve air quality, and enhance biodiversity. However, the significant reduction in the number of potential green roofs due to stringent suitability criteria is consistent with challenges identified in other studies. Similar findings in other cities highlight structural limitations and regulatory barriers for historical buildings as critical factors affecting green roof feasibility, (Silva et al., 2017).

#### 4. CONCLUSIONS

The study aims to map and identify suitable locations for green roofs in Oslo using GIS analysis. It seeks to develop strategies to optimize the environmental benefits of green roofs, considering local conditions and limitations. Literature reviews highlighted several benefits of green roofs, especially in densely populated areas. The GIS-based suitability analysis identified 1% (1208 PRIOGRA and 836 PGRA) of roofs in Oslo as suitable for green roof installations based on roof area and slope criteria and tree density per person criteria. 104 roofs (91 PRIOGRA and 13 PGRA) were considered as most suitable for green roof installations after applying all relevant criteria. These criteria included factors such as roof area and slope, exclusion of cultural heritage buildings and existing green roofs, tree density per person deficit and building age criteria. While many other roofs did not meet all the criteria, they still hold potential for green roof installations and can offer substantial environmental and social benefits.

Stakeholder interviews highlighted practical challenges such as structural limitations, high installation and maintenance costs, and regulatory issues. Despite these challenges, the environmental benefits of green roofs underscore their value in urban sustainability strategies. The findings of this study provide actionable insights for urban planners and policymakers to optimize the environmental gains from green roofs. The study contributes to the knowledge on GI and offers guidance for future sustainable urban development. Establishing standards for roof insulation and load capacity can aid developers in quantifying the associated costs to implement green roofs. Additionally, financial incentives, potentially subsidies or reductions in tax or lower interest on loans, should be introduced to encourage adoption. Innovative funding mechanisms, such as stormwater fees, could also provide additional support. Integrating green roofs with other sustainable technologies, such as solar panels, will optimize roof space and enhance overall environmental performance. Further research should focus on utilizing spatial data to develop the strategies for developing green roof in conjunction with other forms of GI and mapping their potential spatial distribution. Investigating cost-effective installation and maintenance solutions, as well as financial incentives to overcome identified barriers, is crucial. Further inquiries, including interviews with local architects, planning officers, and engineers, should be conducted to explore practical solutions. Long-term studies should evaluate the impacts of green roofs on urban ecosystems and local climate

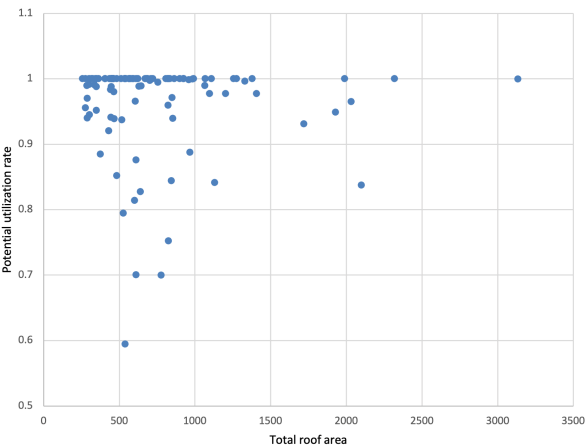
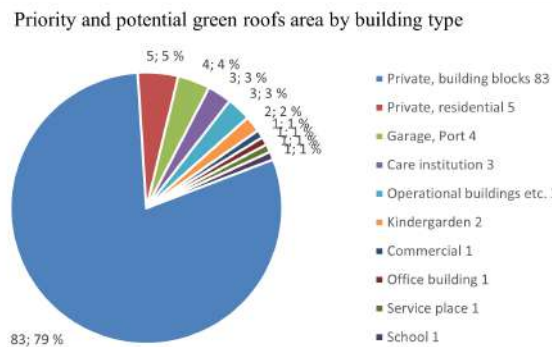


Fig. 4. Priority and potential green roofs area by building type and Potential utilization roof area vs total roof area.

conditions, including their role in mitigating UHI effects, improving stormwater management, and enhancing urban biodiversity. By addressing these areas, future research can provide more evidence-based support and practical solutions to achieve wider scale adoption and effective implementation of green roofs in cities.

#### ACKNOWLEDGEMENTS

This work is an extract of the master thesis of the first author, and we would hereby like to thank all the interviewees for sharing their invaluable insights with us. We have followed the standard procedure following the guidelines of Sikt, the Norwegian Agency for Shared Services in Education and Research related to the personal data related to the interviews with a reference number 190036. We also acknowledge The Planning and Building Agency and The Water and Sewage Agency in Oslo for providing recent reports on existing green roofs and datasets that we have used in this work. (OpenAI, 2024) is used for refining the language of the manuscript.

#### REFERENCES

- Affairs, D.o.E. and Social (2019). World urbanization prospects 2018. Report, United Nations.
- Amorim, J.H., Engardt, M., Johansson, C., Ribeiro, I., and Sannebro, M. (2021). Regulating and cultural ecosystem services of urban green infrastructure in the nordic countries: A systematic review. *International Journal of Environmental Research and Public Health*, 18(3), 1219. URL <https://www.mdpi.com/1660-4601/18/3/1219>.
- Barton, D.N., Venter, Z., Sælthun, N.R., Furuseth, I.S., and Seifert-Dähnn, I. (2021). Brukerfinansiert klimaberedskap? en beregningsmodell for overvannsgelyr i Oslo. *Vann*, 4.
- Berardi, U., GhaffarianHoseini, A., and GhaffarianHoseini, A. (2014). State-of-the-art analysis of the environmental benefits of green roofs. *Applied Energy*, 115, 411–428. doi:10.1016/j.apenergy.2013.10.047. URL <https://www.sciencedirect.com/science/article/pii/S0306261913008775>.
- Bevilacqua, P. (2021). The effectiveness of green roofs in reducing building energy consumptions across different climates. a summary of literature results. *Renewable and Sustainable Energy Reviews*, 151, 111523. doi:10.1016/j.rser.2021.111523. URL <https://www.sciencedirect.com/science/article/pii/S1364032121008017>.
- Bianchini, F. and Hewage, K. (2012). How “green” are the green roofs? lifecycle analysis of green roof materials. *Building and Environment*, 48, 57–65. doi:10.1016/j.buildenv.2011.08.019.
- Busker, T., de Moel, H., Haer, T., Schmeits, M., van den Hurk, B., Myers, K., Cirkel, D.G., and Aerts, J. (2022). Blue-green roofs with forecast-based operation to reduce the impact of weather extremes. *Journal of Environmental Management*, 301, 113750. doi:10.1016/j.jenvman.2021.113750. URL <https://www.sciencedirect.com/science/article/pii/S0301479721018120>.
- Byggforsk, S. (2009). 525.306 terrasser med beplantning på bærende betongdekker [525.306 terraces with planting on load-bearing concrete decks].
- Byggforsk, S. (2013). 544.823 sedumtak [544.823 sedum roofs].
- Deilami, K., Kamruzzaman, M., and Liu, Y. (2018). Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. *International Journal of Applied Earth Observation and Geoinformation*, 67, 30–42. doi:10.1016/j.jag.2017.12.009. URL <https://www.sciencedirect.com/science/article/pii/S0303243417302994>.
- Evaluation, D.I.f.P. (2012). Redegørelse vedrørende skybruddet i storkøbenhavn lørdag den 2. juli 2011 [statement regarding the cloudburst in greater copenhagen on saturday 2 july 2011].
- Jamei, E., Chau, H.W., Seyedmahmoudian, M., and Stojcevski, A. (2021). Review on the cooling potential of green roofs in different climates. *Science of The Total Environment*, 791, 148407. doi:10.1016/j.scitotenv.2021.148407. URL <https://www.sciencedirect.com/science/article/pii/S0048969721034781>.
- Langemeyer, J., Wedgwood, D., McPhearson, T., Baró, F., Madsen, A.L., and Barton, D.N. (2020). Creating urban green infrastructure where it is needed—a spatial ecosystem service-based decision analysis of green roofs in barcelona. *Science of the total environment*, 707, 135487.
- Lönnqvist, J., Hanslin, H.M., Johannessen, B.G., Muthanna, T.M., Viklander, M., and Blecken, G. (2021). Temperatures and precipitation affect

- vegetation dynamics on scandinavian extensive green roofs. *International Journal of Biometeorology*, 65(6), 837–849. doi:10.1007/s00484-020-02060-2. URL 10.1007/s00484-020-02060-2.
- Mentens, J., Raes, D., and Hermy, M. (2006). Green roofs as a tool for solving the rainwater runoff problem in the urbanized 21st century? *Landscape and Urban Planning*, 77(3), 217–226. doi:10.1016/j.landurbplan.2005.02.010. URL <https://www.sciencedirect.com/science/article/pii/S0169204605000496>.
- Nordh, H. and Olafsson, A.S. (2021). Plans for urban green infrastructure in scandinavia. *Journal of Environmental Planning and Management*, 64(5), 883–904. doi:10.1080/09640568.2020.1787960. URL 10.1080/09640568.2020.1787960. 10.1080/09640568.2020.1787960.
- NS-3845 (2020). Ns 3845:2020 blågr ønn faktor - beregningsmetode og vektingsfaktorer [blue-green factor - calculation method and weighting factors].
- OpenAI (2024). Chatgpt (version gpt-4). <https://www.openai.com>. Large language model.
- Oslo-Municipality (2015). Kommuneplan 2015, Oslo mot 2030 [municipal plan 2015, Oslo towards 2030].
- Planning, T. and Building Agency, O. (2022a). Strategi for grønne tak og fasader [strategy for green roofs and facades].
- Planning, T. and Building Agency, O. (2022b). Vedlegg til strategi for grønne tak og fasader [appendix to the strategy for green roofs and facades].
- Planning, T. and Building Agency, O. (2024). Handlingsplan for grønne tak og fasader mot 2030 [action plan for green roofs and facades towards 2030].
- Riksantikvaren (2024). Kulturminner - sefrakbygninger [cultural monuments - sefrak buildings].
- Silva, C.M., Flores-Colen, I., and Antunes, M. (2017). Step-by-step approach to ranking green roof retrofit potential in urban areas: A case study of lisbon, portugal. *Urban forestry & urban greening*, 25, 120–129.
- SSB (2023). Befolkning på rutenett 250 m [population on grid 250 m]. URL <https://kartkatalog.geonorge.no/metadata/befolkning-paa-rutenett-250-m-2023/89bc9a2c-5cb8-4780-8c24-7965c0829696>.
- SSB (2024). Municipal facts, Oslo. URL <https://www.ssb.no/kommunefakta/oslo>.
- TEK17 (2017). *Regulation on Technical Requirements for Construction Works (TEK17)*. Directorate for Building Quality (DiBK), Norway, 1st edition. URL <https://www.dibk.no/regelverk/byggteknisk-forskrift-tek17/14/14-3>.
- Venter, Z.S., Krog, N.H., and Barton, D.N. (2020). Linking green infrastructure to urban heat and human health risk mitigation in oslo, norway. *Science of the Total Environment*, 709, 136193.
- Viecco, M., Jorquera, H., Sharma, A., Bustamante, W., Fernando, H.J.S., and Vera, S. (2021). Green roofs and green walls layouts for improved urban air quality by mitigating particulate matter. *Building and Environment*, 204, 108120. doi:10.1016/j.buildenv.2021.108120. URL <https://www.sciencedirect.com/science/article/pii/S0360132321005217>.
- Wang, L., Wang, H., Wang, Y., Che, Y., Ge, Z., and Mao, L. (2022). The relationship between green roofs and urban biodiversity: a systematic review. *Biodiversity and Conservation*, 31(7), 1771–1796. doi:10.1007/s10531-022-02436-3. URL 10.1007/s10531-022-02436-3.
- Wong, N.H., Tan, C.L., Kolokotsa, D.D., and Takebayashi, H. (2021). Greenery as a mitigation and adaptation strategy to urban heat. *Nature Reviews Earth Environment*, 2(3), 166–181. doi:10.1038/s43017-020-00129-5. URL 10.1038/s43017-020-00129-5.
- Zuo, J., Ma, J., Lin, T., Dong, J., Lin, M., and Luo, J. (2022). Quantitative valuation of green roofs' cooling effects under different urban spatial forms in high-density urban areas. *Building and Environment*, 222, 109367. doi:10.1016/j.buildenv.2022.109367. URL <https://www.sciencedirect.com/science/article/pii/S036013232200600X>.



## Process simulation for biogas upgrading and biomethane recovery using biofilm-based reactors

Vafa Ahmadi. Aryan Bhusal. Gamunu. L. S. Arachchige. Nabin Aryal

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Porsgrunn, Norway (E-mail: [vafa.ahmadi@usn.no](mailto:vafa.ahmadi@usn.no), [258942@student@usn.no](mailto:258942@student@usn.no), [gamunu.arachchige@usn.no](mailto:gamunu.arachchige@usn.no), [nabin.aryal@usn.no](mailto:nabin.aryal@usn.no))

**Abstract:** Microbial biofilm matrices offer numerous benefits in bioprocessing and are crucial in various industrial and remediation processes. They facilitate electron exchange from solid surfaces when they interact with the environment. Emerging technologies such as biofilm-containing trickle bed reactors (TBR) and bioelectrochemical systems (BESs) for carbon dioxide (CO<sub>2</sub>) utilization, mostly rely on microbial biofilm matrices. Metabolic modeling of biofilm-based reactors enables detailed analysis of CO<sub>2</sub> reduction within microorganisms, enhancing reactor efficiency. This study employed simulation models to analyze biomethane synthesis within TBR and BES systems. AQUASIM simulation tool was used for conducting the simulation. Parameters such as non-stoichiometric and stoichiometric ratios of substrates, hydraulic retention time (HRT), biofilm surface area, and applied voltage in BES were varied to evaluate methane (CH<sub>4</sub>) production and microbial biomass growth in TBR and BES. Results demonstrated that 1 day HRT resulted in methanation process failure due to biomass development problem in both TBR and BES. The substrate ratio 1:4 of CO<sub>2</sub> to H<sub>2</sub> increased CH<sub>4</sub> production in the investigated reactors. In BES, in-situ CO<sub>2</sub> and proton (H<sup>+</sup>) generation from oxidation reactions can increase CH<sub>4</sub> production. Whereas in TBR, external H<sub>2</sub> (hydrogen) should be supplied to consume higher amount of CO<sub>2</sub>. The lag phase in TBR was shorter than that in BES because of the greater surface area in TBR. In BES, higher voltage increased the current generation because of development of more biomass on the cathode. The simulation underlines the influence of different variables on biofilm-based reactors, offering critical insights for experimental process design.

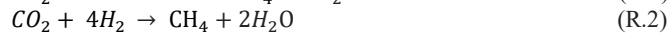
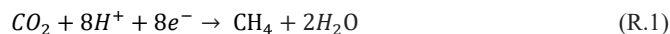
**Keywords:** Biofilms, Trickle bed reactor, Bioelectrochemical system, Methane, Carbon dioxide, and Hydrogen

### 1. INTRODUCTION

In recent decades, bioreactors have gained significant interest for biofuel production, and water and soil remediation due to several environmental advantages. Trickle bed reactors (TBR) which are designed for gas-liquid-solid interaction, can be applied in anaerobic condition for biomethane (CH<sub>4</sub>) production (Germeç et al., 2020). TBR offers high surface area from carrier materials that result higher mass transfer compared to conventional batch and continuous flow reactors (Orgill et al., 2013). Other types of advanced bioreactors are bioelectrochemical systems (BES) that take benefits from electrode-microbe interactions for wastewater treatment and biosynthesis of CH<sub>4</sub> or biochemicals (Pant et al., 2012). Recently, TBR and BES has been intensively applied for CH<sub>4</sub> production from CO<sub>2</sub> or organic waste from wastewater (Ayol et al., 2021).

Hydrogenotrophic methanogenesis in TBR takes place with ex-situ source of H<sub>2</sub>. The process is less complicated, but the biggest limitation is the source of H<sub>2</sub> that should be ecofriendly and economical. In BES, H<sub>2</sub> or H<sup>+</sup> is generated in-situ at response of external voltage to convert CO<sub>2</sub> or organic carbon to CH<sub>4</sub>. The reactor design is more complicated and requires durable electrodes to be applicable in large scale. Despite of complexity of such system, the benefit is that electrons, H<sup>+</sup> and CO<sub>2</sub> are released from anodic oxidation of organic compounds

existing in wastewater. Then two mechanisms can dominate according to the external voltage. R.1 refers to direct electron transfer (DET) mechanism and R.2 is the indirect electron transfer when the cathodic reaction  $2H^+ + 2e^- \rightarrow H_2$  takes place prior to CH<sub>4</sub> production (Eddy et al., 2014; Liu et al., 2019; Nelabhotla and Dinamarca, 2019). R.2 is the governing reaction for CH<sub>4</sub> production in TBR as well.



During anaerobic digestion (AD), the organic matter is broken down by anaerobic microorganisms resulting in CH<sub>4</sub> evolution. However, it is not possible to digest the entire organic matter due to microbial limitations to degrade compounds such as lignocellulose mass and fibrous materials. This results in undigested biomass residues called sludge (Yan et al., 2022). It is not possible to use sludge as fertilizer if it is contaminated by toxic compounds. Therefore, the dried sludge can be utilized in thermal degradation process such as gasification to produce syngas that contains hydrogen (H<sub>2</sub>), CO<sub>2</sub>, CH<sub>4</sub>, carbon monoxide (CO), oxygen (O<sub>2</sub>), nitrogen (N<sub>2</sub>) and other hydrocarbons. H<sub>2</sub>, CO<sub>2</sub> and CO are feasible intermediates which can be fed to TBR or BES reactors for



production of biochemicals or  $\text{CH}_4$  to upgrade biogas (Fericelli, 2011; Maj et al., 2017; Wang et al., 2023).

Other part of the AD effluent is called reject water that comes after mechanical pressing of sludge/ digestate, and contains low organic content, but still has a significant amount of biomass, nitrogen (N) and phosphorus (P) compounds (Fericelli, 2011; Meyer and Wilderer, 2004). The reject water can be used further as nutrient in TBR and BES for biofilm growth. Moreover, it can be treated at anode in BES reactor for organic matter removal (Verma et al., 2023). Autotrophic methanogens which can consume  $\text{CO}_2$  along with  $\text{H}^+$  or  $\text{H}_2$  are dominant in the  $\text{CH}_4$  production process.  $\text{CO}$  in syngas is also a useful energy source and electron donor/acceptor for methanogens through specific pathways (Lim et al., 2022).

Overall, bioprocesses are slow, and variable optimization by experimental studies can usually take from months to years. Process simulation can be a promising approach for process optimization without performing physical lab experiments. There are limited research reports which have done comparative investigation of different biofilm-based processes based on simulation. This work is a conceptual modelling and simulation study for comparison of TBR and BES reactors as different technologies in sludge and reject water management, for  $\text{CH}_4$  recovery. In TBR,  $\text{CO}_2$  to  $\text{H}_2$  ratio from utilizing syngas is typically less than 1:4 which is the theoretical ratio of  $\text{CO}_2$  to  $\text{H}_2$  for complete  $\text{CO}_2$  conversion to  $\text{CH}_4$  (Wang et al. 2023; Eddy et al. 2014). Therefore, external  $\text{H}_2$  source is needed for consumption of the entire supplied  $\text{CO}_2$ . In industrial perspective, if additional  $\text{H}_2$  is supplied from fossil fuels, TBR process may not be ecofriendly.  $\text{H}_2$  source will be sustainable if it is from green sources such as water electrolysis from renewable electricity (Chen et al., 2022). However, this may increase the cost of  $\text{CH}_4$  production in terms of space, energy efficiency, and process complexity.

On the other hand, BES has the advantage of in-situ  $\text{H}_2$  or  $\text{H}^+$  evolution. Such reactors can be single chamber where anode and cathode are placed in the same compartment, or dual chamber where anode and cathode are separated by proton exchange membrane. In double chamber BES, higher voltage can be applied that leads to excessive  $\text{H}^+$  generation. If higher voltage is applied in single chamber BES, it may result in  $\text{O}_2$  evolution that inhibits methanogens. This is not problematic in dual chamber, since  $\text{O}_2$  forms in the anode chamber, and only  $\text{H}^+$  transfers to cathode. The source of  $\text{H}^+$  in low voltage, mostly come from short carbon chain VFAs, but at high voltage, it can be organic nitrogen ( $\text{NH}_4^+$ ), long chain carbohydrates, biomass, and water (Aryal et al., 2020; Batlle-Vilanova et al., 2019). The capacity of  $\text{CO}_2$  conversion to  $\text{CH}_4$  is evaluated to have a comparison, and for better understanding of BES and TBR in handling the  $\text{CO}_2$  according to the availability of  $\text{H}_2$ .

## 2. METHODOLOGY

The concept of this comparative simulation study is shown in schematic diagrams in Figs. 1 and 2. Dried sludge from wastewater treatment plant and AD enters the gasification reactor. The reject water after centrifuge and dryer units will enter to TBR or BES for treatment via microbes. If the process is integrated with TBR (Fig. 1)

$\text{CO}_2$  and  $\text{H}_2$  from gasification enters to TBR for syngas upgrading into  $\text{CH}_4$ . In the alternative of integrating the process with BES reactor (Fig. 2), one input is  $\text{H}_2$  and  $\text{CO}_2$  from gasification. In addition, excess  $\text{CO}_2$  and  $\text{H}^+$  will be available from anodic oxidation of organic compounds in the reject water.

The amount of organic carbon and biomass in the dry sludge was calculated based on the ADM1 Batstone model (Batstone et al., 2002). The model is based on a  $28 \text{ m}^3$  continuous stirred tank reactor with around 7 days HRT. The calculated biomass in AD effluent from the Batstone model was used to calculate the gaseous inputs into TBR and BES. In this stage, only  $\text{H}_2$  and  $\text{CO}_2$  from gasification was used for biomethane production in TBR and BES. The reject water which contained biomass, was used for biofilm formation on the bed of the TBR, and on the electrodes in BES.

AQUASIM (version 2.1) is used as the simulation software which comes with a one-dimensional multi-substrate and multispecies biofilm model. One-dimensional spatial profiles of the microbial species and substrates within the biofilm provides opportunity to forecast real conditions through simulation. Also, it is possible to predict the changes in substrate concentrations, microbial species, and biofilm thickness over time (Reichert, 1998; Wanner and Morgenroth, 2004).

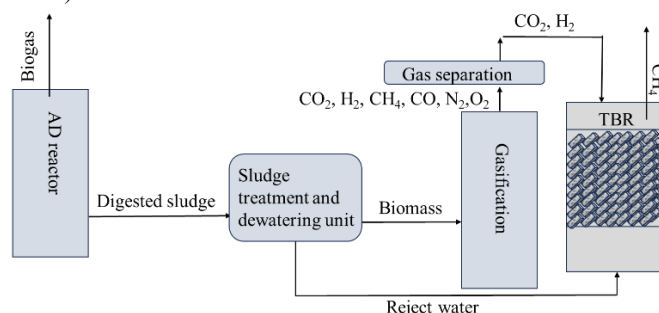


Fig. 1. Schematic diagram of the downstream biogas recovery process with alternative1: trickle bed reactor (TBR) unit. The treated sludge from AD flows to gasification unit.  $\text{CO}_2$  and  $\text{H}_2$  from syngas are applied for simulation of the TBR. Reject water enters the TBR for biofilm formation.

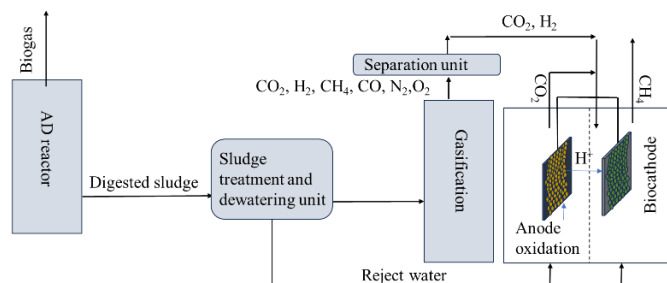


Fig. 2. Dual chamber BES reactor with proton exchange membrane. In case of BES,  $\text{CO}_2$  and  $\text{H}^+$  can be produced from oxidation of organic matter and water at anode.  $\text{CO}_2$  flows out and enters the cathode.  $\text{H}^+$  transfers to cathode through the membrane.

### 2.1 Model assumptions

- The reactors are operated at atmospheric pressure and pH 7.
- In both reactors, the diffusion is considered only on biofilm.

- CO<sub>2</sub> is the only carbon source for cathode which is soluble and in equilibrium with HCO<sub>3</sub><sup>-</sup>.
- All the active biomass is attached. The detached biofilm does not interact with the biofilm reactions.
- Initial biomass fraction is equal for all species.
- In TBR, two types of methanogens grow on the biofilm that grow on CO<sub>2</sub> and H<sub>2</sub> with different yields.
- In BES, two types of methanogens grow on biofilm. One can directly consume CO<sub>2</sub> and H<sub>2</sub>. The other one consumes CO<sub>2</sub> and H<sup>+</sup> via DET mechanism.
- In BES, a cathodic approach is followed, and only cathodic biofilm is involved in simulation of CH<sub>4</sub> production.
- Limitation of the process is the gaseous stream which can be enhanced by continuous flow and gas flow rate. The solubility of H<sub>2</sub> and CO<sub>2</sub> are not equal in atmospheric pressure. Diffusion coefficients can reflect these limitations in the biofilm.
- Every species has the same initial biomass proportion. The yields are assumed as close as possible for the methanogens in TBR and BES.
- Due to operating conditions, the detachment velocity in the simulation as described in Reichert (1998) is taken to be a measure of bacterial decay (K<sub>d</sub>) and biofilm loss.

## 2.2 Simulation approach, inputs, and parameters

Two biofilm-based reactors with continuous flow of gases and biomass are established for the simulation. The flow gases are assumed to come from a gasification reactor which gasify around 1040 mol.m<sup>-3</sup> of organic biomass. The syngas composition from gasification depends on various operating conditions such as temperature, pressure, feedstock and other factors. In this work, according to reported values for sewage sludge, the gasification reactor product is assumed to contain 20% CH<sub>4</sub>, 15% CO<sub>2</sub>, 20% H<sub>2</sub>, 15% CO, 25% N<sub>2</sub>, and 5% trace elements (Wang et al., 2023). If the gasification input is 1040 mol.m<sup>-3</sup> of dry sludge with biomass structure formula of CH<sub>1.8</sub>O<sub>0.5</sub>N<sub>0.2</sub> (Heijnen and Kleerebezem, 1999), 312 mol CO<sub>2</sub> and 416 mol H<sub>2</sub> is obtained according to these considerations. This is assumed as the non-stoichiometric ratio of CO<sub>2</sub> to H<sub>2</sub> flowing into the TBR, and into the cathode chamber of BES.

Calculation of the concentration of excess CO<sub>2</sub> and H<sup>+</sup> in BES was done based on the oxidation half-reactions of organic compounds such as volatile fatty acids (VFAs), from the effluent of ADM1 model. Half-cell voltage inputs below -0.8 V vs SHE, leads to oxidation of organic matter, while it is often not enough for water oxidation. Therefore, the amount of generated H<sup>+</sup> is limited mostly by the organic matter such as short chain VFAs. Low voltage may not provide enough H<sup>+</sup> for reduction of entire CO<sub>2</sub> that flows to the cathode. Thus, higher voltage can be applied for providing more H<sup>+</sup> from oxidation of long carbon chain organic matter and water splitting.

Tab.1 gives the non-stoichiometric and stoichiometric ratios of CO<sub>2</sub>, H<sub>2</sub> and H<sup>+</sup> to the TBR and BES reactor. In TBR reactor, H<sub>2</sub> was considered from external source. For BES, -0.8 and -1.5 V vs SHE was assumed which the lower voltage provides limited H<sup>+</sup>, and the higher voltage can generate enough H<sup>+</sup> for the complete reduction of CO<sub>2</sub> input.

Table 1. CO<sub>2</sub>, H<sub>2</sub> and H<sup>+</sup> inputs for TBR and BES

CO <sub>2</sub> to H <sub>2</sub> ratio from gasifier (mol.m <sup>-3</sup> )	CO <sub>2</sub> to H <sup>+</sup> from oxidation of organics at -0.8 V vs SHE (mol.m <sup>-3</sup> )	Excess H <sub>2</sub> for TBR, and H <sup>+</sup> generated from organics and water oxidation at -1.5 V vs SHE (mol.m <sup>-3</sup> )
CO <sub>2</sub> : 312 H <sub>2</sub> : 416	CO <sub>2</sub> : 42 H <sup>+</sup> : 246	H <sub>2</sub> : 832 H <sup>+</sup> : 1754

Some of the parameters are contemplated equal for both TBR and BES as shown in Tab. 2. The inlet flow rate (Q<sub>in</sub>), initial biomass concentration, the initial biofilm thickness, and the reactor volume are equal in both reactors. It should be noted that the volume of cathode in BES is equal to the total volume of TBR. Since the inputs are of high quantity, lower than a specific volume is not possible to apply. The volume of dual chamber BES was two times greater than TBR. The typical TBR bed area to reactor volume ratio is assumed between 100-1000 m<sup>2</sup>.m<sup>-3</sup> (Manjrekar and Mills, 2022). The cathode size to reactor volume in BES is assumed between 10-100 m<sup>2</sup>.m<sup>-3</sup> (Rabaey and Verstraete, 2005). So, the surface area is adjusted according to these general ratios.

For the autotrophic microbes in TBR, the biomass yield is 0.05-0.3 (Eddy et al., 2014; Thauer et al., 2008). Thus, the yield was considered in the middle range between 0.1 and 0.15. For the BES reactor, the yield can be lower, because the imposed electricity can enhance microbial growth. So, yield of 0.083 (Ahmadi and Aryal, 2024) was taken for electroactive microbes via DET which consume H<sup>+</sup> and CO<sub>2</sub>, and 0.1 for H<sub>2</sub> and CO<sub>2</sub> methanogens. Since the biomass yield has a big impact on CO<sub>2</sub> reduction, both reactors must have close biomass yield to be comparable. The yields which fitted the simulation, were obtained by trial and error.

## 2.3 Model expressions

In TBR simulation, multiplicative Monod equation can be used for calculation of CH<sub>4</sub> production and methanogenic biomass growth ( $\frac{d[X_{CH_4}]}{dt}$ ). Here, H<sub>2</sub> is the electron donor to the microbes, and CO<sub>2</sub> is the only electron acceptor as shown in (1).

$$\frac{d[X_{CH_4}]}{dt} = X_{CH_4} \cdot \left( \mu_{X_{CH_4}}^{max} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H_2}}{K_{H_2} + S_{H_2}} - k_{d,CH_4} \right) \quad (1)$$

Where  $X_{CH_4}$  refers to methanogenic biomass content,  $\mu_{X_{CH_4}}^{max}$  is the maximum growth rate of methanogens,  $S_{CO_2}$  is the molar concentration of CO<sub>2</sub>,  $K_{CO_2}$  is the half-saturation constant of CO<sub>2</sub>,  $S_{H_2}$  is the molar concentration of H<sub>2</sub>,  $K_{H_2}$  is the half-saturation constant of H<sub>2</sub>, and  $k_{d,CH_4}$  is the decay rate of methanogens. The values of each parameter with units are given in Tab. 2.

In BES simulation, methanogens can consume the in-situ generated H<sup>+</sup> via DET. So, the electron acceptor in Monod expression is H<sup>+</sup> together with CO<sub>2</sub>. Both components are limiting factors for CH<sub>4</sub> production. Cathode performs as the

electron donor. The Nernst term in (2) presents the role of electron donor in electroactive CH<sub>4</sub> production. In theory, cathodic reactions are the opposite of anode reactions in terms of the signs.

Table 2. Parameters required for simulation of TBR and BES

Parameter, Unit	Value	Ref
Diffusivity of CH <sub>4</sub> , m <sup>2</sup> · d <sup>-1</sup>	1.296 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of CO <sub>2</sub> , m <sup>2</sup> · d <sup>-1</sup>	1.658 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of H <sub>2</sub> , m <sup>2</sup> · d <sup>-1</sup>	4.43 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of H <sup>+</sup> , m <sup>2</sup> · d <sup>-1</sup>	8.04 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of biomass, m <sup>2</sup> · d <sup>-1</sup>	1 · 10 <sup>-7</sup>	(Ahmadi and Dinamarca, 2022)
Biomass density, mol · m <sup>-3</sup>	222	(Ahmadi and Dinamarca, 2022)
Half-saturation constant of CO <sub>2</sub> , mol · m <sup>-3</sup>	1	(Eddy et al., 2014)
Half-saturation constant of H <sub>2</sub> , mol · m <sup>-3</sup>	0.02	(Eddy et al., 2014)
Half-saturation constant of H <sup>+</sup> , mol · m <sup>-3</sup>	1 · 10 <sup>-4</sup>	(Eddy et al., 2014)
Max growth rate of methanogens 1 in TBR, d <sup>-1</sup>	1.15	Assumed*
Max growth rate of methanogens 2 in TBR, d <sup>-1</sup>	1.24	Assumed*
Max growth rate of methanogens 1 in BES, d <sup>-1</sup>	1.32	Assumed*
Max growth rate of methanogens 2 in BES, d <sup>-1</sup>	1.15	Assumed*
Yield of methanogens_1 in TBR	0.15	Assumed*
Yield of methanogens_2 in TBR	0.12	Assumed*
Yield of methanogens_1 in BES	0.083	Assumed*
Yield of methanogens_2 in BES	0.1	Assumed*
Boundary layer resistance, m	1 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Initial biofilm thickness, m	1 · 10 <sup>-9</sup>	(Ahmadi and Dinamarca, 2022)
Half-maximum growth voltage, V	-0.4	Assumed*
Applied potential on cathode, V	-0.8, -1.5	Assumed*
Cell synthesis coefficient	0.54, 0.4	(Eddy et al., 2014; Ahmadi and Aryal, 2024)
Cathode compartment volume of the BES, m <sup>3</sup>	4	Assumed
TBR reactor volume, m <sup>3</sup>	4	Assumed
Cathodic biofilm surface area, m <sup>2</sup>	400, 2000, 4000	Assumed
Trickle bed surface area, m <sup>2</sup>	40, 200, 400	Assumed

Therefore, the Nernst term which is originally obtained from anodic reactions and anode respiring bacteria (ARB), will be opposite sign in the cathodic approach (Ahmadi and Dinamarca, 2022; Eddy et al., 2014; Rittmann and McCarty, 2020; Torres et al., 2008). The biomass growth rate of electroactive microbes ( $\frac{d[X_{CH_4,el}]}{dt}$ ) can be calculated via (2).

$$\frac{d[X_{CH_4,el}]}{dt} = X_{CH_4} \cdot \left( \mu_{CH_4}^{max} \cdot \frac{1}{1 + \exp\left[\frac{(E_{app} - E_{KA})F}{RT}\right]} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} - k_{d,CH_4} \right) \quad (2)$$

Where  $E_{app}$  is the cathodic voltage,  $F$  is the Faraday constant (96485.3 C · mol<sup>-1</sup>),  $R$  is the gas constant (8.314 J · mol<sup>-1</sup> · K<sup>-1</sup>),  $T$  is the temperature (K),  $S_{H^+}$  is the molar concentration of H<sup>+</sup>,  $K_{H^+}$  is the half-saturation concentration of H<sup>+</sup>.

In the Nernst term in (2), the term  $E_{KA}$  is the voltage when the growth rate of microbes is half of the maximum growth rate. For cathodic microbes, this value falls in the reductive voltage region between the open circuit voltage (OCV) of cathode, and the optimum voltage of biocathode. Various work reported that higher CH<sub>4</sub> production takes place when the voltage is -0.8 and higher (Tremblay et al., 2019). This is due to H<sup>+</sup> evolution from heavy chain organics, inorganic matter, and water oxidation. Assuming cathodic voltage -0.8 V and OCV of the cathode is -0.25,  $E_{KA}$  can fall between -0.25 and -0.8 V. In this work,  $E_{KA}$  was taken as -0.4 V vs SHE (Torres et al., 2008).

The calculated Nernst term (Fig. 3) shows that when  $E_{app}$  is equal to  $E_{KA}$ , the Nernst term becomes equal to 0.5. So, the growth rate is half of its maximum. The approach is according to the kinetic study carried on by Kato Marcus et al., (2007). At  $E_{app} = -0.8$  V, the Nernst term becomes equal to 1, giving the maximum growth rate. From -0.8 to -1.5 V, the Nernst term stays equal to 1. This means the maximum growth rate will be dependent on higher H<sup>+</sup> evolution. In dual chamber BES reactors, -1.5 V vs SHE results in massive H<sup>+</sup> generation which can be beneficial for higher CH<sub>4</sub> production in the cathode chamber.

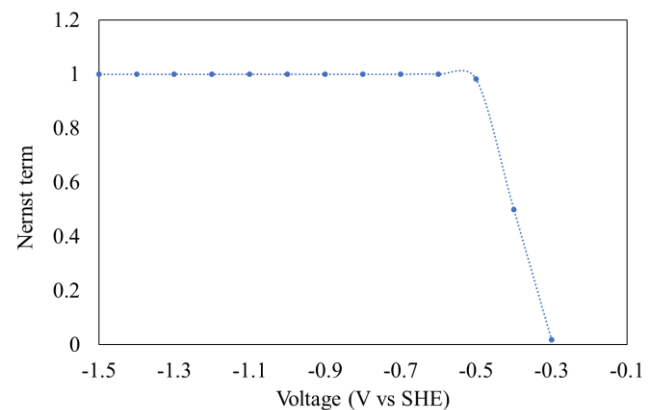


Fig. 3. The resulted values for Nernst term ( $\frac{1}{1 + \exp\left[\frac{(E_{app} - E_{KA})F}{RT}\right]}$ ) at reductive applied voltage when  $E_{KA} = -0.4$  V.



CH<sub>4</sub> production ( $\frac{d[S_{CH_4}]}{dt}$ ) in both TBR and BES was calculated according to (3).

$$\frac{d[S_{CH_4}]}{dt} = \left( \frac{d[X_{CH_4}]}{dt} \right) / Y_{CH_4} \quad (3)$$

Where  $Y_{CH_4}$  refers to methanogenic biomass yield based on CO<sub>2</sub> and H<sub>2</sub> or H<sup>+</sup> consumption.

Change in current density,  $j$  ( $\frac{dj}{dt}$ , A·m<sup>-2</sup>) over time correlates to electroactive biomass concentration ( $\frac{d[X]}{dt}$ ) and the biofilm thickness ( $L_f$ ) which can be calculated by (4) (Torres et al., 2008).

$$\frac{dj}{dt} = \frac{d[X]}{dt} \cdot \gamma \cdot L_f \cdot (f_s^0 - 1) \quad (4)$$

Where  $\gamma$  is 8 which is the number of electrons required for CH<sub>4</sub> generation. The term  $f_s^0$  is the fraction of mol substrate per mol electron equivalent used for cell synthesis. At high voltage, the cell synthesis yield becomes lower. The relationship between  $f_s^0$  and voltage is assessed in another work considering the thermodynamics of the electrochemical reactions (Ahmadi and Aryal, 2024). Moreover, higher voltage increases the current density to the cathode (Aryal et al., 2016; Tremblay et al., 2019). The share of the current density that is consumed for electroactive biofilm growth can be obtained from (4).

### 3. RESULTS AND DISCUSSIONS

For TBR with non-stoichiometric inputs of 1:1.33 CO<sub>2</sub> to H<sub>2</sub> (Fig. 4), the plots are grouped for 3 surface areas with 3 different HRTs. The HRT has a bigger impact than the surface area according to the chosen parameters. The simulation shows that CH<sub>4</sub> production starts with a longer lag phase with 1 day HRT, then a steep methane production happens, but it ends with reactor failure that can be due to the detachment velocity. The corresponding lines for 1 day HRT are those with the highest methane production. Nevertheless, the process could not continue for 50 days, and collapses after 43 days. 4000 m<sup>2</sup> surface area helps reducing the startup time of CH<sub>4</sub> production. Higher surface area can be a better choice when handling higher H<sub>2</sub> inputs.

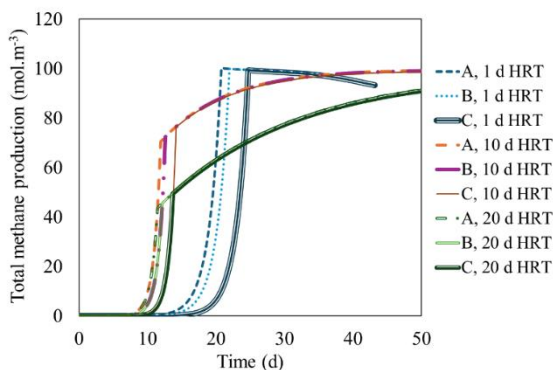


Fig. 4. Methane production in TBR with non-stoichiometric CO<sub>2</sub> to H<sub>2</sub> input. A: 4000 m<sup>2</sup>, B: 2000 m<sup>2</sup>, C: 400 m<sup>2</sup>.

Moreover, 10 days HRT has the shortest startup time, even shorter than 20 days HRT. Also, 10 days HRT leads to higher CH<sub>4</sub> production compared to 20 days HRT. So, 20 days HRT is not efficient, and 1 day HRT is too short for the process. Therefore, to avoid process failure, and for high CO<sub>2</sub> conversion, 10 days HRT with 4000 m<sup>2</sup> surface area (1000 m<sup>2</sup>·m<sup>-3</sup>) was taken as optimum design parameters for TBR with the defined process condition for further evaluation.

For non-stoichiometric conditions in BES (Fig. 5), the oxidation of organics at -0.8 V vs SHE, provides H<sup>+</sup> for the cathode which leads to higher CH<sub>4</sub> production compared to TBR. In BES as well as in TBR, 1 day HRT leads to process failure. 10 days HRT in this reactor is better for the process with a reasonably short lag phase of 9 days. Also, the corresponding CH<sub>4</sub> production indicates that all the H<sub>2</sub> and H<sup>+</sup> available for cathode is consumed for CO<sub>2</sub> reduction. The 400 m<sup>2</sup> surface area (100 m<sup>2</sup>·m<sup>-3</sup>) is advantageous because of a faster startup time which can be reduced to 7 days. Thus, 10 days HRT and 400 m<sup>2</sup> surface area can be taken for cathode for further assessment of stoichiometric inputs.

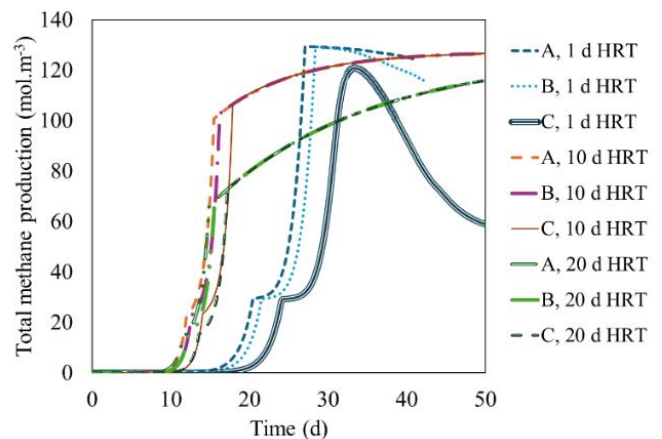


Fig. 5. Methane production in BES with non-stoichiometric CO<sub>2</sub> to H<sub>2</sub> or H<sup>+</sup> input. A: 400 m<sup>2</sup>, B: 200 m<sup>2</sup>, C: 40 m<sup>2</sup>.

Nevertheless, to point out the reason of process deficiency in 1 day HRT for both TBR and BES, Figure 6 illustrates the biomass growth profiles in 1, 10, and 20 days HRT. As depicted in Fig. 6, in 1 day HRT, the biomass growth happens after 10 days with a fast slope, but a sharp decrement in biomass concentration occurs after 15 days in TBR, and after 28 days in BES. This means that failure in CH<sub>4</sub> production is relevant to biomass growth profile in 1 day HRT in both reactors. A possible biomass washout due to fast flow rate can be the reason for the loss of CH<sub>4</sub> production in TBR and BES. Moreover, washout happens earlier in TBR compared to BES reactor. Therefore, 10 and 20 days HRT gives better stability in BES and TBR. In practical experiments also, washout due to fast HRT is reported to cause production failure (He et al., 2024).

The biomass growth for both reactors in 10 days HRT is stable, and results in higher CH<sub>4</sub> formation (Figs. 4 and 5). The biomass growth profile peaks at day 9 in TBR, and reaches the highest at day 14 in BES. In 10 days HRT, the amount of biomass on the biofilm is higher than that in 20 days HRT, which results in higher CH<sub>4</sub> production.

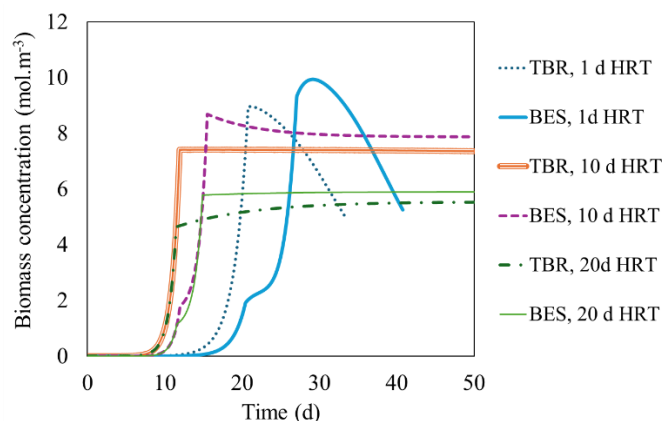


Fig. 6. Biomass concentration in TBR reactor with 4000 m<sup>2</sup> biofilm surface area, and in BES reactor with 400 m<sup>2</sup> biofilm surface area in non-stoichiometric condition at 1, 10, and 20 days HRT.

For evaluating TBR and BES in CO<sub>2</sub> reduction with stoichiometric input values, Figure 7 shows the results for 10 days HRT with 4000 m<sup>2</sup> biofilm area in TBR, and with 400 m<sup>2</sup> biofilm area in BES reactor. In Fig. 7, CH<sub>4</sub> production starts faster in TBR than that in BES. The production starts at day 8 in TBR, and at day 11 in BES reactor. Nevertheless, the amount of CH<sub>4</sub> which can be obtained by BES is higher than that in TBR. One reason is the surplus CO<sub>2</sub> which is generated from the oxidation of organic matter in BES. The other reason is the lower biomass yield in BES. In BES, as explained earlier, the biomass yield will be thermodynamically lower because the external electrical energy assists the cathodic biofilm to grow. So that, lower amount of CO<sub>2</sub> is consumed for cell synthesis. On the other hand, since the biomass yield and the maximum growth rate of microbes are considered close values, the biomass concentration reaches 22.88 mol.m<sup>-3</sup> in TBR, and 19.3 mol.m<sup>-3</sup> in BES. With stoichiometric H<sub>2</sub> input, 299 mol.m<sup>-3</sup> CH<sub>4</sub> will be produced, and the rest of CO<sub>2</sub> will be consumed for cell synthesis. In BES, 340.5 mol.m<sup>-3</sup> CH<sub>4</sub> will be produced, and the remaining CO<sub>2</sub> is consumed for cell synthesis. As it is obvious, in the BES reactor, H<sup>+</sup> is responsible for utilizing a bigger share of the CO<sub>2</sub> inlet. So, CH<sub>4</sub> that is generated from the electroactive methanogens is higher than the methanogens which consume H<sub>2</sub> for CO<sub>2</sub> reduction to CH<sub>4</sub>.

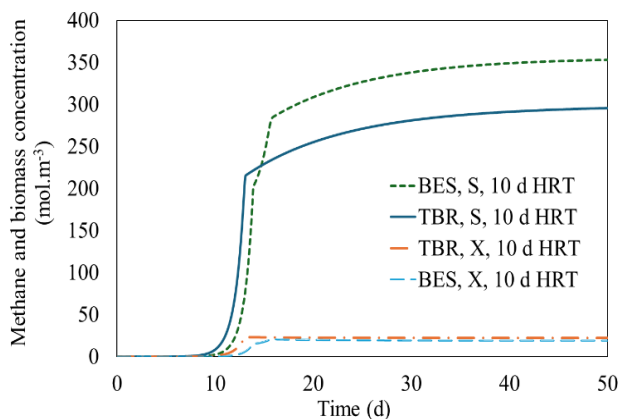


Fig. 7. CH<sub>4</sub> production and biomass concentration in TBR and BES reactor for stoichiometric values in 10 days HRT.

The current density was calculated using the electroactive biomass concentration together with the biofilm thickness which was obtained from the simulation. In the dual chamber BES, the generated H<sup>+</sup> at low and high voltage, is the result of current density and electron flow from anode to cathode. The amount of current density which correlates to electricity consumption for biomass growth can be calculated over time. However, at lower reductive voltage, lower biomass can aggregate on biofilm that corresponds to less CH<sub>4</sub> formation. Figure 8 shows that the reductive voltage for organic matter oxidation corresponds to -0.03 mA.m<sup>-2</sup> current density at -0.8 V vs SHE. Nevertheless, CO<sub>2</sub> reduction is low with respect to low H<sup>+</sup> generation. To generate higher H<sup>+</sup> to consume more CO<sub>2</sub>, -1.3 mA.m<sup>-2</sup> current density is obtained. That means the amount of biomass which grows on biocathode, requires higher current density for complete CO<sub>2</sub> reduction. In a double chamber reactor, higher voltage results in higher H<sup>+</sup> generation (Pisciotta et al., 2012).

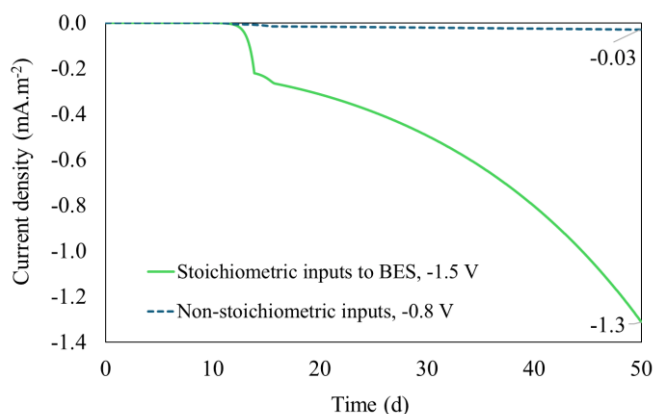


Fig. 8. Current density of the BES reactor at -0.8 and -1.5 V vs SHE corresponding to the current generation for non-stoichiometric and stoichiometric H<sup>+</sup> generation in the reactor.

#### 4. CONCLUSIONS

This work carried on a preliminary study of the concept of TBR and BES reactors integrated with an AD reactor for downstream sludge treatment using gasifier and reject water treatment. The simulation showed that 4000 m<sup>2</sup> surface area for TBR and 400 m<sup>2</sup> cathode surface area for BES is beneficial for reducing the starting time of CH<sub>4</sub> production. 1 day HRT leads to operation failure. Therefore, in the defined operating conditions in this simulation work, in both processes, 10 days HRT led to stable and optimum CH<sub>4</sub> production. TBR is more straightforward than BES, however, greater amount of CH<sub>4</sub> can be achieved in BES reactor due to excess CO<sub>2</sub> and H<sup>+</sup> from BES process. The complexities of TBR are the source of excess H<sub>2</sub> that should be provided from green sources. Therefore, economical aspects of in-situ and ex-situ H<sub>2</sub> production should be studied further. Overall, both processes seem to be promising in valorizing CO<sub>2</sub> to CH<sub>4</sub>. Moreover, CO is an electron donor source. So, in the next step of the study, CO will be studied together with CO<sub>2</sub> and H<sub>2</sub> in the biofilm processes.



## REFERENCES

- Ahmadi, V., and Aryal, N. (2024). Evaluation of the relationship between voltage and microbial yield in a bioelectrochemical reactor for optimization of methane production, in: 18th IWA World Conference on Anaerobic Digestion – IWA. IWA, Istanbul.
- Ahmadi, V., and Dinamarca, C. (2022). Simulation of the Effect of Local Electric Potential and Substrate Concentration on CO<sub>2</sub> Reduction via Microbial Electrosynthesis.
- Aryal, N., Ghimire, N., and Bajracharya, S. (2020). Coupling of microbial electrosynthesis with anaerobic digestion for waste valorization, in: *Advances in Bioenergy*. Elsevier, pp. 101–127.
- Aryal, N., Halder, A., Tremblay, P.-L., Chi, Q., and Zhang, T. (2016). Enhanced microbial electrosynthesis with three-dimensional graphene functionalized cathodes fabricated via solvothermal synthesis. *Electrochim Acta* 217, 117–122.
- Ayol, A., Peixoto, L., Keskin, T., and Abubackar, H.N. (2021). Reactor designs and configurations for biological and bioelectrochemical C1 gas conversion: A review. *Int J Environ Res Public Health* 18, 11683.
- Batlle-Vilanova, P., Rovira-Alsina, L., Puig, S., Balaguer, M.D., Icaran, P., Monsalvo, V.M., Rogalla, F., and Colprim, J. (2019). Biogas upgrading, CO<sub>2</sub> valorisation and economic revaluation of bioelectrochemical systems through anodic chlorine production in the framework of wastewater treatment plants. *Science of The Total Environment* 690, 352–360. doi.10.1016/j.scitotenv.2019.06.361
- Batstone, D., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S., Rozzi, A., Sanders, W., Siegrist, H., and Vavilin, V. (2002). Anaerobic digestion model No 1 (ADM1). *Water Sci Technol* 45, 65–73.
- Chen, Z., Wei, W., Song, L., and Ni, B.-J. (2022). Hybrid water electrolysis: A new sustainable avenue for energy-saving hydrogen production. *Sustainable Horizons* 1, 100002.
- Eddy, M., Abu-Orf, M., Bowden, G., Burton, F.L., Pfrang, W., Stensel, H.D., Tchobanoglous, G., and Tsuchihashi, R., (Firm), A. (2014). *Wastewater engineering: treatment and resource recovery*. McGraw Hill Education.
- Fericelli, P.D. (2011). Comparison of sludge treatment by gasification vs. incineration, in: Ninth LACCEI Latin American and Caribbean Conference (LACCEI'2011), Engineering for a Smart Planet, Innovation, Information Technology and Computational Tools for Sustainable Development.
- Germec, M., Demirci, A., and Turhan, I. (2020). Biofilm reactors for value-added products production: An in-depth review. *Biocatal Agric Biotechnol* 27, 101662. doi.10.1016/j.bcab.2020.101662
- He, Huiban, Wang, Z., Wang, W., He, Haoxing, Yan, J., Wang, H., Cui, Z., and Yuan, X. (2024). Mitigating short-circuits through synergistic temperature and hydraulic retention time control for enhancing methane yield in continuous stirred-tank reactors. *Energy* 289, 129914. doi. 10.1016/j.energy.2023.129914
- Heijnen, J.J., and Kleerebezem, R. (1999). Bioenergetics of microbial growth. *Encyclopedia of bioprocess technology: Fermentation, biocatalysis and bioseparation*, 1, 267-291.
- Kato Marcus, A., Torres, C.I., and Rittmann, B.E. (2007). Conduction-based modeling of the biofilm anode of a microbial fuel cell. *Biotechnol Bioeng* 98, 1171–1182.
- Lim, J.K., Yang, J.-I., Kim, Y.J., Park, Y.-J., and Kim, Y.H. (2022). Bioconversion of CO to formate by artificially designed carbon monoxide: formate oxidoreductase in hyperthermophilic archaea. *Commun Biol* 5, 539.
- Liu, C., Sun, D., Zhao, Z., Dang, Y., and Holmes, D.E. (2019). Methanoxithrix enhances biogas upgrading in microbial electrolysis cell via direct electron transfer. *Bioresour Technol* 291, 121877.
- Maj, I., Kudlek, E., Ostrowski, P., and Pronobis, M. (2017). Investigation of sewage sludge gasification with use of flue gas as a gasifying agent, in: E3S Web of Conferences. EDP Sciences, p. 02001.
- Manjrekar, O., and Mills, P. (2022). Trickle Bed Reactors. pp. 533–588. doi.10.1002/9783527812066.ch9
- Meyer, S.S., and Wilderer, P.A. (2004). Reject water: Treating of process water in large wastewater treatment plants in Germany—A case study. *Journal of Environmental Science and Health, Part A* 39, 1645–1654.
- Nelabhotla, A.B.T., and Dinamarca, C. (2019). Bioelectrochemical CO<sub>2</sub> reduction to methane: MES integration in biogas production processes. *Applied Sciences* 9, 1056.
- Orgill, J.J., Atiyeh, H.K., Devarapalli, M., Phillips, J.R., Lewis, R.S., and Huhnke, R.L. (2013). A comparison of mass transfer coefficients between trickle-bed, hollow fiber membrane and stirred tank reactors. *Bioresour Technol* 133, 340–346.
- Pant, D., Singh, A., Van Bogaert, G., Olsen, S.I., Nigam, P.S., Diels, L., and Vanbroekhoven, K. (2012). Bioelectrochemical systems (BES) for sustainable energy production and product recovery from organic wastes and industrial wastewaters. *RSC Adv* 2, 1248–1263.
- Pisciotta, J.M., Zaybak, Z., Call, D.F., Nam, J.-Y., and Logan, B.E. (2012). Enrichment of microbial electrolysis cell biocathodes from sediment microbial fuel cell bioanodes. *Appl Environ Microbiol* 78, 5212–5219.
- Rabaey, K., and Verstraete, W. (2005). Microbial fuel cells: novel biotechnology for energy generation. *Trends Biotechnol* 23, 291–298.
- Reichert, P. (1998). Aquasim 2.0-user manual. Swiss Federal Institute for Environmental Science and Technology. Dübendorf, Switzerland.
- Rittmann, B.E., and McCarty, P.L. (2020). *Environmental Biotechnology: Principles and Applications*, 2nd Edition. ed. McGraw-Hill Education, New York.

- Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008). *Methanogenic archaea: ecologically relevant differences in energy conservation*. *Nat Rev Microbiol* 6, 579–591.
- Torres, C.I., Marcus, A.K., Parameswaran, P., and Rittmann, B.E. (2008). Kinetic Experiments for Evaluating the Nernst–Monod Model for Anode-Respiring Bacteria (ARB) in a Biofilm Anode. *Environ Sci Technol* 42, 6593–6597. doi. 10.1021/es800970w
- Tremblay, P.-L., Faraghiparapari, N., and Zhang, T. (2019). Accelerated H<sub>2</sub> evolution during microbial electrosynthesis with *Sporomusa ovata*. *Catalysts* 9, 166.
- Verma, S., Kuila, A., and Jacob, S. (2023). Role of biofilms in waste water treatment. *Appl Biochem Biotechnol* 195, 5618–5642.
- Wang, B., Gupta, R., Bei, L., Wan, Q., and Sun, L. (2023). A review on gasification of municipal solid waste (MSW): Syngas production, tar formation, mineral transformation and industrial challenges. *Int J Hydrogen Energy*.
- Wanner, O., and Morgenroth, E. (2004). Biofilm modeling with AQUASIM. *Water Science and Technology* 49, 137–144.
- Yan, W., Xu, H., Lu, D., and Zhou, Y. (2022). Effects of sludge thermal hydrolysis pretreatment on anaerobic digestion and downstream processes: mechanism, challenges and solutions. *Bioresour Technol* 344, 126248. doi. 10.1016/j.biortech.2021.126248

# Modeling and Simulation of Full-Scale Sequential Batch Reactor Biological Process Using GPS-X

I.K.B. Stenberg\*, R.S. Bekele\*\*\*, G.S. Arachchige\* E. Janka\*

\*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway  
Kjølnes ring 56, 3918 porsgrunn, Norway (e-mail: [eshetu.j.wakjera@usn.no](mailto:eshetu.j.wakjera@usn.no)).

\*\*Department of Urban Environmental Management, College of Business Technology and Vocational Education, Kotebe  
University of Education, Addis Ababa, Ethiopia

**Abstract:** Sequencing Batch Reactors (SBR) are effective for treating reject water and high-strength industrial wastewater. The Knardalstrand wastewater treatment plant in Porsgrunn recently upgraded the plant with two full-scale SBRs with volumes of 115 m<sup>3</sup> and 100.4 m<sup>3</sup>, focusing on simultaneous nitrification and denitrification. This project aimed to model and simulate these processes using GPS-X software to optimize the municipal wastewater treatment. Two SBR models – simple and advanced – were developed for ammonia nitrogen (NH<sub>4</sub>-N) removal. Results indicated that the advanced model outperformed the simple model, achieving a 75% nitrogen removal efficiency at a dissolved oxygen (DO) setpoint of 1.5 mg/L through shortcut partial nitrification-denitrification. The advanced SBR model's ammonia removal efficiency through nitrification increased with an increase in dissolved oxygen (DO) concentration. In all the simulations there was high nitrite (NO<sub>2</sub><sup>-</sup>) accumulation in the reactors, which could be due to the partial nitrification coupled with denitrification. The simulation has showed the presence of a simultaneous nitrification-denitrification process in the full-scale SBR plant. Enhancing model accuracy with quality data and optimizing DO levels can significantly reduce operational energy costs.

**Keywords:** GPS-X, Sequential Batch Reactor, Nitrification, Denitrification, model, Simulation.

## 1. INTRODUCTION

The need for efficient wastewater treatment systems is driven by the rapid growth of world's population, urban development and industrialization (Naidoo and Olaniran, 2013; Singh et al., 2023). Wastewater often has high organic, nitrogen and phosphorus contents and when these wastewater constituents are discharged to water bodies without efficient treatment the consequences can have serious ecological effects (Kesari et al., 2021; Singh et al., 2023). High concentrations of nitrogen and phosphorus in discharged wastewater can lead to eutrophication, resulting in significant blooms of planktonic algae or phytoplankton. This algal blooming increases the quantity of organic matter in the water body and result in depletion of oxygen content in the water. The process of eutrophication can have disastrous impacts in the aquatic ecosystem where certain species may perish because the dissolved oxygen (DO) content in water drops below a critical level (Bhagwati and Ahamad, 2019).

Wastewater treatment technologies have been developed to handle the complex organic and nutrient pollutants and meet the discharge regulatory requirements. The Sequential Batch Reactor (SBR) is one of these novel technologies that has shown to be an effective and flexible method to remove organics and nutrients from wastewater (Azeez et al., 2023). The broad use of SBR system in both municipal and industrial context has been driven by its flexibility to change influent characteristics, high treatment efficiency, small areal footprint, and low energy use (Fernandes et al., 2013).

The SBR processes has mainly five phases (Fig. 1). The first is **Fill phase** where raw wastewater is the influent to SBR. The

second is **React phase** where dissolved oxygen (DO) and mixing starts in SBR process. The third is **Settle phase** where the activated sludge is separated from the liquid by the process of settling. The fourth is **Draw phase** where clear supernatant is removed using a decanter, and fifth is **Idle phase** where a sludge wasting process.

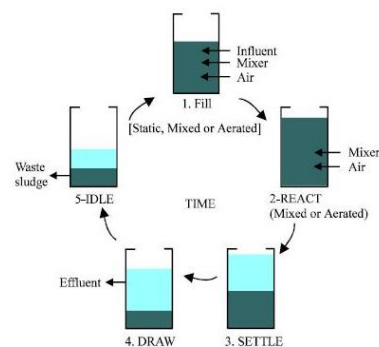


Fig. 1. The five phases of SBR operation for the one cycle periods of fill, react, settle, draw and idle (Irvine and Ketchum, 2004).

Modeling and simulation of full-scale SBR is vital to understand the plant process and control of important process parameters. Several models have been applied in SBR process specially using the activated sludge model (ASM) (Man et al., 2017; Zhou et al., 2013). However, few works have been done in SBR modeling and simulation using a GPS-X modeling platform. GPS-X is a modeling and simulation program, especially for planning and optimizing large-scale or full-scale treatment plants and processes. The biological, chemical, and physical unit activities of wastewater treatment processes can be built into sophisticated models by the software and

simulated under different operating situations. In addition, GPS-X provides an extensive library of pre-defined process models and kinetic equations that may greatly accelerate modeling and serve as a strong basis for dynamic simulation and real-time data integration.

For modeling and simulation of full-scale wastewater treatment plants, GPS-X was applied in the simulation of nitrogen removal, to optimize textile wastewater decision support, for energy consumption and cost minimization of full-scale wastewater treatment plants (Cao et al., 2021; Sadri Moghaddam and Pirali, 2021; Sean et al., 2020; Sid et al., 2017; Wondim et al., 2023). The input data and parameters have significant impact on the accuracy of the simulation outcomes, even though the GPS-X program uses sophisticated mathematical methods and simulation techniques. Hence, validation of the model output with the actual data is vital (Cao et al., 2021; Dolatshah et al., 2024; Łagód et al., 2019). In general GPS-X is a useful tool with an intuitive interface and an extensive library of process models for modeling and simulating wastewater treatment processes in SBR.

Therefore, the aim of this study was to model and simulate the two full-scale SBR plant at Porsgrunn wastewater treatment plant, Norway. The result of the modeling result aims to support process control decisions system and treatment efficiency of SBR for a robust wastewater treatment reactor with small areal footprint and energy use.

## 2. MATERIALS AND METHODS

### 2.1 Simple SBR model setup in GPS-X

The model environment in GPS-X has several unit process objects available for the SBR. The simple SBR model unit has a fixed order of phases, and the user can determine the duration of these phases. The model is simplified and focuses on the side stream treatment using the SBR unit. Therefore, the model only consists of the SBR unit and its inflows and outflows (Fig. 2). The influent flow of the real SBR was a combination of the reject water and the water from thickener. The reject water and its influent characteristics are listed below in table 1, together with the default values in GPS-X. The physical characteristics the full-scale SBR was working volume of 115 m<sup>3</sup>, a max water filling height of 4.1 m and a reactor surface area of 28.05 m<sup>2</sup>. The simulation was set to simulate for 30 days by setting the stopping time to 30.0 [d] in the GPS-X simulation setup, with a communication interval set to 0.05 [d]. The wastewater temperature was set to 28 °C with a set pH value to 7.7.

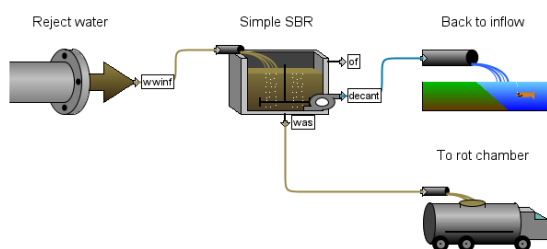


Fig. 2. The model environment of the SBR model in GPS-X with the corresponding side streams reject water influent, effluent and model objects.

### 2.2 Advanced SBR model setup in GPS-X

To implement a model that is more representative to the real SBR at Knarrdalstrand wastewater treatment plant (WWTP), the advanced SBR model unit in GPS-X was used. The advanced model unit allows for more freedom in parameter setup, order and number of phases. Additionally, it allows to decide the phase conditions in terms of aeration, mixing, settling and flowrates. The physical characteristics of the advanced SBR model is identical to the simple SBR model. The aeration set up was using a dissolved oxygen (DO) controller with a proportional-integral-derivative (PID) controller type in velocity form having a derivative kick protection turned off. Fifteen operating phases different in duration (minutes), mixing and DO (mgO<sub>2</sub>/L) concentration were setup.

The Knardalstrand WWTP includes more physical, chemical and biological process unit objects. The new process unit objects include a grit chamber, Ferric chloride addition for coagulation, sedimentation tank, thickener, anaerobic digestion (AD) reactor, centrifuge, and buffer tanks. The buffer tanks control the inflow rate since the full-scale SBR did not receive a continuous inflow. Hence, the advanced model environment was developed based on the real WWTP at Knarrdalstrand including the SBR, centrifuge and thickener (Fig. 3).

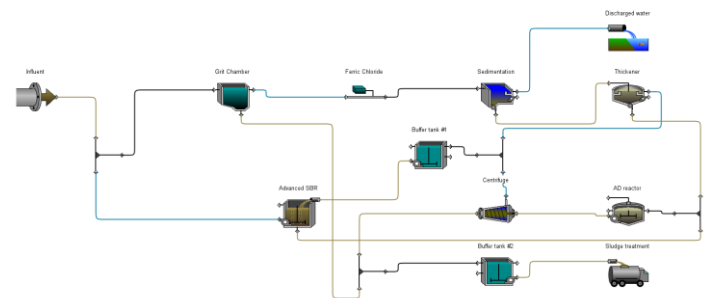


Fig. 3. The model environment of the advanced SBR model in GPS-X with all physical, chemical and biological process unit objects in the main and side streams.

### 2.3 Equation for nitrogen removal efficiency

The nitrogen removal efficiency in the advanced SBR model was calculated based on the total nitrogen mass balance. It assumed nitrogen is only removed as nitrogen gas through denitrification not other biological processes such as anaerobic ammonium oxidation (ANAMMOX). Therefore, the nitrogen removal efficiency was calculated using equations (1)-(3) (Pathak et al., 2022).

Reject water from AD is commonly characterized for high ammonium nitrogen concentration ( $568 \pm 76.7$  mg/L) due to mineralization of organic nitrogen, protein degradation, location of nitrification and concentration effect during dewatering. In equation 2 ammonium nitrogen considered as the primary nitrogen species entering to the SBR. In the SBR, nitrification is the major process for the biological oxidation of ammonia into nitrite and nitrate. Besides the ammonium nitrogen, nitrite nitrogen and nitrate nitrogen considered in equation 3 are major nitrogen species in the SBR effluent (Noutsopoulos et al., 2018;).

$$\text{Nitrogen Removal Efficiency (\%)} = \frac{C_{N_{in}} - C_{N_{out}}}{C_{N_{in}}} \quad (1)$$

Where:

$$C_{N_{in}} = C_{NH_4^+_{in}} \quad (2)$$

$$C_{N_{out}} = C_{NH_4^+_{out}} + C_{NO_3^-_{out}} + C_{NO_2^-_{out}} \quad (3)$$

Table 1. The influent reject water characteristics used for both simple and advanced SBR model and simulation.

Parameter	Default value	Measured value
Total COD	430 gCOD/m <sup>3</sup>	2500 gCOD/m <sup>3</sup>
Ammonia nitrogen	25 gN/m <sup>3</sup>	600 gN/m <sup>3</sup>
TKN*	40 gN/m <sup>3</sup>	685 gN/m <sup>3</sup>
Total phosphorus	10 gP/m <sup>3</sup>	13 gP/m <sup>3</sup>
pH	7	7.7
Temperature	20 °C	28 °C

\*TKN is total Kjeldahl nitrogen which includes organic nitrogen and total ammonia nitrogen

### 3. RESULTS AND DISCUSSION

#### 3.1 Simulating the Simple SBR model

The simple SBR model was used to simulate ammonia removal from the reject water with influent characteristics described in table 1. Figure 4 shows the SBR influent and effluent flow balance for three cycles.

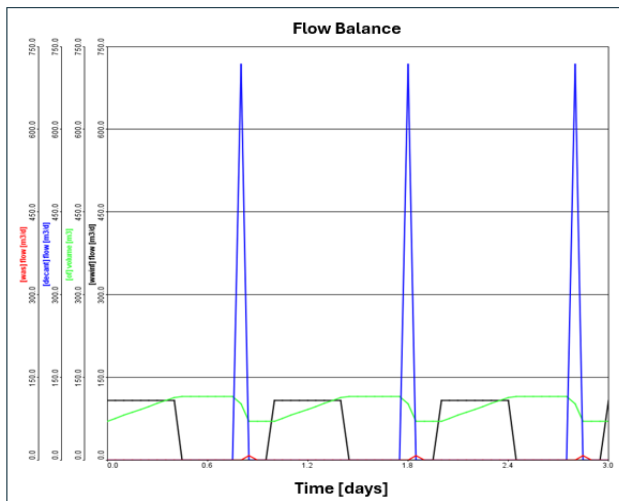


Fig. 4. The simulation environment in GPS-X with the influent and effluent flow balance of the simple SBR model. The plotted black line is the influent flow, the blue plotted line being the effluent (decant) flow and the red plotted line is the sludge waste flow. The green line is the hydraulic volume of the SRR.

The 30-days simulations of the simple SBR model have shown that the average effluent concentrations ammonia nitrogen concentration was 228 mgN/L and the average ammonia removal efficiency found was 62% (Fig. 5). The accumulation of nitrite in the effluent (Fig. 6) could be due to the partial

nitrification (PN) where the ammonium oxidizing bacteria (AOB) are dominant over the nitrite oxidizing bacteria (NOB) and the nitrite concentration increases (Duan et al., 2020). The median growth rates for AOB and NOB at 20 °C are 0.74 and 0.65 d<sup>-1</sup>, respectively. PN is common in SBR as well as operational parameters such as DO, temperature and pH significantly favors PN in SBR (Liu et al., 2020).

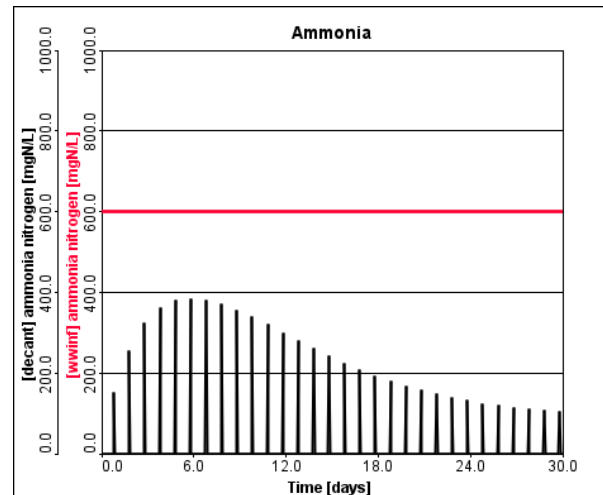


Fig. 5. Simulation of ammonia removal in SBR. The constant influent concentration of ammonia (red line) and the effluent ammonia concentration (black). Simulation was done at DO setpoint of 2.0 mg/L.

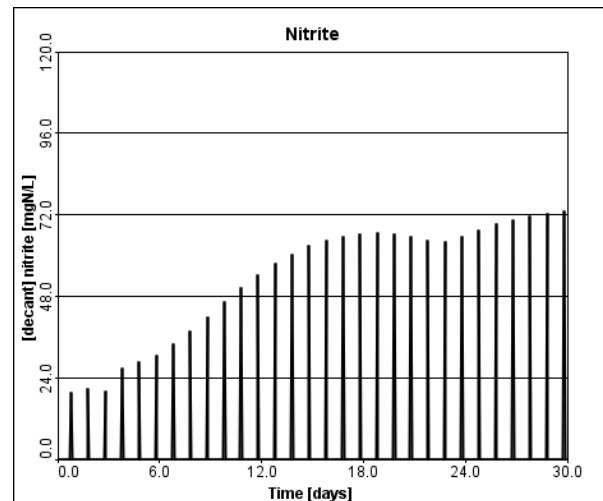


Fig. 6. Simulation ammonia removal where the nitrite [mgN/L] concentration in effluent. Simulation was done at DO setpoint of 2.0 mg/L.

#### 3.2 Simulating the advanced SBR model

The flow balance in the SBR was setup in the simulation environment in such a way that the advanced SBR model was operating according to the cycle settings of fifteen phases different in duration (minutes), mixing and DO (mgO<sub>2</sub>/L) concentration were setup. Figure 7 depicts the influent, effluent and sludge wasting flow balance simulated for 3 cycles.

The advanced SBR model was simulated for 30-days with influent ammonia concentration of 600 mgN/L. The average concentration effluent ammonia was 6 mgN/L with ammonia removal efficiency above 90% for all DO concentration



setpoints simulated (Fig. 8). Furthermore, the nitrite concentrations in the effluent reach the highest concentration during the third cycle at 222 mgN/L and then after it was decreased and stabilized (Fig. 9).

The advanced SBR model was more accurate replication of the real SBR at Knarrdalstrand WWTP. Where in the model development in the flow balance was successfully set up in the simulation environment to match with the SBR at Knarrdalstrand WWTP with all phases of the cycle. The real plant has an average ammonia removal efficiency of approximately 85%. The simulation result of the advanced SBR model predicted higher with a small margin than the treatment efficiency of the real SBR at Knarrdalstrand WWTP. This was expected when the simulated environment has been considered without disturbance from external factors. However, the advanced SBR model can be improved to the real reactor by experimentally determining the physical characteristics of the reject wastewater and by calibration of the model key process parameters (Sadri Moghaddam and Pirali, 2021).

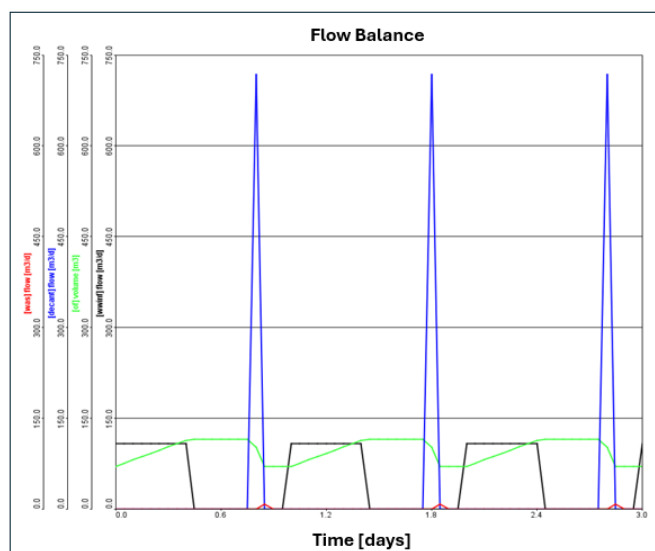


Fig. 7. The flow balance of the advanced SBR model. The black line is the influent flow, the blue line is the decant flow and the red is the waste flow. The green plotted line shows the hydraulic volume of the reactor.

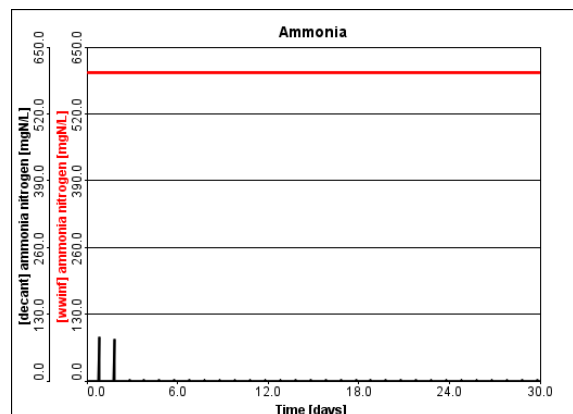


Fig. 8. Simulation of ammonia removal in advanced SBR model. The constant influent concentration of ammonia (red line) and the effluent ammonia concentration (black). Simulation was done at DO setpoint of 2.0 mg/L.

The presence of high concentration of nitrite in the effluent has shown that there was partial nitrification. This indicates that not all of the nitrite has been converted to nitrate where the dominant species in the biological process were the ammonia oxidizing bacteria (AOB) (van Niel et al., 1992).

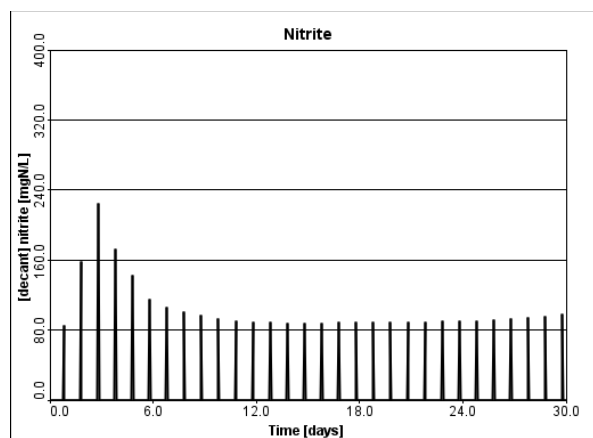


Fig. 9. Simulation of the effluent concentration of nitrite [mgN/L] in the advanced SBR model. Simulation was done at DO setpoint of 1.5 mg/L.

### 3.3 Simulation of different aeration setpoints in the advanced SBR model

To investigate the impact of dissolved oxygen (DO) concentration on nitrification and denitrification in the advanced SBR model, simulations were done using three DO setpoints. Optimizing the amount of DO required for ammonia removal presents a significant opportunity to minimize operational costs in wastewater treatment by reducing the energy expenses (Sean et al., 2020; Sid et al., 2017).

Particularly, the advanced SBR model has greater flexibility in its aeration settings than the simple model. It allowed to aerate the model more closely to the real SBR at Knarrdalstrand WWTP. Hence, it was able alternate between aerating at 2.0 mg/L DO during the aeration phases and 0.6 mg/L DO during mixing phases. This improved the ammonia removal efficiency from 62% in the simple SBR model to 95% in the advanced SBR model.

The result of the advanced model was validated from the real data from the Knarrdalstrand WWTP. For instance, the ammonia concentration of influent reject water and the effluent decant flow of the real WWTP at Knarrdalstrand was 367.3 mgN/L and 52.3 mgN/L, respectively. Based on these the average ammonia removal efficiency was approximately 85% where the model predicted 10% higher efficiency. Furthermore, the advanced SBR model can be improved to the real reactor by experimentally determining more influent characteristics of the reject and thickener water as well as parameters related with the physical characteristics of the real reactor.

Moreover, the average ammonia removal efficiency of the advanced SBR model simulation at 1.5 mg/L and 1.0 mg/L DO setpoints while maintaining 0.6 mg/L DO setpoint during mixing was 94% and 90%, respectively. Although these are averaged values of the ammonia removal efficiency, the overall trend for all these DO scenarios was reduced efficiency

in the first few cycles before it eventually stabilized at higher efficiencies. Our simulation shows that the main difference observed between the model running at the different DO setpoints were the amount of time the model needs to stabilize. The simulation at 2.0 mg/L DO had the lowest time needed for the model to stabilize with only two cycles. Whereas the simulation at 1.5 mg/L DO and 1 mg/L DO have taken three and five cycles to stabilize, respectively. Study reported that different DO concentrations have shown effect on the long-term stability of partial nitrification process at room temperature. Where AOB activity was significantly higher than NOB activity at DO of 2.5 mg/L (Cui et al., 2020). Table 2 summarizes the simulation results of the three DO setpoints in the advanced SBR model.

Table 2 The simulation result in the three DO simulated with the advanced SBR model. The average ammonia concentration in the effluent, average ammonia removal efficiency, the average COD concentration in the effluent and the average COD removal efficiency for all scenarios.

DO setpoint (mg/L)	Av. Effluent ammonia (mg/L)	Av. Ammonia removal efficiency (%)	Av. Effluent COD (mgCOD/L)	Av. COD removal efficiency
2.0	5.9	95	123	95
1.5	9.9	94	123	95
1.0	19.9	91	123	95

When compared with the influent ammonia concentration of 600 mgN/L, the ammonia concentration in the effluent were significantly lower in all DO setpoint scenarios. Moreover, the presence of high nitrite concentration than nitrate in the effluent has shown there was partial nitrification. The simulation at 2.0 mg/L DO had an effluent with high concentration of nitrite and with small concentration of nitrate. The average concentration of nitrite and nitrate were 102 mgN/L and 23 mgN/L, respectively. Hence, this indicates that not all of the nitrite has been being converted to nitrate through nitrification where the process was partial nitrification (Duan et al., 2020). Partial nitrification is common in SBR reactors when operation parameter such as DO, pH and temperature favors the process (Liu et al., 2020). Moreover, in the partial nitrification the dominant species in process are AOB that has higher bacterial growth rate (Liu et al., 2020; van Niel et al., 1992). The growth of NOB was decreasing with each cycle, while the growth AOB as well as the denitrifiers remained stable and dominated in the process. Studies shows that the NOB suppression occurring due to nitrite competition between NOB and denitrifiers instead of oxygen competition (Xu et al., 2021).

The analysis of nitrogen removal efficiency through denitrification based on the nitrogen mass balance (equation 1-3) as well as the simulation result of nitrification and denitrification rates (Fig. 10) strongly confirm that denitrification process has occurred in the SBR. From the nitrogen mass balance analysis, the nitrogen removal efficiency through denitrification in the advanced SBR model simulated with 2.0 mg/L DO was 72%. Which means that 72% of the influent ammonia across the 30-day simulated period

has been converted to nitrogen gas ( $N_2$ ) through the denitrification process. Comparatively, simulations conducted with 1.5 and 1.0 mg/L DO setpoints had the nitrogen removal efficiency of 75% and 72% through denitrification, respectively. The results were not significantly different, although simulations with the 1.5 mg/L DO setpoint had higher amount of nitrogen removed through denitrification. Since denitrification is anoxic process it is known that denitrifying bacteria (heterotrophs) thrive best in anoxic conditions (Song et al., 2021).

Moreover, simultaneous nitrification and denitrification (SND) is an advantageous bioprocess that allows the complete removal of ammonia nitrogen (Di Capua et al., 2022; Janka et al., 2022). From the nitrogen mass balance study and the simulation result it can be concluded that nitrification and denitrification process occurred in the Knardalstrand WWTP SBR and the process is the SND process.

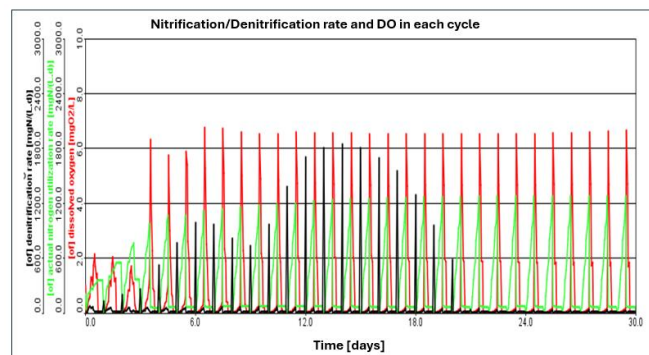


Fig. 10. The simulation of denitrification and nitrification rates for 30-days simulation period in the advanced SBR model at DO 2.0 mg/L

Hence, in the SND process, there is a huge possibility to reduce the carbon and energy consumption with a simultaneous removal of both nitrogen and phosphorus. In general, SND is cost-effective due to low DO and energy requirements (Di Capua et al., 2022; James and Vijayanandan, 2023).

### 3.4 Energy cost reduction

GPS-X allow for wastewater treatment plant operational cost estimation in the model and simulation environment (Sadri Moghaddam and Pirali, 2021; Sean et al., 2020; Sid et al., 2017). Therefore, simulation was done at different DO setpoints to investigate the energy requirement of the SBR model. In full-scale wastewater treatment plant, the major cost of the treatment process is aeration, which occupies 49% the operational cost (Pryce et al., 2022). In our simulation, the cost estimation was based on the use of blower only. The study found that the cumulative energy required for the blower running at 2 mg/L DO setpoint over 30 days simulation period was 1641 kWh. In the model the energy cost of 1 kWh was assumed 0.1\$ which resulted in the cumulative energy cost for the SBR blower was 164.1\$.

When further simulations were conducted with the DO setpoints 1.5 mg/L and 1 mg/L, respectively. The cumulative energy required for the blower operating at 1.5 mg/L DO setpoint for 30 days simulation period was 1512 kWh. With

the same energy cost the cumulative energy cost for the blower has reduced to 151.2\$ which was 7.8 % energy cost reduction. For the blower operating at 1 mg/L DO setpoint for 30 days simulation period, the cumulative energy required was 1378 kWh which has reduced the energy cost to 137.8\$ which was 16% energy cost reduction.

Hence, the energy cost reduction simulated at different DO setpoint has shown that there is a substantial opportunity to reduce the energy cost of full-scale SBR plant. The overall operation cost can be minimized by optimizing the amount of DO used for the SBR process, while maintaining the treatment efficiency. Our simulation study showed that switching the blower operation from 2 mg/L to 1.5 mg/L DO saves approximately 12% energy cost, yearly. However, the choice should be compensated with the process efficiency in removing nitrogen and organics constituents in the wastewater. For instance in similar study optimization of the SBR model process parameters such as optimum air flow into the aeration tank saved 91.5% of energy in the process that enable decision makers for the best course of action (Wondim et al., 2023).

#### 4. CONCLUSIONS

The biological process in SBR can be model and simulated using the GPS-X software. In this study SBR models have been developed for ammonia removal of the reject water at Knarrdalstrand WWTP. The advanced SBR model in GPS-X was found to be the most robust and efficient model that predicted the actual process condition of the real SBR at treatment plant. However, the advanced SBR model with the complete wastewater treatment process units needs more validation work to improve the model accuracy. Even though, the SBR model environment developed in GPS-X functions properly regarding aeration cycle and phase operations.

#### REFERENCES

- Azeez, G.K., Al-Jaberi, F.Y., Ahmed, S.A. and Hussain, A.A. (2023). Sequencing batch reactor (SBR) technology in wastewater treatment: A mini-review. *AIP Conference Proceedings*, 2806(1), 030014.
- Bhagowati, B. and Ahamad, K.U. (2019). A review on lake eutrophication dynamics and recent developments in lake modeling. *Ecology & Hydrobiology*, 19(1), 155-166.
- Cao, J., Yang, E., Xu, C., Zhang, T., Xu, R., Fu, B., Feng, Q., Fang, F. and Luo, J. (2021). Model-based strategy for nitrogen removal enhancement in full-scale wastewater treatment plants by GPS-X integrated with response surface methodology. *Science of The Total Environment*, 769, 144851.
- Cui, B., Yang, Q., Liu, X., Huang, S., Yang, Y. and Liu, Z. (2020). The effect of dissolved oxygen concentration on long-term stability of partial nitrification process. *Journal of Environmental Sciences*, 90, 343-351.
- Di Capua, F., Iannacone, F., Sabba, F. and Esposito, G. (2022). Simultaneous nitrification–denitrification in biofilm systems for wastewater treatment: Key factors, potential routes, and engineered applications. *Bioresource Technology*, 361, 127702.
- Dolatshah, M., Asadi, A., Gholami, F. and Nazari, S. (2024). Development and modeling of an integrated fixed-film activated sludge (IFAS) system for simultaneous nitrogen and carbon removal from an industrial estate wastewater. *Biotechnology Reports*, 41, e00831.
- Duan, Y., Liu, Y., Zhang, M., Li, Y., Zhu, W., Hao, M. and Ma, S. (2020). Start-up and operational performance of the partial nitrification process in a sequencing batch reactor (SBR) coupled with a micro-aeration system. *Bioresource Technology*, 296, 122311.
- Fernandes, H., Jungles, M.K., Hoffmann, H., Antonio, R.V. and Costa, R.H.R. (2013). Full-scale sequencing batch reactor (SBR) for domestic wastewater: Performance and diversity of microbial communities. *Bioresource Technology*, 132, 262-268.
- Irvine, R.L. and Ketchum, L.H. (2004). "The sequencing batch reactor and batch operation for the optimal treatment of wastewater.". *Afr. J. Biotechnol.*, 5(12), 1167–1179.
- James, S.N. and Vijayanandan, A. (2023). Recent advances in simultaneous nitrification and denitrification for

The advanced SBR model simulation with a DO setpoint of 2.0 mg/L has shown the highest average ammonia removal efficiency while the efficiency decreased slightly as the DO setpoint decreased. The simulated three DO setpoint scenarios and simulation of nitrification and denitrification in the advanced SBR model has shown that DO at 1.5 mg/L had the higher nitrogen removal efficiency through denitrification.

Generally, the model validation shows that the advanced SBR model was more effective in ammonia removal with a little higher than the ammonia removal efficiency of the real full-scale SBR plant at Knarrdalstrand. Hence, further research on sufficient data on reject waster physical and biochemical characteristics and accurate assumptions of the physical characteristics of the real SBR is vital to make the model more representative and robust in the predictive capability. Furthermore, sensitivity analysis of model parameters with different operational settings i.e. dissolved oxygen levels, pH, sludge retention time, and temperature are needed. Further work is also needed to optimize energy consumption.

#### ACKNOWLEDGEMENTS

This research project was supported by the financial support of Norwegian Directorate for Higher Education and Skills, University of South-Eastern Norway [Grant number WASH4ONEHEALTH/NORPART-2021/10070]. The full-scale SBR plant operation data and wastewater samples were accessed from Knarrdalstrand wastewater treatment plant, Porsgrunn. All authors acknowledge the contributions from the staff, especially Katrine Willer Strand and Christian Halvorsen.

- nitrogen and micropollutant removal: a review. *Biodegradation*, 34(2), 103-123.
- Janka, E., Pathak, S., Rasti, A., Gyawali, S. and Wang, S. (2022). Simultaneous Heterotrophic Nitrification and Aerobic Denitrification of Water after Sludge Dewatering in Two Sequential Moving Bed Biofilm Reactors (MBBR). *International Journal of Environmental Research and Public Health*, 19(3), 1841.
- Kesari, K.K., Soni, R., Jamal, Q.M.S., Tripathi, P., Lal, J.A., Jha, N.K., Siddiqui, M.H., Kumar, P., Tripathi, V. and Janne R. (2021). Wastewater Treatment and Reuse: a Review of its Applications and Health Implications. *Water, Air, & Soil Pollution*, 232(5), 208.
- Łagód, G., Piotrowicz, A., Cel, W., Zaburko, J. and Drewnowski, J. (2019). Modeling sequencing batch reactor operational conditions depending on oxygen concentration. *IOP Conference Series: Materials Science and Engineering*, 710(1), 012026.
- Liu, X., Kim, M., Nakhla, G., Andalib, M. and Fang, Y. (2020). Partial nitrification-reactor configurations, and operational conditions: Performance analysis. *Journal of Environmental Chemical Engineering*, 8(4), 103984.
- Man, Y., Shen, W., Chen, X., Long, Z. and Pons, M.-N. (2017). Modeling and simulation of the industrial sequencing batch reactor wastewater treatment process for cleaner production in pulp and paper mills. *Journal of Cleaner Production*, 167, 643-652.
- Naidoo, S. and Olaniran, A.O. (2013). Treated wastewater effluent as a source of microbial pollution of surface water resources. *Int J Environ Res Public Health*, 11(1), 249-70.
- Pathak, S., Wang, S. and Janka, E. (2022). Achieving Partial Nitritation in Anammox Start-Up Environment, *Water*.
- Pryce, D., Kapelan, Z. and Memon, F.A. (2022). Economic evaluation of a small wastewater treatment plant under different design and operation scenarios by life cycle costing. *Development Engineering*, 7, 100103.
- Sadri Moghaddam, S. and Pirali, M.R. (2021). Modeling and calibration of a full-scale wastewater treatment plant using GPS-X model (A case study of Tehran). *Numerical Methods in Civil Engineering*, 5(4), 67-76.
- Sean, W.-Y., Chu, Y.-Y., Mallu, L.L., Chen, J.-G. and Liu, H.-Y. (2020). Energy consumption analysis in wastewater treatment plants using simulation and SCADA system: Case study in northern Taiwan. *Journal of Cleaner Production*, 276, 124248.
- Sid, S., Volant, A., Lesage, G. and Heran, M. (2017). Cost minimization in a full-scale conventional wastewater treatment plant: associated costs of biological energy consumption versus sludge production. *Water Sci Technol*, 76(9-10), 2473-2481.
- Singh, B.J., Chakraborty, A. and Sehgal, R. (2023). A systematic review of industrial wastewater management: Evaluating challenges and enablers. *J Environ Manage*, 348, 119230.
- Song, T., Zhang, X., Li, Ji., Wu, X., Feng, H. and Dong, W. (2021). A review of research progress of heterotrophic nitrification and aerobic denitrification microorganisms (HNADMs). *Science of The Total Environment*, 801, 149319.
- van Niel, E.W., Braber, K.J., Robertson, L.A. and Kuenen, J.G. (1992). Heterotrophic nitrification and aerobic denitrification in *Alcaligenes faecalis* strain TUD. *Antonie Van Leeuwenhoek*, 62(3), 231-7.
- Wondim, T.T., Dzwauro, R.B., Aklog, D., Janka, E. and Samarakoon, G. (2023). Enhancing Textile Wastewater Treatment Performance: Optimization and Troubleshooting (Decision Support) via GPS-X Model, *Processes*.
- Xu, G., Zhang, Z. and Gao, F. (2021). Effect of COD/N ratios and DO concentrations on the NOB suppression in a multi-cycle SBR. *Journal of Environmental Chemical Engineering*, 9(4), 105735.
- Zhou, M., Gong, J., Yang, C. and Pu, W. (2013). Simulation of the performance of aerobic granular sludge SBR using modified ASM3 model. *Bioresource Technology*, 127, 473-481.

## Kinetic modelling and simulation of bioanode and biocathode in a bioelectrochemical cell for carbon dioxide reduction

Vafa Ahmadi. Nabin Aryal

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Porsgrunn, Norway (E-mail: [vafa.ahmadi@usn.no](mailto:vafa.ahmadi@usn.no), [nabin.aryal@usn.no](mailto:nabin.aryal@usn.no))*

**Abstract:** Bioelectrochemical systems (BESs) have garnered extensive research attention for their biosynthesis and environmental remediation applications. One of the challenges to upscaling BES for carbon dioxide (CO<sub>2</sub>) methanation is energy-efficient process development. Investigations are ongoing to determine the relationship between the yield of electroactive microorganisms, the key candidates for electrochemical reactions with external electricity input. Consequently, simulating processes, particularly with biocathode for biosynthesis and bioanode for remediation, gives crucial insights for designing efficient BESs. The framework for establishing Nernst-Monod equations for modelling BES, starts from bioanode, where anode respiring bacteria (ARB) oxidate organic carbon compounds to CO<sub>2</sub>, and generate the proton (H<sup>+</sup>). In this work, kinetic modelling was applied to calculate the biomass yield of ARBs corresponding to the applied anodic voltage. The generated CO<sub>2</sub> and H<sup>+</sup> from the anode determined the biomass yield of electroactive methanogens and acetogens on the cathode. Two biofilm models were established for anodic and cathodic biofilm growth in Aquasim simulation tool. Results showed that the concentration of organic carbon compound (acetate) available for ARB, had a significant impact on the biofilm thickness and biomass concentration on the biofilm, especially at + 0.3 V. The optimum anode voltage which released the highest CO<sub>2</sub> and H<sup>+</sup>, was + 0.3 V. The anodic and cathodic biofilm thickness reached respectively 3 mm and 55 μm at + 0.3V and 10 g.L<sup>-1</sup> acetate input to the anode chamber. Moreover, methanogens surpassed acetogens on the biocathode for CO<sub>2</sub> reduction to methane rather than acetate. In addition, acetate consumption rate by ARB at anode was remarkably faster than acetate production at cathode.

**Keywords:** Carbon dioxide, Anode respiring bacteria, Methanogens, Acetogens, Biocathode, Bioanode, Methane and Acetate.

### 1. INTRODUCTION

Biomethane production from anaerobic digestion (AD) is one of the strong alternatives to fossil fuels. However, AD may face limitations due to multiple process-related factors, in particular high pH fluctuation, low productivity, or inhibition effect due to volatile fatty acids (VFA). Understanding the limitations of the process and the factors affecting microbial growth is essential for increasing biomethane production. Bioelectrochemical systems (BESs) are one of the emerging technologies that benefit from biofilms with electroactive microbes that consume electricity. BES is a promising technology to overcome the limitations of AD process. The electricity input aids the microbes and enhances the breakdown of organic matter. Meanwhile, it can stop unwanted side reactions which limit production. Therefore, the purity of target products such as methane (CH<sub>4</sub>), acetic acid or other valuable organic carbon compounds can be increased (Aryal et al., 2020). The methods currently available for the electrochemical reduction of CO<sub>2</sub> need external overpotentials due to high energy demand, to overcome ohmic losses. If enough electric potential is provided, bacteria and archaea can overcome potential losses in BES, where CO<sub>2</sub> can be reduced to valuable energy carriers

such as CH<sub>4</sub> or other biochemicals (Rittmann and McCarty 2020).

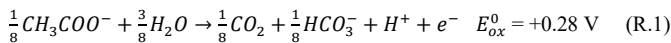
The anode potential (oxidation voltage) can be correlated with the metabolic activities of biocatalysts and the corresponding energy dissipation. In microbial electrochemical cells (MECs), the applied electricity can help microorganisms obtain electromotive forces for better growth and maintenance (Torres et al., 2008). In MEC, a dense biofilm on anode oxidizes the organic and inorganic molecules if enough energy is available (Rabaey and Verstraete, 2005). The external energy supports and enhances the microbial community of anode-oxidizing bacteria (ARB) due to the microbial selectivity process. Besides, the activity of methanogenic microbial biomass at cathode will be enhanced. Autotrophic methanogens will become more abundant in the biofilm matrix, and a mixture of autotrophic and heterotrophic microbial biomass communities will also increase in the suspended media of the reactor.

Generally, direct electron transfer (DET) or mediated electron transfer is well accepted when abiotic electrodes and electroactive microbes interact with each other. In mediated electron transfer theories, the kinetics of electron transfer starts from ARB. These bacteria oxidize organic compounds, which

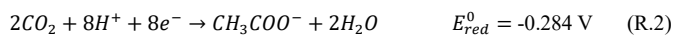


provide electrons for cathode. ARBs transfer electrons to the anode via anodic respiration. Like methanogens, ARBs have a constrained range of electron donors. The most popular and beneficial electron donor by far is acetate. Certain ARBs can also use hydrogen ( $H_2$ ), formate, ethanol, propionate, glucose, and other sugars as electron donors. Anode respiration is a novel and distinct kind of catabolism. A general electron donor is used, but the electron acceptor is substituted with a solid acceptor, completing the catabolic reaction. *Geobacter sulfurreducens* and *Shewanella* species are the two most well-known ARBs. However, it is yet unclear how ARBs accomplish extracellular electron transfer (EET) biochemically (Rittmann and McCarty 2020).

BES reactor configuration can be divided into multi-chamber and single-chamber reactors. In dual chamber reactors, the anode and cathode are separated by an ion exchange membrane, preventing possible anodic/cathodic reversible reactions. However, in membrane-less single-chamber reactors, the anode and cathode are in the same chamber (Aryal et al., 2022). Therefore, in the mixed microbial environment of the reactor, it is possible that an organic matter, such as acetate, is oxidized at the anode via electrochemical reactions or by ARB, and then reproduced at cathode by acetogenic microbes (May et al., 2016; Yates et al., 2017). The oxidation of acetate by ARBs at the anode can be faster than the reproduction of acetate at the cathode. Additionally, the competition between methanogens and acetogens in the cathode chamber is another important area of study. In dual-chamber reactors, it is possible to measure the acetate consumption rate by ARBs and the acetate production rate by autotrophic acetogens. The half-reaction of acetate oxidation (R.1) shows that acetate is oxidized to  $CO_2$ ,  $HCO_3^-$ , and  $H^+$  (Rittmann and McCarty 2020):



The reproduction of acetate can occur at the cathode via (R.2).



Methanogens at the cathode produce  $CH_4$  via (R.3) through the DET mechanism:



This work simulates ARB growth based on acetate consumption at different anodic voltages. The  $H^+$  and  $CO_2$  released from acetate oxidation by electroactive ARBs will be used to calculate the growth of cathodic biofilm by autotrophic methanogenic species for  $CH_4$  production. Furthermore, the reproduction of acetate by acetogenic species has been studied to compare the acetate consumption rate with the acetate reproduction rate.

## 2. METHODOLOGY

The electron acceptor of the ARBs has no concentration because it is a solid anode. Microorganisms are considered in direct contact with the electrode (Rittmann and McCarty 2020). ARBs utilize the electron donor molecules, which in

this work is acetate. The activity of the electroactive cathodic biofilm is limited by the number of electrons and  $H^+$  flowing to the cathode for  $CH_4$  and acetate production (Fig. 1). The process scheme involves wastewater containing acetate and biomass flowing to the anode with 1 day hydraulic retention time (HRT) (Ahmadi and Dinamarca, 2022). Additionally, biomass-containing wastewater enters the cathode chamber with 1 day HRT. In the BES reactor illustrated in Fig. 1, voltage is applied to the anode. The anode potential creates an electron pull on the cathode, which absorbs the electrons flowing from bioanode to the biocathode.

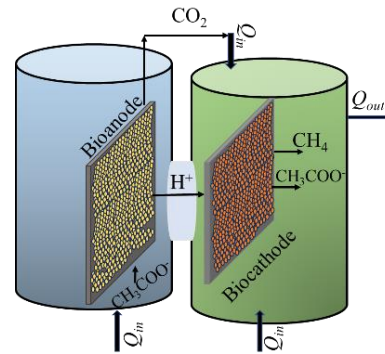


Fig. 1. Schematic diagram of the  $CO_2$  reduction with a dual-chamber BES reactor separated by ion exchange membrane. The ARB bioanode oxidizes acetate. Then  $H^+$  and  $CO_2$  flow to biocathode for  $CH_4$  and acetate production.

In a BES, the anode acts as the electron acceptor instead of an organic electron acceptor. The cathode serves as the electron donor instead of an organic electron donor. Therefore, it can be hypothesized that in a BES system, two complete biological redox reactions take place, as illustrated in Fig. 2.

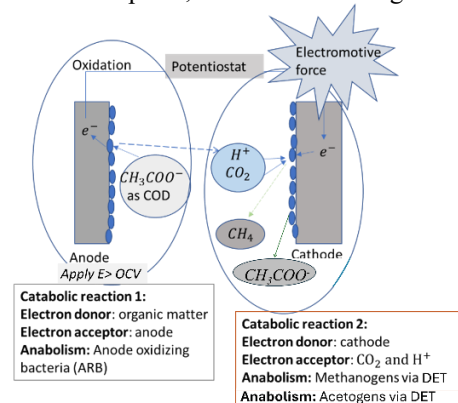


Fig. 2. Graphical view of catabolism and anabolism in a BES reactor. Anabolism refers to cell growth, while catabolism refers to product formation. OCV is the open circuit voltage of anode; any voltage higher than this will be relatively positive.

### 2.1 Simulation parameters and inputs

Tab. 1 provides all the parameters, expressions, and inputs for simulating the model. The biomass formula is considered to be  $CH_{1.8}O_{0.5}N_{0.2}$  (Heijnen and Kleerebezem, 1999). Aquasim version 2.1 (Reichert, 1998) was used to simulate the kinetics of the bioanode and biocathode for  $CO_2$  reduction. Furthermore, the acetate input ( $S_{ac}$ ) to the anode chamber is assumed to be 1, 5, and 10  $g \cdot L^{-1}$ . The molar concentrations of acetate are 16.66, 83.33, and 160.66  $mol \cdot m^{-3}$ , respectively. The

CO<sub>2</sub> and H<sup>+</sup> inputs to the cathode are obtained after simulating the anodic biofilm kinetics. Then, the CO<sub>2</sub> and H<sup>+</sup> produced during steady-state anodic oxidation will be used for CH<sub>4</sub> and acetate production at cathode.

## 2.2 Model assumptions

-The model consists of two biofilms: one for ARB attachment in the anode chamber and the second, the cathodic biofilm (or biocathode), for the formation of methanogenic and acetogenic biofilms.

- The substrate is acetate for organic-oxidizing bacteria, such as *Acetobacter acetae* (ARB1), and sulfate-reducing bacteria (SRB), assumed as ARB2, that grow on acetic acid (Tremblay et al., 2019).

- The substrates for biocathode are CO<sub>2</sub> and H<sup>+</sup> from acetate oxidation at the anode. *Sporomusa ovata* can be considered as acetogens, and *methanobacterium* as methanogens on the cathode (Zhang et al., 2017).

- pH is considered 7 and the pressure is 1 atm.

- CO<sub>2</sub> and HCO<sub>3</sub><sup>-</sup> are in equilibrium, so the input to the cathode is considered to be CO<sub>2</sub>.

- The biofilm decay is accounted by considering a detachment velocity in the system.

- BES oxidation and reduction are specifically emphasized.

- No electrons are lost in other reactions, and potential losses in the system are ignored.

- The impact of electrode material and the self-potential of the electrodes are not considered.

## 2.3 Model expressions

The simulation in this work started at the anode, where ARBs oxidize the chemical oxygen demand (COD) as acetate. The Nernst-Monod equation is used to calculate the ARB biomass growth rate ( $\frac{d[X_{ARB}]}{dt}$ ) and the variation in the concentration of acetate ( $\frac{d[S_{ac}]}{dt}$ ), as well as the produced CO<sub>2</sub> and H<sup>+</sup>. Equations 1 and 2 illustrate the anodic model expressions (Korth et al. 2015; Fischer 2017; Rittmann and McCarty 2020).

$$\frac{d[X_{ARB}]}{dt} = X_{ARB} \cdot \left( \mu_{X_{ARB}}^{max} \cdot \emptyset_{an} \cdot \frac{S_{ac}}{K_{S_{ac}} + S_{ac}} \cdot \frac{1}{1 + \exp\left[-\frac{(E_{app} - E_{KA})F}{RT}\right]} - k_{d,ac} \right) \quad (1)$$

$$\frac{d[S_{ac}]}{dt} = - \left( \frac{d[X_{ac}]}{dt} \right) / Y_{ac} \quad (2)$$

Where  $X_{ARB}$  is the ARB biomass concentration,  $\mu_{X_{ARB}}^{max}$  is the maximum growth rate of ARB,  $\emptyset_{an}$  is the fraction of electroactive biomass in the biofilm, assumed 0.75,  $S_{ac}$  is the acetate molar concentration,  $K_{S_{ac}}$  is the half-saturation constant of acetate,  $E_{app}$  is the applied voltage,  $E_{KA}$  is the potential at which the biomass growth rate is half of its maximum,  $k_{d,ac}$  is the ARB decay rate,  $F$  is the Faraday constant (96485.3 C. mol<sup>-1</sup>),  $R$  is the universal gas constant (8.314 J. mol<sup>-1</sup>. K<sup>-1</sup>),  $T$  is temperature (K), and  $Y_{ac}$  is the ARB biomass yield based on acetate consumption. All the terms with values are given fully in Tab.1.

In R1, the Nernst term is a critical factor for the electroactive biomass growth. Every electrode material has a natural voltage relative to its surrounding electrolyte, known as the open circuit voltage (OCV).

Table 1. Model parameters for the simulation

Parameter, unit	Value	Ref
Diffusivity of acetate, m <sup>2</sup> · d <sup>-1</sup>	9.41 · 10 <sup>-5</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of CH <sub>4</sub> , m <sup>2</sup> · d <sup>-1</sup>	1.296 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of CO <sub>2</sub> , m <sup>2</sup> · d <sup>-1</sup>	1.658 · 10 <sup>-4</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of H <sup>+</sup> , m <sup>2</sup> · d <sup>-1</sup>	8.04 · 10 <sup>-5</sup>	(Ahmadi and Dinamarca, 2022)
Diffusivity of biomass, m <sup>2</sup> · d <sup>-1</sup>	1 · 10 <sup>-7</sup>	(Ahmadi and Dinamarca, 2022)
Biomass density, mol · m <sup>-3</sup>	220	(Ahmadi and Dinamarca, 2022)
Half saturation concentration of CO <sub>2</sub> , mol · m <sup>-3</sup>	3.8	(Eddy et al., 2014)
Half saturation concentration of H <sup>+</sup> , mol · m <sup>-3</sup>	0.0002	(Eddy et al., 2014)
Half saturation concentration of acetate, mol · m <sup>-3</sup>	0.03	(Eddy et al., 2014)
Max growth rate of ARB1, d <sup>-1</sup>	12	Assumed
Max growth rate of ARB2, d <sup>-1</sup>	2.4	Assumed
Max growth rate of methanogens, d <sup>-1</sup>	2.28	Assumed
Max growth rate of acetogens, d <sup>-1</sup>	1.008	Assumed
Acetogenic growth yield	0.02	Assumed
Methanogenic growth yield	0.083	Assumed
ARB1 growth yield	0.0384	Assumed
ARB2 growth yield	0.28	Assumed
Initial biofilm thickness, m	1 · 10 <sup>-10</sup>	This work
Half maximum growth voltage of ARB1, V	-0.2	(Heijnen and Kleerebezem, 1999)
Half maximum growth voltage of ARB2, V	-0.1	(Heijnen and Kleerebezem, 1999)
Inflow, m <sup>3</sup> · d <sup>-1</sup>	0.1	This work
Biofilm reactor volume, m <sup>3</sup>	0.1	This work
Biofilm surface area, m <sup>2</sup>	10	This work
Anode voltage, V	-0.2, -0.1, 0.3	This work
Acetic acid concentration, mol · m <sup>-3</sup>	16.66, 83.33, 166.6	This work
CO <sub>2</sub> concentration, mol · m <sup>-3</sup>	Simulation results	This work
H <sup>+</sup> concentration, mol · m <sup>-3</sup>	Simulation results	This work

An applied voltage higher than OCV increases oxidation on the electrode. Therefore, the applied voltage in this work is relatively voltage with respect to OCV. The OCV of the anode is crucial in determining the applied voltage and the potential range within which  $E_{KA}$  falls.

Figure 3 shows the calculated Nernst term with respect to the anodic voltage. According to Rittmann and McCarty (2020), when  $E_{KA}$  is higher, more voltage is required to reach the

maximum Nernst term of 1. Thus, lower  $E_{kA}$  is more advantageous for the BES in terms of energy savings.

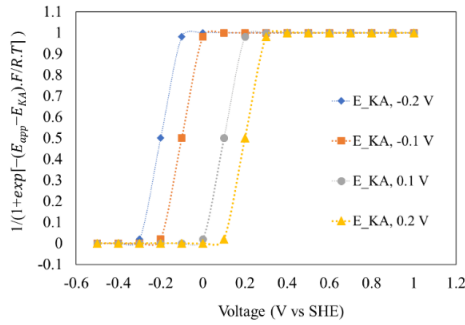


Fig. 3. Variation of the Nernst term with respect to the anode voltage at different  $E_{kA}$  values, ranging from -0.2 to +0.2 V.

The dynamic current generation ( $\frac{dj}{dt}$ ,  $A \cdot m^{-2}$ ) via ARB biomass on anode (current density,  $A \cdot m^{-2}$ ) can be obtained by (6) (Torres et al., 2008).

$$\frac{dj}{dt} = \frac{d[X_{ARB}]}{dt} \cdot \gamma \cdot L_f \cdot (f_s^0 - 1) \quad (6)$$

where,  $\gamma$  represents the equal electron production, which is 8 for acetate oxidation.  $L_f$  is the biofilm thickness (m). The term  $f_s^0$  refers to electron equivalent cell synthesis coefficient, which can be obtained by thermodynamic equations as fully described by Rittmann and McCarty (2020). At lower voltage,  $f_s^0$  is reduced. At higher voltage,  $f_s^0$  increases leading to a higher biomass population on the anode. This concept is explained in a thermodynamic evaluation of biomass growth with respect to the voltage (Ahmadi and Aryal, 2024).

In BES, voltage is applied either to either the anode or cathode. This creates a potential difference between electrodes, enabling redox reactions. Therefore, to establish the methanogenic and acetogenic biomass growth rates on the cathode ( $\frac{d[X_{CH_4}]}{dt}$  and  $\frac{d[X_{ac}]}{dt}$ , respectively), using the multiplicative Monod equation without the Nernst term is a more reasonable approach, as shown in equations (3), (4), and (5). The number of electrons and  $H^+$  from acetate oxidation is equal. Thus, when  $H^+$  is available, it indicates that enough electrons have been delivered to undergo  $CO_2$  reduction (Korth et al. 2015; Fischer 2017; Rittmann and McCarty 2020).

$$\frac{d[X_{CH_4}]}{dt} = X_{CH_4} \cdot \left( \mu_{X_{CH_4}}^{max} \cdot \emptyset_{cat} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} - k_{d,CH_4} \right) \quad (3)$$

$$\frac{d[X_{ac}]}{dt} = X_{ac} \cdot \left( \mu_{X_{ac}}^{max} \cdot \emptyset_{cat} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} - k_{d,ac} \right) \quad (4)$$

$$\frac{d[S_{CH_4}]}{dt} = \left( \frac{d[X_{CH_4}]}{dt} \right) / Y_{CH_4} \quad (5)$$

where  $X_{CH_4}$  is the methanogenic biomass content,  $\mu_{X_{CH_4}}^{max}$  is the maximum growth rate of methanogens,  $\emptyset_{cat}$  is the fraction of electroactive biomass on cathode,  $S_{CO_2}$  is the molar concentration of  $CO_2$ ,  $K_{CO_2}$  is the half-saturation constant of  $CO_2$ ,  $S_{H^+}$  is the molar concentration of  $H^+$ ,  $K_{H^+}$  is the half-saturation constant of  $H^+$ ,  $k_{d,CH_4}$  is the decay rate of methanogens,  $X_{ac}$  is the acetogenic biomass content,  $\mu_{X_{ac}}^{max}$  is the maximum growth rate of acetogens,  $k_{d,ac}$  is the decay rate

of acetogens,  $\frac{d[S_{CH_4}]}{dt}$  is the methane production rate in molar concentration,  $Y_{CH_4}$  is the methanogenic biomass yield based on  $CO_2$  and  $H^+$  consumption.

### 3. RESULTS AND DISCUSSION

#### 3.1 Concentration profiles at anode

Figure 4 shows the total acetate consumption by ARBs in the anode chamber at different concentrations. The concentration profiles show the relation between the anodic voltage and acetate consumption. When the anode voltage is equal to  $E_{kA}$  of ARB, the growth rate of ARB is half of its maximum value. As a result, it takes approximately 17 days for ARB to consume all the acetate. By increasing the voltage to -0.1 V, acetate consumption accelerates, with all the acetate being consumed after 5 days. An even higher anodic voltage results in faster acetate consumption, completely depleting the acetate in 2 days. At higher acetate concentrations, the consumption profile follows the same order, but it takes longer to deplete the total acetate at -0.2 V. However, increasing the voltage helps reduce the acetate depletion time at higher concentrations.

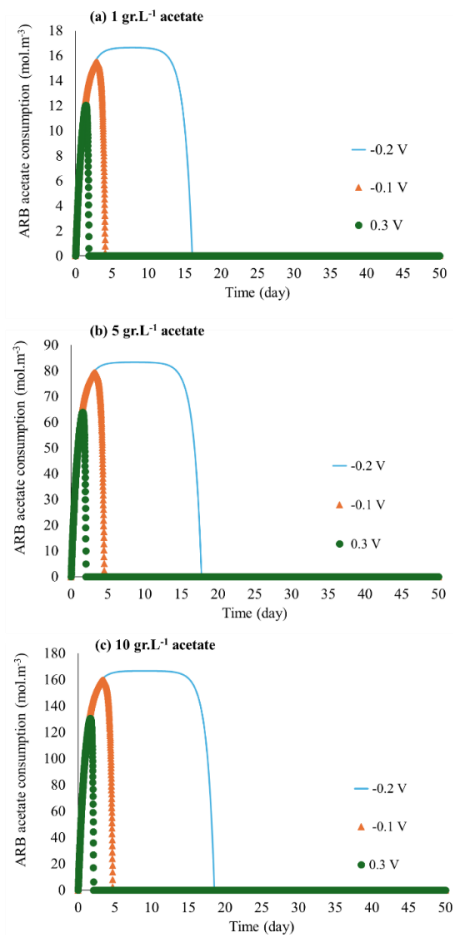


Fig. 4. Acetate consumption by ARB biofilm for different acetate concentration in anode chamber. The applied anode voltage is -0.2, -0.1 and +0.3 V.

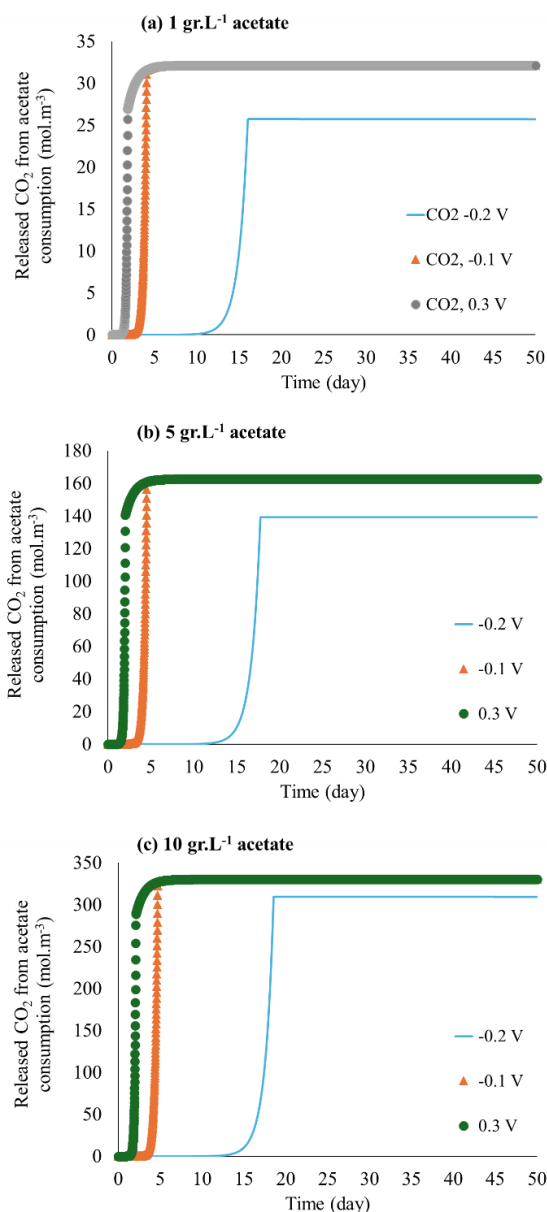


Fig. 5. The released CO<sub>2</sub> from ARB oxidation of acetate. The applied anode voltage is -0.2, -0.1 and +0.3 V.

The correspondence concentration profiles of CO<sub>2</sub> (Fig. 5) and H<sup>+</sup> (Fig. 6) reflects the same trend as at higher concentration of acetate, higher amount of CO<sub>2</sub> and H<sup>+</sup> will be released that will flow to cathode to be consumed by acetogens and methanogens. Also, the graphs show that higher anode voltage helps shortening the H<sup>+</sup> evolution.

The ARB biofilm thickness (Fig. 7) at anode is more dependent on the availability of substrate for the ARB than the applied voltage at anode. Nevertheless, +0.3 V significantly increases the biofilm thickness to 3 mm while at -0.2 V, the biofilm thickness will be 1.5 mm where the applied voltage is equal to E<sub>KA</sub> of one ARB species. At -0.1 V, the thickness reaches 2.2 mm for 10 g.L<sup>-1</sup> acetate input. The graphs show that at low substrate concentration, the effect of voltage on the biofilm thickness is not significant. For 1 g.L<sup>-1</sup> acetate input, the biofilm thickness will be 55, 57, 59 μm at -0.2 V, -0.1 V and +0.3V respectively. For 5 g.L<sup>-1</sup> acetate input, the biofilm

thickness reaches 110, 117 and 121 μm correspondingly for -0.2, -0.1 and +0.3 V.

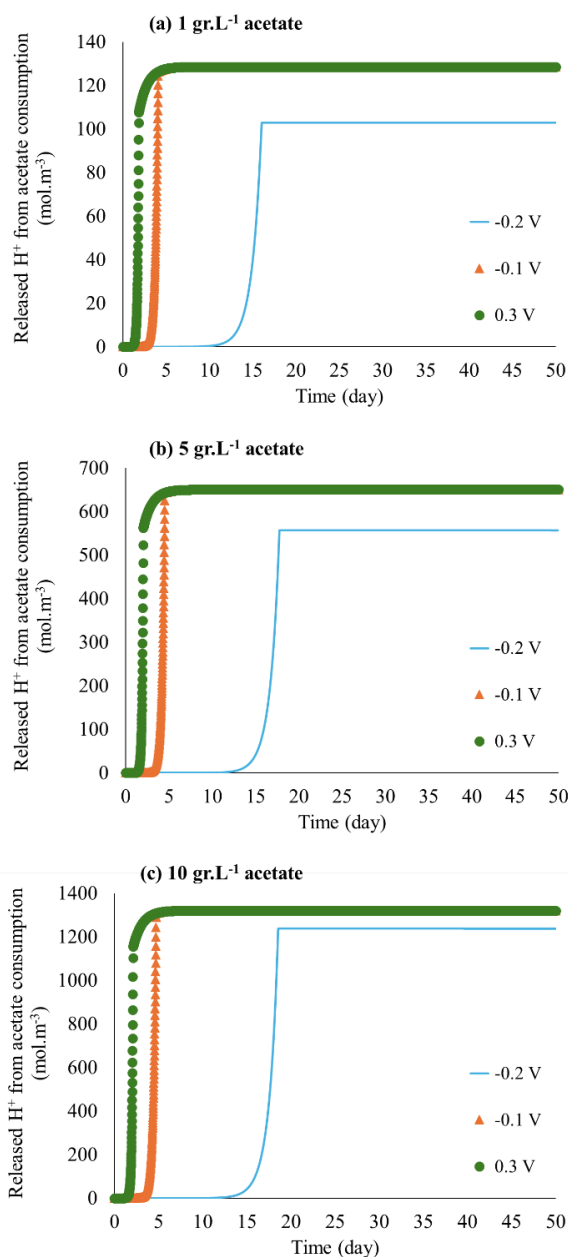


Fig. 6. The released H<sup>+</sup> from ARB oxidation of acetate. The applied anode voltage is -0.2, -0.1 and +0.3 V.

Since 10 g.L<sup>-1</sup> acetate input to anode resulted in the highest biofilm thickness, the ARB biomass concentration (Fig. 8) is provided for different voltages and 10 g.L<sup>-1</sup> acetate concentration. Here in Fig. 8 also, 10 g.L<sup>-1</sup> acetate input at higher voltage of 0.3 V, increases the biomass concentration on anode. The reason is the Nernst term in Eq.1 will be equal to 1 which gives the highest biomass concentration.



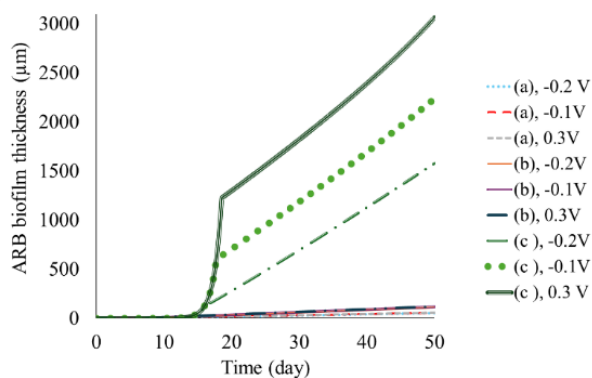


Fig. 7. ARB biofilm thickness at different acetate concentrations; (a) 1 g.L<sup>-1</sup>, (b) 5 g.L<sup>-1</sup>, (c) 10 g.L<sup>-1</sup> acetate at bioanode. The applied anode voltage is -0.2, -0.1 and +0.3 V.

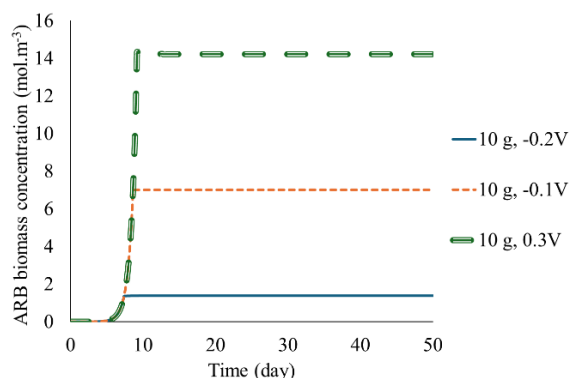


Fig. 8. The anode respiring bacteria (ARB) concentration at -0.2, -0.1, and 0.3V for 10 g.L<sup>-1</sup> acetate input.

In the current density plots in Fig. 9, the bioelectrochemical current density which can be generated directly by the ARB, is the highest at +0.3V when the acetate input is 10 g.L<sup>-1</sup> in the anode chamber. The biomass concentration is also higher (Fig. 8); therefore, microbes can break down more substrate to electrons, CO<sub>2</sub> and H<sup>+</sup>. Also, in a bioanode, at the start of the biofilm formation, the current density is very low in the micro ampere levels, and after day 10, when the biomass starts to form significantly, the current density increases faster. These simulation results fit the experimental observations in other research work. The bioelectrochemical current density is low at the starting days and increases over time as a dense biofilm is formed on the electrode surface (Li et al., 2020). Moreover, Figure 9 indicates clearly that increasing the applied voltage to BES will not increase the bioelectrochemical current density if not enough substrate is available for bio-oxidation. This is one of the problems which happens in experimental research that high voltage at a low organic matter liquid, leads to many problems such as electrode failure, or the electron generation will not be mainly due to the oxidation of organic matter. In these conditions, increasing the voltage leads to the electrochemical oxidation of the buffer compounds, ammonium, salts, minerals, or water (Sivalingam et al., 2020; Wagner et al., 2010)

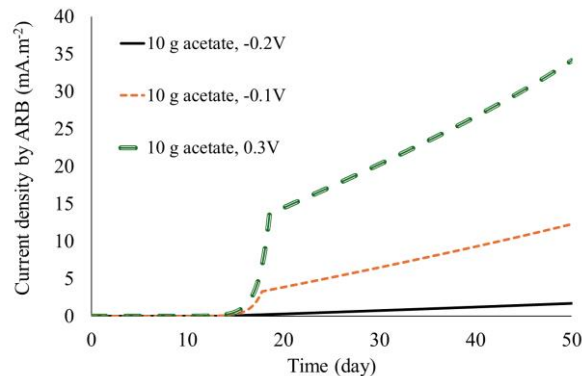


Fig. 9. Current generation by ARB from acetate consumption at 0.3V anode voltage.

### 3.2. Concentration profiles at cathode

The corresponding CO<sub>2</sub> consumption profiles via biocathode is presented in Fig. 10. The profiles show that the amount of CO<sub>2</sub> consumed by the cathodic biofilm significantly depends on the activity of ARB. In addition to the dependence on the applied anodic voltage and acetate concentration in the anodic chamber, the methanogens and acetogens on the biofilm need more days before the biofilm gets saturated by CO<sub>2</sub> and H<sup>+</sup>, and afterwards, CH<sub>4</sub> and acetate production starts. Here, is the illustration of the reason for lag phase in BES reactors to reach stable and maximum CH<sub>4</sub> production. The H<sup>+</sup> consumption profile is also presented in Fig. 11. In the CO<sub>2</sub> and H<sup>+</sup> profiles, it is depicted that the availability of H<sup>+</sup> and CO<sub>2</sub> for cathode, depends both on the applied voltage and acetate concentration. So, lack of organic matter in the anode chamber can be a more significant reason than the applied voltage. Also, the profiles shift with respect to the applied voltage again shows that at higher anode voltage, the CO<sub>2</sub> and H<sup>+</sup> will be available faster for the cathodic microbes. Therefore, the biocathode becomes stable faster at +0.3 V vs SHE compared to lower voltage.

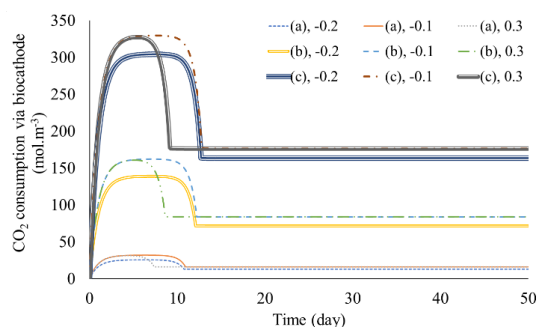


Fig. 10. CO<sub>2</sub> consumption via biocathode, in case of (a) 1 g.L<sup>-1</sup>, (b) 5 g.L<sup>-1</sup> and (c) 10 g.L<sup>-1</sup> acetate for bioanode. The applied anode voltage is -0.2, -0.1 and +0.3 V.

In the H<sup>+</sup> consumption profiles in Fig. 11, the H<sup>+</sup> will be completely consumed by the biocathode while CO<sub>2</sub> remains in the system. This is because according to the acetate oxidation reaction, two moles CO<sub>2</sub> and eight moles H<sup>+</sup> will be available in case of assuming all the CO<sub>2</sub> and H<sup>+</sup> will be available for Doieight moles H<sup>+</sup>. That is why a part of the CO<sub>2</sub> will remain unconsumed. For consumption of more CO<sub>2</sub>, H<sup>+</sup> should be provided at a higher voltage from other sources such as water, phosphate, sulphate, or ammonium (Rabaey et al., 2009).



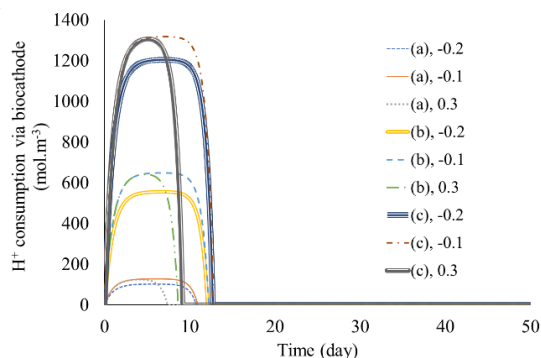


Fig. 11. H<sup>+</sup> consumption via biocathode, in case of (a) 1 g.L<sup>-1</sup>, (b) 5 g.L<sup>-1</sup> and (c) 10 g.L<sup>-1</sup> acetate for bioanode. The applied anode voltage is -0.2, -0.1 and +0.3 V.

CH<sub>4</sub> production profiles in Fig. 12 show that there is a direct relation between the acetate input to anode and the applied anodic voltage with CH<sub>4</sub> production. CH<sub>4</sub> will be higher and faster as a result of 10 g.L<sup>-1</sup> acetate and +0.3 V applied voltage at anode. It should be noted that +0.3 V is a relative value with respect to the open circuit potential (OCV) of the anode. At +0.3 V at anode, the released CO<sub>2</sub> from oxidation of 10 g.L<sup>-1</sup> acetate, the amount of CH<sub>4</sub> will be the highest according to the parameters in this work, and CH<sub>4</sub> production starts after 6 days. The startup time of CH<sub>4</sub> production is influenced by the voltage. At all acetate concentrations, +0.3 V at anode, leads to higher H<sup>+</sup> and CO<sub>2</sub> release, which reduces the startup time of the process.

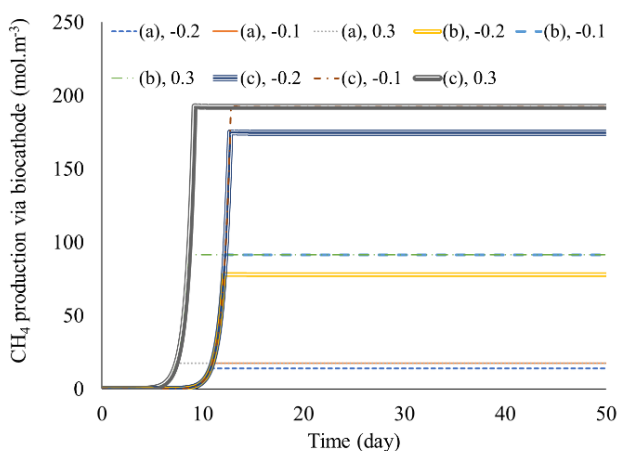


Fig. 12. CH<sub>4</sub> production via biocathode, (a) in case of 1 g.L<sup>-1</sup> acetate, (b) 5 g.L<sup>-1</sup> acetate and (c) 10 g.L<sup>-1</sup> available acetate for bioanode. The applied anode voltage is -0.2, -0.1 V and +0.3.

Figure 13 shows the amount of acetate production in the cathode is very insignificant compared to CH<sub>4</sub> production. Also, acetate production reaches a peak of 0.007 mol.m<sup>-3</sup>, then it starts to decrease until reaching 0 until day 25. For acetate production, two moles of CO<sub>2</sub> should be consumed. Another reason is the assumed yield and max growth rate of methanogens, which greatly impacts the results. However, in numerous experimental works, CH<sub>4</sub> production surpasses acetate production in a BES reactor.

The results clearly illustrate that acetate oxidation rate in Fig. 4, is much faster than acetate production in Fig. 13. Acetate production is negligible compared to acetate consumption. In

many research work, if acetate is the final product, methanogenic inhibitor is used to stop methanogens. Therefore, acetogens will not have competing species, and acetate content will increase in the system (Sivalingam et al., 2022).

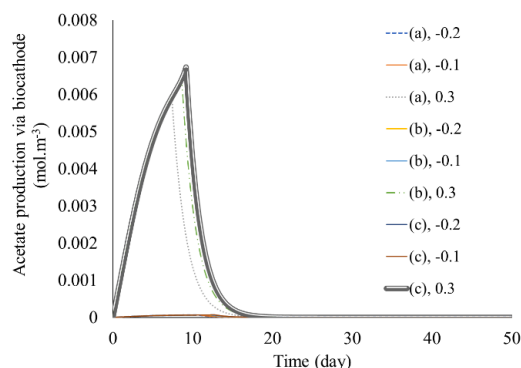


Fig. 13. Acetate production via biocathode, in case of (a) 1 g.L<sup>-1</sup>, (b) 5 g.L<sup>-1</sup> and (c) 10 g.L<sup>-1</sup> acetate input to anode. The applied anode voltage is -0.2, -0.1 and +0.3 V.

The cathodic biofilm thickness (Fig. 14) shows that the higher concentration of acetate which is oxidized at anode leads to higher CO<sub>2</sub> and H<sup>+</sup> input to cathode. Therefore, the cathodic biofilm is thicker when the acetate input to anode is 10 g.L<sup>-1</sup> because it releases higher concentration of CO<sub>2</sub> and H<sup>+</sup>. Also, +0.3 V anode voltage led to a thicker biofilm at all concentrations. Also, at +0.3 V anode voltage, the cathodic biofilm starts to grow faster compared to lower voltage. Also, since the growth of autotrophic microbes which grow on inorganic carbon is slower than organic carbon consumption, the ARB biofilm at anode at 10 g.L<sup>-1</sup> acetate input and +0.3V is 3 mm which is 5 times thicker than the cathodic biofilm 55μm. This is reasonable because organic matter is easier to consume compared to inorganic carbon source (CO<sub>2</sub>) for microbes (Abreu et al. 2022; Rittmann and McCarty 2020).

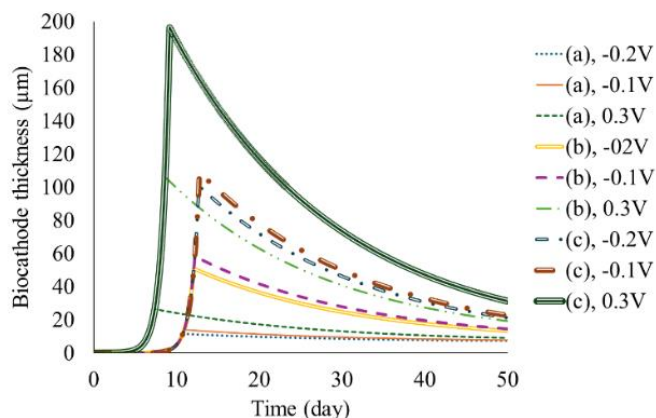


Fig. 14. The cathodic biofilm thickness at different CO<sub>2</sub> and H<sup>+</sup> concentrations resulted from acetate oxidation at anode at -0.2, -0.1 and +0.3 V. (a): 1 g.L<sup>-1</sup>, (b): 5 g.L<sup>-1</sup>, (c): 10 g.L<sup>-1</sup> acetate input to anode.

#### 4. CONCLUSIONS

This simulation work gives a critical insight on the kinetics of bioanode and biocathode in a dual chamber BES reactor. It

should be noted that the model in this work is made according to a single electron donor acetate. In a BES, other sources of electrons and  $H^+$  can take part in oxidation reactions which could be bioelectrochemical or electrochemical. Assessment of the effect of other electron donors on the kinetics of  $CH_4$  production can be considered for further work. The simulation gives a clear vision about the importance of concentration of electron donors and the relation between biomass growth and voltage. Therefore, the simulation will be used for process design and future validation of the results. The kinetic simulation helps regulating the reactor conditions to obtain the optimum biomass on anode and cathode to carry on bioelectrochemical reactions.

## REFERENCES

- Abreu, A.P., Morais, R.C., Teixeira, J.A., Nunes, J. (2022). A comparison between microalgal autotrophic growth and metabolite accumulation with heterotrophic, mixotrophic and photoheterotrophic cultivation modes. *Renewable and Sustainable Energy Reviews* 159, 112247. doi. 10.1016/j.rser.2022.112247
- Ahmadi, V., Aryal, N. (2024). Evaluation of the relationship between voltage and microbial yield in a bioelectrochemical reactor for optimization of methane production, in: 18th IWA World Conference on Anaerobic Digestion – IWA. IWA, Istanbul.
- Ahmadi, V., Dinamarca, C. (2022). Simulation of the Effect of Local Electric Potential and Substrate Concentration on  $CO_2$  Reduction via Microbial Electrosynthesis.
- Aryal, N., Ghimire, N., Bajracharya, S. (2020). Coupling of microbial electrosynthesis with anaerobic digestion for waste valorization, in: *Advances in Bioenergy*. Elsevier, pp. 101–127.
- Aryal, N., Zhang, Y., Bajracharya, S., Pant, D., Chen, X. (2022). Microbial electrochemical approaches of carbon dioxide utilization for biogas upgrading. *Chemosphere* 291, 132843.
- Eddy, M., Abu-Orf, M., Bowden, G., Burton, F.L., Pfrang, W., Stensel, H.D., Tchobanoglous, G., Tsuchihashi, R., (Firm), A. (2014). *Wastewater engineering: treatment and resource recovery*. McGraw Hill Education.
- Fischer, K. (2017). Mathematical modelling of electron transfer modes in electroactive microorganisms.
- Heijnen, J.J., Kleerebezem, R. (1999). Bioenergetics of microbial growth. *Encyclopedia of bioprocess technology: Fermentation, biocatalysis, and bioseparation* 1, 267–291.
- Korth, B., Rosa, L.F.M., Harnisch, F., Picioreanu, C. (2015). A framework for modeling electroactive microbial biofilms performing direct electron transfer. *Bioelectrochemistry* 106, 194–206.
- Li, M., Li, Y.-W., Yu, X.-L., Guo, J.-J., Xiang, L., Liu, B.-L., Zhao, H.-M., Xu, M.-Y., Feng, N.-X., Yu, P.-F. (2020). Improved bio-electricity production in bio-electrochemical reactor for wastewater treatment using biomass carbon derived from sludge supported carbon felt anode. *Science of the total environment* 726, 138573.
- May, H.D., Evans, P.J., LaBelle, E. V. (2016). The bioelectrosynthesis of acetate. *Curr Opin Biotechnol* 42, 225–233.
- Rabaey, K., Angenent, L., Schroder, U., Keller, J. (2009). *Bioelectrochemical systems*. IWA publishing.
- Rabaey, K., Verstraete, W. (2005). *Microbial fuel cells: novel biotechnology for energy generation*. Trends Biotechnol 23, 291–298.
- Reichert, P. (1998). *Aquasim 2.0-user manual*. Swiss Federal Institute for Environmental Science and Technology. Dubendorf, Switzerland.
- Rittmann, B.E., McCarty, P.L. (2020). *Environmental Biotechnology: Principles and Applications*, 2nd Edition. ed. McGraw-Hill Education, New York.
- Sivalingam, V., Ahmadi, V., Babafemi, O., Dinamarca, C. (2020). Integrating syngas fermentation into a single-cell microbial electrosynthesis (MES) reactor. *Catalysts* 11, 40.
- Sivalingam, V., Winkler, D., Haugen, T., Wentzel, A., Dinamarca, C. (2022). Syngas fermentation and microbial electrosynthesis integration as a single process unit. *Bioresour Technol* 356, 127314.
- Torres, C.I., Marcus, A.K., Parameswaran, P., Rittmann, B.E. (2008). Kinetic Experiments for Evaluating the Nernst–Monod Model for Anode-Respiring Bacteria (ARB) in a Biofilm Anode. *Environ Sci Technol* 42, 6593–6597. doi. 10.1021/es800970w
- Tremblay, P.-L., Faraghiparapari, N., Zhang, T. (2019). Accelerated  $H_2$  evolution during microbial electrosynthesis with *Sporomusa ovata*. *Catalysts* 9, 166.
- Wagner, R.C., Call, D.F., Logan, B.E. (2010). Optimal set anode potentials vary in bioelectrochemical systems. *Environ Sci Technol* 44, 6036–6041.
- Yates, M.D., Ma, L., Sack, J., Golden, J.P., Strycharz-Glaven, S.M., Yates, S.R., Tender, L.M. (2017). Microbial electrochemical energy storage and recovery in a combined electrotrophic and electrogenic biofilm. *Environ Sci Technol Lett* 4, 374–379.
- Zhang, G., Feng, S., Jiao, Y., Lee, D.-J., Xin, Y., Sun, H. (2017). Cathodic reducing bacteria of dual-chambered microbial fuel cell. *Int J Hydrogen Energy* 42, 27607–27617.

## Alternative fuels for the maritime industry and their impact on flue gas composition

Mostafa Abedini\*, Mohammad Rakibul Hasan Chowdhury\*, Saman Pershen\*, Joachim Rød\*,  
Ishmael Nii Nyarko Solomon\*, Pramod Ghimire\*\*, Nabin Aryal\*

\*University of South-Eastern Norway, Department of Process, Energy and Environmental Technology  
Campus Porsgrunn (email: [259007](mailto:259007), [258940](mailto:258940), [258928](mailto:258928), [229995](mailto:229995), [250872@student.usn.no](mailto:250872@student.usn.no)  
[nabin.aryal@usn.no](mailto:nabin.aryal@usn.no))

\*\*Kongsberg Digital, Horten, Norway (email: [pramod.ghimire@kongsbergdigital.com](mailto:pramod.ghimire@kongsbergdigital.com))

**Abstract:** The maritime industry contributes to 80-90% of global trade and is on an increasing trend. However, it is also responsible for substantial amounts of greenhouse gas (GHG) emissions such as carbon dioxide (CO<sub>2</sub>), nitrogen oxides (NO<sub>x</sub>), sulfur oxides (SO<sub>x</sub>), carbon monoxide (CO), and hydrocarbons (HC). Therefore, industries are searching for alternative solutions to reduce GHG emissions by using alternative fuels. This study presents a novel investigation exploring the performance of various alternative marine fuels such as liquefied natural gas (LNG), methanol (MeOH), ammonia (NH<sub>3</sub>), and hydrogen (H<sub>2</sub>) in terms of combustion and emissions. Such comprehensive evaluation is limited in literature, making this study uniquely valuable in contributing to the field. The study assesses the impact of different equivalence ratios on emissions for the studied fuel profiles using Cantera and Aspen HYSYS simulations. Results show that CO<sub>2</sub> peaks at the stoichiometric ratio, with CO rising from 0.8 to 1.1. Non-carbon fuels like NH<sub>3</sub> and H<sub>2</sub> emit fewer GHGs than carbonaceous fuels such as LNG and MeOH. H<sub>2</sub> has the highest energy release at 87.21 MJ per kg, while NH<sub>3</sub> shows lower emission levels, suggesting its potential as a sustainable maritime fuel. This research emphasizes the significance of choosing the right fuel to mitigate maritime emissions, highlighting NH<sub>3</sub> and H<sub>2</sub> as promising alternatives.

**Keywords:** Fuel, Ammonia, Flue gas, Emissions, Aspen HYSYS, Python, Simulation, and Cantera.

### 1. INTRODUCTION

The global maritime sector is crucial in facilitating international transportation and trade. Maritime transportation, while efficient and relatively clean per unit of material transported, has gained attention due to its fuel efficiency and growth projections. It is expected to increase at an annual rate of 5.3% between 2010 and 2035. Nevertheless, it also significantly contributes to greenhouse gas (GHG) emissions, which are expected to double or triple by 2050 if no measures are taken. Other than environmental damage, these emissions can lead to health concerns, with shipping particulate matter (PM) emissions linked to thousands of cardiopulmonary and lung cancer deaths globally (Moirangthem and Baxter, 2016).

According to the International Maritime Organization (IMO), maritime transport contributes nearly 2.5% of GHG emissions worldwide and generates one billion tons of carbon dioxide (CO<sub>2</sub>) annually. Consequently, maritime industry is actively seeking alternatives to mitigate GHG emissions.

In April 2018, the IMO established ambitious targets through the Marine Environmental Protection Committee (MEPC) resolution MEPC.304(72) to decarbonize the global fleet. This strategy outlines initial goals to reduce the average CO<sub>2</sub> emissions per transport work by a minimum of 40% from 2008 levels by 2030 and by 70% by 2050. Additionally, this target aims to decrease the total annual GHG emissions from shipping by at least 50 % by 2050. Achieving these objectives

involves employing technical and operational approaches, as well as exploring alternative fuels (ABS, 2021).

In the pursuit of deconcentrating the shipping industry before 2050, advancements in ship technology primarily concentrate on newly constructed vessels. However, considering the extended operational lifespan of ships, it is evident that a substantial portion, approximately 20% of the global fleet according to certain projections, will continue to operate beyond 2050 despite being originally designed for fossil fuel propulsion as shown in Fig. 1. It is imperative to address the decarbonization of these existing vessels as a vital component of the broader maritime energy transition. One viable approach involves retrofitting these ships to operate on carbon-neutral or zero-carbon fuels. This retrofitting process may necessitate modifications to the vessel's engine, tanks, pipework, systems, and overall structure. This strategy acknowledges the importance of adapting existing vessels to align with sustainable and environmentally friendly energy sources, contributing significantly to the overall objective of reducing carbon emissions in the maritime sector (LR, 2023).

The maritime industry is currently experiencing a notable transition in fuel technology. In 2023, half of the ordered tonnage is equipped to utilize LNG, LPG, or MeOH in dual-fuel engines. This represents a significant increase compared to one-third of the tonnage on order in 2022. The shift towards alternative fuels is evident not only in new builds but also in existing vessels. Presently, 6.52% of the tonnage of

operational ships can operate on alternative fuels, reflecting an increase from 5.5% in the previous year (DNV, 2023). This trend underscores the industry's commitment to adopting more sustainable and eco-friendly fuel options across both new and existing maritime assets.

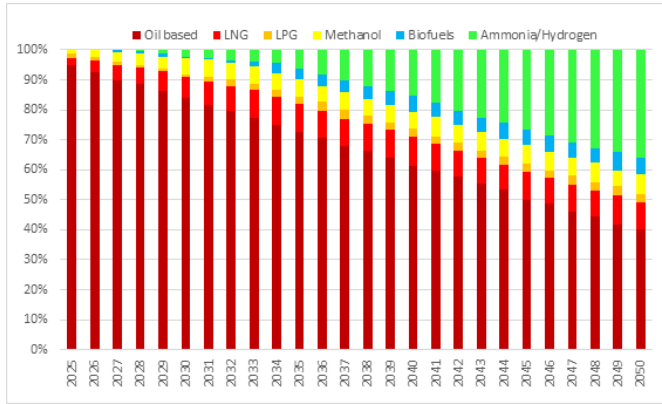


Fig. 1. Predicted marine fuel use to 2050 (ABS, 2021).

Furthermore, the production of alternative fuels has been rigorously investigated to meet the demands of the maritime industry and other sectors (Aryal et al., 2021). Renewable-based alternative fuels and chemicals are also recognized for their environmental benefits (Gadkari et al., 2021). Recently, simulation-based studies are considered to solve problems in different sectors as they enable realistic exploration of real-world problems in a safer, more cost-effective, and efficient manner (Ghimire, et al., 2021a). Further, efficiency and emissions estimation using various models for different ship types using conventional fuels are discussed in (Ghimire, et al., 2022) and (Ghimire, et al., 2024).

This study aims to investigate the emissions of LNG, MeOH, NH<sub>3</sub>, and H<sub>2</sub> as alternative fuels in marine diesel engines. In this simulation-based study, the amount of NO<sub>x</sub>, CO, CO<sub>2</sub> and GHGs are compared based on the same amount of energy produced by each fuel. Aspen HYSYS is used to model an internal combustion diesel engine, and parameter optimization is done for limiting the emissions for each fuel by changing the equivalence ratio. Cantera (Python Code) is also employed to model the combustion process and calculate the chemical potential of every element present in the flue gas.

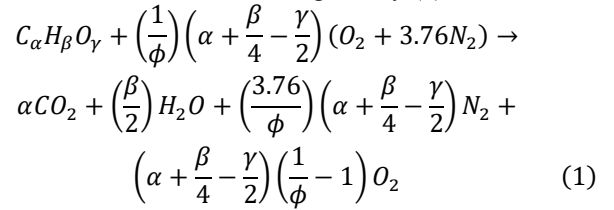
## 2. CONCEPT AND GOVERNING EQUATIONS

In this section, the process of production emissions in an internal combustion engine (ICE) is discussed theoretically, and the effective parameters of its reduction are explained through academic concepts. To create a base for evaluating and comparing the pollution of different fuels, scientific relationships are presented, and the most important emitted gases from a marine diesel engine are defined to compare alternative marine fuels.

### 2.1 Chemical reactions in the combustion process

The combustion process is a chemical reaction between a fuel and an oxidizing agent, typically the oxygen (O<sub>2</sub>), from the air. This reaction releases heat and produces combustion

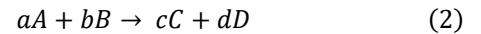
byproducts, such as CO<sub>2</sub>, water vapor, and other gases, depending on the composition of the fuel. The general chemical reaction of fuel and air is given by (1):



In equation (1),  $\phi$  is equivalence ratio, and it is defined as the actual FAR (fuel-air ratio) per stoichiometric FAR. When  $\phi$  is one, it means that the mixture is in stoichiometric condition, but when  $\phi$  is smaller than one, it means that there is more air available to burn the fuel and in common, the mixture is called a lean mixture in conditions where  $\phi > 1$ , the amount of fuel is more than the required air, and as a result, some fuels remain unburned after combustion. The mixture is called rich in the last situation (McAllister et al., 2011). The equivalence ratio plays a critical role in flame temperature as well as emission production. Hence, (1) is used to control the amount of feed (reactants) in an engine to control the flame temperature and flue gas concentrations. Although it shows the general chemical reaction, in practice, a series of sequential or simultaneous reactions are encountered. In hydrocarbons fuel, by increasing the number of carbons the species and steps of elementary reactions increased rapidly. For instance, combustion of CH<sub>4</sub> has 53 species in reaction mechanism and C<sub>8</sub>H<sub>8</sub> has 857 species (McAllister et al., 2011). In real-world scenarios, each of these reactions is time-dependent and is affected mainly by temperature, that will be discussed later.

### 2.2 Chemical kinetics

Chemical kinetics involves the study of the rates at which chemical reactions occur. The rate of reactions defines the speed of species consumption and production. Combustion chemistry exhibits two significant characteristics not typically seen in other chemical systems. Initially, the speed at which combustion reactions occur is highly influenced by temperature. Additionally, a considerable amount of heat is released during a chemical reaction, which affects the temperature. To elaborate it, consider (2) describe a general elementary reaction, and  $a$ ,  $b$ ,  $c$  and  $d$  are stoichiometry coefficient.



According to (2) the rate of reaction progress ( $\dot{q}_{RXT}$ ) is calculated by equation (3) that in which the Arrhenius rate constant  $k$  is the constant of proportionality and calculated from (4). The expression for the rate at consumption of reactant A, ( $\hat{r}_A$ ) is then provided by (5).

$$\dot{q}_{RXT} = k[A]^a[B]^b \quad (3)$$

$$k = A_0 T^b \exp\left(-\frac{E_a}{R_u T}\right) = A_0 T^b \exp\left(-\frac{T_a}{T}\right) \quad (4)$$

$$\frac{d[A]}{dt} = \hat{r}_A = -a \cdot \dot{q}_{RXT} \quad (5)$$

In (3) the activation energy ( $E_a = T_a \hat{R}_u$ ) is the minimum energy that required to have successful collision to result in a successful reaction. Coefficient b is for collision. As it can be concluded the rate of reaction or in other way rate of consumption or production of species is directly related to temperature. In practice,  $k$  is derived from experimental data. For instance, Figure 2 illustrates how someone can calculate the Arrhenius rate constant of some elementary reactions of burning  $\text{CH}_4$  by using (6) and plotting the test data.

$$\ln k = \ln A_0 - \frac{E_a}{\hat{R}_u T} \quad (6)$$

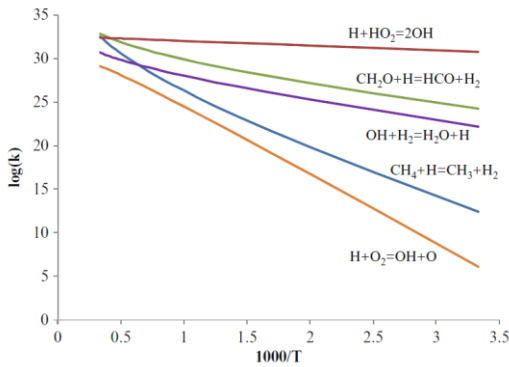
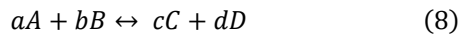


Fig. 2.  $k$  value for different elementary reactions of burning methane ( $\text{CH}_4$ ) (McAllister et al., 2011).

In equilibrium condition, reaction rate is calculated by forward and backward reaction rate constants as  $k_f$  and  $k_b$ . In (7) equilibrium constant is based on concentration or partial pressure defined.

$$K_c = \frac{k_f}{k_b} \quad (7)$$

For example, the following equilibrium reaction is taken for calculation:



In this way, the reaction progress is defined by the following relation:

$$\dot{q}_{RXT} = k_f[A]^a[B]^b - k_b[C]^c[D]^d \quad (9)$$

Therefore, equilibrium constant can be derived based on thermodynamics properties in the reaction as follows:

$$k_c = \frac{k_f}{k_b} = \frac{[C]^c \cdot [D]^d}{[A]^a \cdot [B]^b} = K_p(T) \left( \frac{\hat{R}_u T}{101.3 \text{ kPa}} \right)^{a+b-c-d} \quad (10)$$

In that (10)  $K_p(T)$ , is the equilibrium constant based on partial pressures and can be defined by the following equation:

$$K_p(T) = \exp \left\{ \frac{a\hat{g}_A^0 + b\hat{g}_B^0 - c\hat{g}_C^0 - d\hat{g}_D^0}{\hat{R}_u T} \right\} \quad (11)$$

Where  $\hat{g}_A^0$ , is the Gibbs free energy and can be found in the thermodynamic tables.  $K_p(T)$  is unitless and temperature dependent.

Therefore, in a combustion process of a general fuel, the rate of progress can be computed from activation energy, temperature, and concentration by the following (12):

$$\begin{aligned} \dot{q}_{RXT} &= A_0 \exp \left( -\frac{E_a}{\hat{R}_u T} \right) [\text{Fuel}]^a [\text{O}_2]^b = \\ &= A_0 \exp \left( -\frac{E_a}{\hat{R}_u T} \right) x_{\text{fuel}}^a x_{\text{O}_2}^b \end{aligned} \quad (12)$$

To use these equations and calculate emissions for each fuel at different temperature and pressure condition, commercial and scientific tools such as Aspen HYSYS and Cantera (Python) are developed to ease the simulation of combustion processes. In this study, authors used both software to model the combustion process in a marine diesel engine. The methods and the results will be discussed in the next chapter. In the next section the emission from combusting fuels is discussed.

### 2.3 Emissions

Emissions in maritime industries refers to the release of GHG and air pollutants from ships that transport goods and passengers across the world's oceans. Common types of emissions are  $\text{CO}_2$ ,  $\text{NO}_x$ ,  $\text{SO}_x$ , PM and  $\text{CH}_4$ ; They are dependent on the flame temperature.

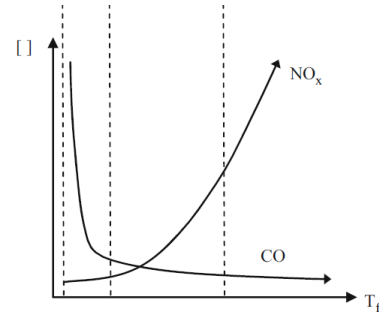


Fig. 3.  $\text{NO}_x$  and CO concentration versus flame temperature ( $T_f$ ) (McAllister et al., 2011).

Figure 3 shows the challenges of achieving lean combustion, emphasizing issues with flame stability at low temperatures. In low temperatures, less  $\text{NO}_x$  emissions is produced. Lean blow off is the condition where the combustion flame transitions from a lean to a condition where combustion cannot be sustained (McAllister et al., 2011).

The calculated emissions are the results of emissions formed in the elementary reactions. Using the Arrhenius constant ( $k$ ,  $K$ ,  $A_0$ ,  $E_a$ ) that is achieved from the chemical kinetics mechanism, and having pressure of chamber can calculate the combustion engine's production rate or emission rate.

#### 2.4.1 Greenhouse gases (GHG) emissions

$\text{CO}_2\text{e}$ , or  $\text{CO}_2$  equivalent, is a measurement unit used to quantify the impact of various GHG on global warming and



climate change. It allows scientists to standardize measurements of gases like CO<sub>2</sub>, CH<sub>4</sub>, nitrous oxide, and synthetic gases. CO<sub>2e</sub> helps in understanding the contributions of different gases to rising temperatures and environmental changes. The main GHG included in CO<sub>2e</sub> measurements are CO<sub>2</sub>, CH<sub>4</sub>, nitrous oxide, hydrofluorocarbons, perfluorocarbons, sulfur hexafluoride, and nitrogen trifluoride. CO<sub>2e</sub> is used to create a standardized metric that simplifies the comparison of the global warming potential of various gases.

Table 1 shows the gases typically encompassed within the CO<sub>2e</sub> (carbon dioxide equivalent) measurement. (13) is total sum of this CO<sub>2e</sub> that is mostly reported in kg unit (US EPA, 2023).

$$CO_{2e} = \sum(GHG \text{ emission} \times GWP) \quad (13)$$

Table 1: Global warming potential (GPW) for GHG.

GHG	GPW
Carbon dioxide (CO <sub>2</sub> )	1
Methane (CH <sub>4</sub> )	25
Nitrous oxide (N <sub>2</sub> O)	298

#### 2.4.2 Alternative fuel selection

According to the definition of this project, fuels - LNG, MeOH, NH<sub>3</sub>, and H<sub>2</sub>- in pure conditions were selected to model in the simulation. The chosen fuels are common fuels that researchers are working on as pure or mixture fuel for compression ignition (CI) engines and they are extensively expected to be utilized in marine applications (ABS, 2021). Their distinctive properties make them adaptable options for sustainable energy solutions. NH<sub>3</sub>, renowned for emissions reduction, aligns with green initiatives. LNG, a cleaner-burning natural gas, is favored for its lower emissions profile. Biofuel, derived from renewable sources, offers an eco-friendly alternative. MeOH, a liquid fuel, is valued for its versatility and low carbon footprint. H<sub>2</sub>, a clean energy carrier, stands out for its potential in fuel cell applications, promising a greener marine industry (Moirangthem and Baxter, 2016).

LNG is a promising maritime fuel, reducing CO<sub>2</sub> emissions by 33.7%. Case studies reveal substantial environmental gains, with LNG leading in emissions, cost, and engine adaptability. However, high bunkering station costs hinder industry acceptance. Ongoing research focuses on optimizing infrastructure for improved viability (Wang et al., 2023).

MeOH emerges as a promising alcohol fuel for maritime use, offering cleaner energy and reduced emissions. Its high-octane rating, compatibility with existing engines, and ease of integration make it a viable solution. Advanced engines, like MAN B&W LGIM, showcase its potential for improved performance and efficiency in maritime applications (MAN, 2021; Tian et al., 2022; Ghimire et al., 2021b). NH<sub>3</sub> has the potential use in maritime transport for decarbonization faces challenges such as regulatory changes and the absence of ready-to-sail NH<sub>3</sub>-fueled ships. NH<sub>3</sub> engines, researched since 1900, are costlier than conventional LNG and diesel engines. Storage, safety concerns, and emission control methods impact NH<sub>3</sub> feasibility. Different NH<sub>3</sub> grades and storage types exist, each with specific considerations. NH<sub>3</sub> can serve as an energy

carrier, particularly in fuel cells, offering advantages over H<sub>2</sub>. Combustion challenges include NH<sub>3</sub>'s properties, flammability, and NO<sub>x</sub> emissions. Overall, NH<sub>3</sub>'s adoption hinges on overcoming technological, safety, and regulatory hurdles (Reiter and Kong, 2008).

H<sub>2</sub> potential as a green marine fuel is evident, but challenges in emissions during production and low energy density complicate its viability. Grey H<sub>2</sub>, which is generated through the processing of alternative fossil fuels or natural gas, dominates production (75%), limiting emissions reduction by producing about 70-gram CO<sub>2</sub> per MJ energy from H<sub>2</sub>. However, its exceptional energy content could enhance efficiency, but volume challenges and cryogenic storage requirements pose obstacles. H<sub>2</sub> blends like HLNG or HCNG offer alternatives. Key concerns include safety, storage, and development costs (ABS, 2021).

### 3. SIMULATION OF COMBUSTION CHAMBER

In this simulation study utilizing Cantera, which is open-source software, and Aspen HYSYS, the focus is on modeling the combustion chamber to assess the emissions of five alternative fuels – LNG, MeOH, NH<sub>3</sub>, H<sub>2</sub>, and Bioethanol. By considering optimal engine conditions for each fuel, the project delves into the intricate interplay of thermodynamic properties, combustion kinetics, and emissions. Cantera, with its open-source suite for chemical kinetics and thermodynamics, complements HYSYS, a widely used process simulation software. The examination of flue gas emissions provides valuable insights into the environmental impact of diverse maritime fuels, crucial for advancing sustainable and efficient combustion processes.

#### 3.1 Aspen HYSYS

Aspen HYSYS is a process simulation software used to model and design chemical processes. It is widely used in the oil and gas industry to simulate various processes such as distillation, heat exchange, and chemical reactions. A built-in reactor was used to explore the scope of Aspen HYSYS for evaluating emissions from combustion. Figure 4 illustrates the set-up of the simulation, which was based on work done by (Suyitno et al., 2019).

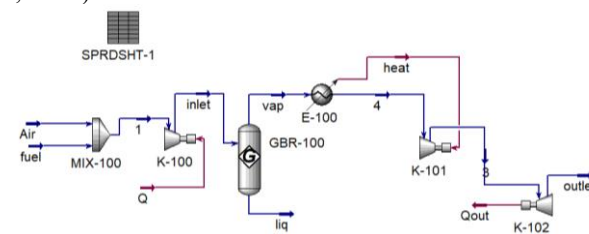


Fig. 4. Simulation of combustion in Aspen HYSYS.

#### 3.2 Cantera

Cantera is an open-source suite of software tools for solving problems involving chemical kinetics, thermodynamics, and transport processes. It is designed to aid scientists and engineers in modelling and simulating a wide range of chemical phenomena, such as combustion, catalysis,

atmospheric chemistry, and materials science. Cantera provides a comprehensive set of features, including a database of chemical species and reactions, thermodynamic and transport property calculations, a solver for kinetic equations, and support for various input and output formats. Cantera version 3.0.0 was used for this project in Python as interface of Cantera (Reiter and Kong, 2008).

#### 4. SIMULATIONS RESULTS

In this section, the Cantera and Aspen HYSYS simulation results for all four alternative fuels are discussed. For the simulations, the initial combustion condition considered for the simulation was taken from state-of-the-art research papers, which are shown in Table 2. The composition of LNG is considered to be the same as LNG from Qatar: 89.87% CH<sub>4</sub>, 6.65% Ethane, 2.30% Propane, and 0.98% Butane (Kanbur et al., 2017). However, as GRI30 does not contain butane's reaction mechanism and chemical kinetics, butane was not considered in the simulation; it was replaced by N<sub>2</sub>

Table 2: Initial conditions of the fuel combustion in engine.

Fuel	Temperature (K)	Initial Pressure (bar)	Reference
LNG	300	30	(Yao et al., 2022)
MeOH	513	75	(Verhelst et al., 2019)
NH <sub>3</sub>	563	39	(Reiter and Kong 2008)
H <sub>2</sub>	300	16.5	(ABS, 2021)

Moreover, as in this project the design of the engine is out of the scope of the project, the simulation considers the optimal engine conditions. In HYSYS, the reactor uses the Gibbs free energy method, but Cantera uses chemical kinetic mechanisms such as the GRI30 library database, which is based on experimental reaction rates developed by scientific institutes.

##### 4.1 LNG Simulation

The CO<sub>2</sub> and CO emissions of burning LNG is shown in Fig. 5 in various equivalence ratios. The results of HYSYS and Cantera are similar, with small differences.

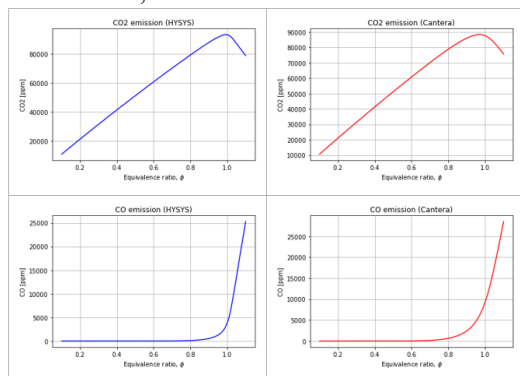


Fig. 5. CO<sub>2</sub> and CO from LNG in simulation.

Looking at the figures, the maximum CO<sub>2</sub> emissions is about 90000 ppm in both simulations that produced in about stoichiometric reaction. Additionally, it is obvious that by

increasing the equivalence ratio, CO concentration rises rapidly. Whereas Figure 6 compares the emission parts of NO<sub>2</sub>, NO, and N<sub>2</sub>O for LNG in two simulations with HYSYS and Cantera. In lean mixture, a major part of emission from LNG is NO. It peaks at the equivalence ratio of about 0.8, producing 3500 ppm and 8000 ppm emission in HYSYS and Cantera, respectively. N<sub>2</sub>O emission peaks at a fuel/air ratio of 0.6. N<sub>2</sub>O emission is very low in both simulations.

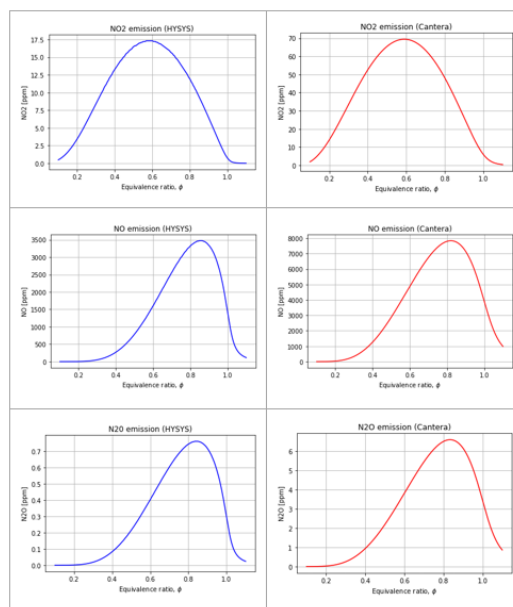


Fig. 6. NO<sub>x</sub> and N<sub>2</sub>O from LNG in HYSYS and Cantera simulation.

##### 4.2 Methanol Simulation

Figure 7 illustrates CO<sub>2</sub> and CO emission of MeOH in various equivalence ratio.

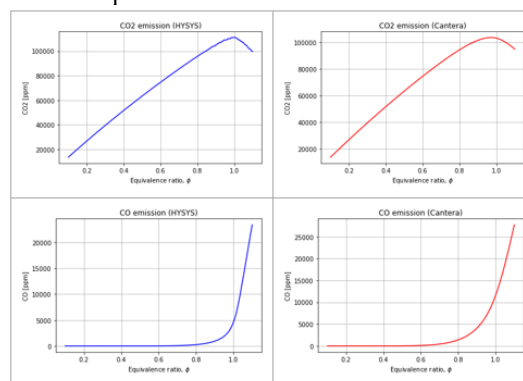


Fig. 8. CO<sub>2</sub> and CO from MeOH in HYSYS and Cantera.

The graphs are almost the same with negligible differences in quantities that would be because of the concept of solving chemical reactions in both utilized software. The maximum amount of CO<sub>2</sub> is 110000 ppm and 103000 ppm in HYSYS and Cantera respectively. Whereas Figure 8 shows comparison between NO<sub>2</sub>, NO, and N<sub>2</sub>O emissions for MeOH in two simulations with HYSYS and Cantera. In lean mixture, the NO<sub>2</sub> has the highest peak at FAR of 0.5, but as the mixture gets richer the NO, and N<sub>2</sub>O emissions increase up to the equivalence ratio of 0.8, and after that they again decrease.

The maximum NO emissions were calculated at about 5000 ppm and 9000 ppm in Aspen HYSYS and Cantera, respectively.

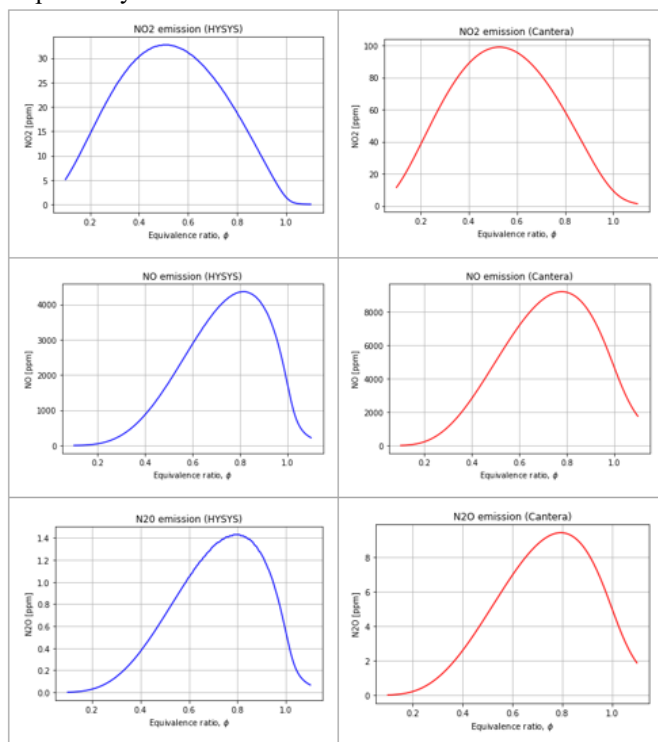


Fig. 8. NO<sub>x</sub> and N<sub>2</sub>O emissions from MeOH combustion in HYSYS and Cantera simulation.

### 4.3 Ammonia Simulation

Simulation was based on an experiment where NH<sub>3</sub> was used in a two-stroke engine (Ichikawa et al., 2023). where initial pressure is 39 bars and temperature are 563 K. The FAR is changed from 0.1 to 1.1 to see how emissions depend on this factor. Figure 9 shows that running in a lean mixture is needed to reduce NH<sub>3</sub> slip from combustion. The combustion of NH<sub>3</sub> produces no carbon in the emissions and only NO<sub>x</sub> and N<sub>2</sub>O is considered shown in Fig. 10. NO<sub>2</sub> emissions are the highest around the equivalent ratio of 0.5, and NO emissions are the highest around FAR of 0.8, the same as the N<sub>2</sub>O emissions. Similarities of results from Aspen HYSYS and Cantera simulation are observed here, also like all other fuels. The general trend that is seen for all fuels is that the CO<sub>2</sub> emission peaks at the stoichiometric ratio, where CO starts to increase rapidly after that.

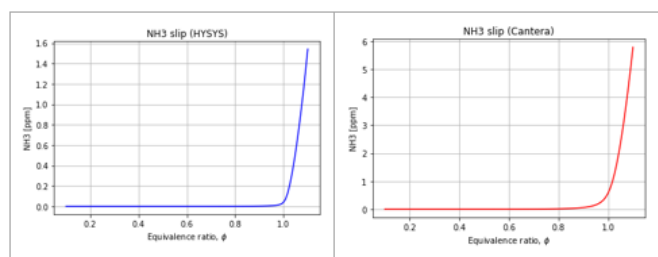


Fig. 9. NH<sub>3</sub> slip from NH<sub>3</sub> combustion in HYSYS and Cantera simulation.

NO is the major part in NO<sub>x</sub> emission and usually peaks at about equivalence ratio of 0.8, while NO<sub>2</sub> is usually maximized between 0.4 and 0.6 in all fuels. Generally, emissions decrease with reducing equivalence ratio.

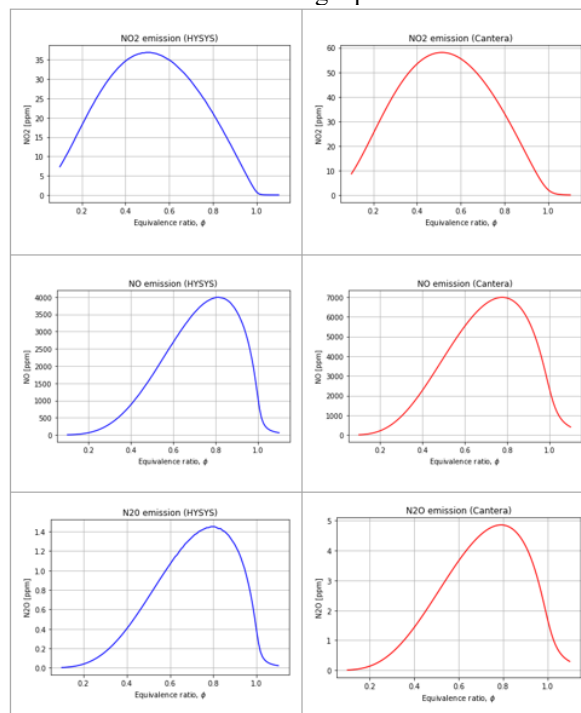


Fig. 10. NO<sub>x</sub> and N<sub>2</sub>O emissions from Ammonia combustion in HYSYS and Cantera simulation.

### 4.4 Hydrogen Simulation

Figure 11 illustrates flue gas emission of burning H<sub>2</sub> in various equivalence ratio.

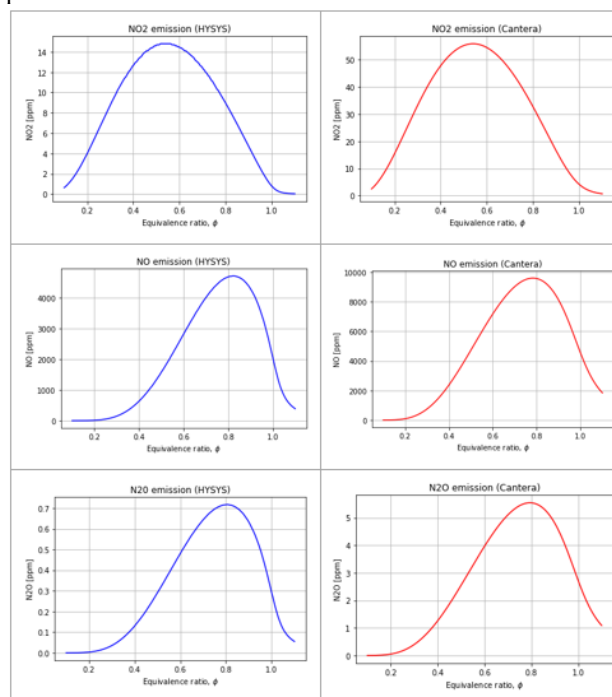


Fig. 11. NO<sub>x</sub> and N<sub>2</sub>O emissions from H<sub>2</sub> combustion in HYSYS and Cantera simulation.

The similarities of graphs between Aspen HYSYS and Cantera show validation of the simulations. Generally, the emission concentrations are higher in Cantera’s results than HYSYS ones. At FAR of 0.8 NO emission is maximized with about 5000 ppm in HYSYS simulation and 10000 ppm in Cantera modelling. NO<sub>2</sub> emission is always below 55 ppm in both simulations, and N<sub>2</sub>O emission is negligible. However, clean H<sub>2</sub> does not produce any CO or CO<sub>2</sub>, which is a great advantage of using H<sub>2</sub> as a fuel instead of fossil fuels. Table 2 illustrates flue gas emissions from burning H<sub>2</sub> in various equivalence ratios. The similarities of the graphs between Aspen HYSYS and Cantera show that the simulations are validated.

### 5. COMPARING FUELS AND DISCUSSION

The flue gas emissions produced by various fuels were investigated under the initial conditions specified in Table 2, with an equivalence ratio of 1 maintained for all fuels.

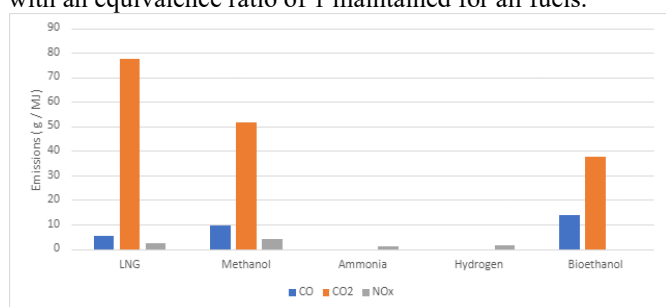


Fig. 12. Emissions (g/MJ) to produce 1 MJ of energy.

By standardizing the energy output across different fuels, the corresponding emissions were effectively compared.

Table 3. Emissions from fuels while producing 1 MJ energy.

Fuel	Required fuel (kg/MJ)	CO (g/MJ)	CO <sub>2</sub> (g/MJ)	CO <sub>2</sub> e (g/MJ)	NO <sub>x</sub> (g/MJ)
LNG	0.04	5.27	77.48	2157.18	2.35
MeOH	0.06	9.64	51.64	1444.41	4.26
NH <sub>3</sub>	0.08	0.00	0.00	11.91	1.25
H <sub>2</sub>	0.01	0.00	0.00	35.14	1.72

Figure 12 demonstrates that carbon-based fuels exhibit the highest emissions, with LNG having the greatest impact, followed by MeOH. The results of this analysis are also presented in Table 3.

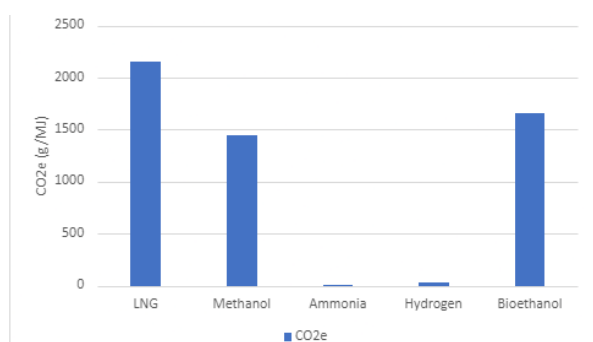


Fig. 13. CO<sub>2</sub> equivalence for different fuels (g/MJ).

Figure 13 presents a comparison of fuels based on their CO<sub>2</sub> equivalence. It is evident that carbon-containing fuels, such as LNG and MeOH, generate significantly higher CO<sub>2</sub> equivalence compared to NH<sub>3</sub> and H<sub>2</sub>. However, fuel consumption also needs to be taken into consideration when suggesting fuel composition due to its effect on cost estimations.

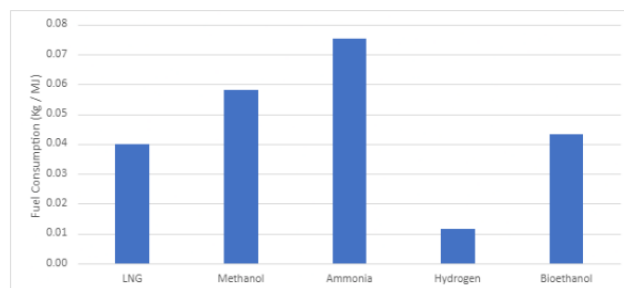


Fig. 14: Comparison of alternative marine fuels based on required mass for producing 1 MJ energy (kg/MJ).

Figure 14 shows what amount (kg) of fuel is needed to produce 1 MJ of energy where it can be observed a quite opposite scenario. LNG overperforms other fuels in terms of fuel consumption, H<sub>2</sub> is the only exception which seems most efficient in terms of fuel consumption too. However, fuel consumption also needs to be taken into consideration when suggesting fuel composition due to its effect on cost estimations.

### 6. CONCLUSIONS

The study sheds light on emissions from various marine fuels, contributing valuable data for sustainable energy choices in maritime industries. The comparison of GHG levels shows that when considering all the emissions, the non-carbon fuels still release fewer emissions per unit of energy. Based on different initial conditions of all the alternative maritime fuels, H<sub>2</sub> overperforms the other fuels in terms of fuel consumption. Nonetheless, the infrastructure cost of using H<sub>2</sub> should also be considered as one of the decision factors. Although, among all studied fuels, NH<sub>3</sub> shows good potential as an alternative fuel for the marine industries by producing the least emissions, attention should be given to its safety issues. However, the decision on fuel selection should also consider fuel consumption. The research can be expanded to use complex models of engines under different conditions, different ranges of fuel mixtures, and infrastructure as a future work scope. Fuel blends, for example, biofuel blends, can be explored with different compositions and sources. Emissions during the production process of the fuels need to be incorporated to reduce the overall carbon footprint. Additionally, a multi-criteria decision analysis study can be conducted for optimum fuel selection based on criteria like emissions, cost, and efficiency.

### ACKNOWLEDGEMENTS

Nabin Aryal expresses gratitude to the Norwegian Partnership Program for Global Academic Cooperation (NORPART) for funding the Re-Tech project.

## REFERENCES

- ABS (2021) SUSTAINABILITY WHITEPAPER: HYDROGEN AS MARINE FUEL. American Bureau of Shipping (ABS).
- Aryal, N., Ottosen, L. D. M., Kofoed, M. V. W., and Pant, D. (2021). *Emerging Technologies and Biological Systems for Biogas Upgrading* (N. Aryal, L. D. M. Ottosen, M. V. W. Kofoed, and D. Pant (eds.); I). Elsevier. <https://www.elsevier.com/books/emerging-technologies-and-biological-systems-for-biogas-upgrading/aryal/978-0-12-822808-1>
- DNV (2023) *Energy Transition Outlook, MARITIME FORECAST TO 2050*. Det Norske Veritas (DNV).
- US EPA (2023) *Greenhouse Gases Equivalencies Calculator - Calculations and References*, US EPA. <https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references>.
- Ghimire, P., Zadeh, M., Pedersen, E., Thorstensen, J., (2021) a. *Dynamic Modeling, Simulation, and Testing of a Marine DC Hybrid Power System in IEEE Transactions on Transportation Electrification*, 7, 905–919. doi:10.1109/TTE.2020.3023896.
- Ghimire, P., Zadeh, M., Thorstensen, J., Pedersen, E., (2021) b. *Data-Driven Efficiency Modeling and Analysis of All-Electric Ship Powertrain; A Comparison of Power System Architectures in IEEE Transactions on Transportation Electrification*, doi: 10.1109/TTE.2021.3123886.
- Ghimire, P., Karimi, S., Zadeh, M., Nagalingam, K.K., Pedersen, E., (2022). *Model-based efficiency and emissions evaluation of a marine hybrid power system with load profile in Electrical Power System Research*, 212. doi:10.1016/j.eprs.2022.108530.
- Ghimire, P., Zadeh, M., Thapa, S., Thorstensen, J., Pedersen, E., (2024). *Operational Efficiency and Emissions Assessment of Ship Hybrid Power Systems with Battery; Effect of Control Strategies in IEEE Transactions on Transportation Electrification*, doi:10.1109/TTE.2024.3365351.
- Gadkari, S., Mirza Beigi, B. H., Aryal, N., and Sadhukhan, J. (2021). *Microbial electrosynthesis: is it sustainable for bioproduction of acetic acid? RSC Advances*, 11(17), 9921–9932. doi:10.1039/d1ra00920f
- Ichikawa, Y., Suzuki, T., Yamamoto, H., and Takahashi, M. (2023) *Experimental study of combustion process of NH<sub>3</sub> stratified spray using imaging methods for NH<sub>3</sub> fueled large two-stroke marine engine, Applications in energy and combustion science*, 13, 100119-. doi:10.1016/j.jaecs.2023.100119.
- Kanbur, B.B., Yilmaz, M., Ersoz, M., and Kilicarslan, A. (2017) *Cold utilization systems of LNG: A review, Renewable and Sustainable Energy Reviews*, 79, 1171-1188,. doi:10.1016/j.rser.2017.05.161.
- LR (2023) *ENGINE RETROFIT REPORT: Applying alternative fuels to existing ships*. Lloyd's Register.
- MAN (2021) *The Methanol fuelled MAN B&W LGIM Engine*. MAN Energy Solutions. [https://www.man-es.com/docs/default-source/document-sync-archive/the-methanol-fuelled-man-b-w-lgim-engine-eng.pdf?sfvrsn=36b925d2\\_7](https://www.man-es.com/docs/default-source/document-sync-archive/the-methanol-fuelled-man-b-w-lgim-engine-eng.pdf?sfvrsn=36b925d2_7).
- McAllister, S., Chen, J.-Y. and Fernandez-Pello, A.C. (2011) *Fundamentals of Combustion Processes*. New York, NY: Springer (Mechanical Engineering Series). doi:10.1007/978-1-4419-7943-8.
- Moirangthem, K. and Baxter, D. (2016) *Alternative fuels for marine and inland waterways*. LU: Publications Office of the European Union. doi:10.2790/227559 (Accessed: 12 November 2023).
- Reiter, A.J. and Kong, S.-C. (2008) *Demonstration of Compression-Ignition Engine Combustion Using Ammonia in Reducing Greenhouse Gas Emissions, Energy and Fuels*, 22(5), 2963–2971. doi:10.1021/ef800140f.
- Suyitno, Haryanto, A., Widayat, T. and Sutrisno, A. (2019) *The Effect of Biogas Pressure in the Performance and Emission of Spark Ignition Engine, IOP Conference Series: Materials Science and Engineering*, 694(1), 012022. doi:10.1088/1757-899X/694/1/012022.
- Tian, Z., Wang, X., Li, H., and Chen, Y. (2022) *The effect of methanol production and application in internal combustion engines on emissions in the context of carbon neutrality: A review, Fuel*, 320, 123902. doi:10.1016/j.fuel.2022.123902.
- Verhelst, S., Turner, J.W.G., Sileghem, L., and Vancoillie, J. (2019) *Methanol as a fuel for internal combustion engines, Progress in Energy and Combustion Science*, 70, 43–88. doi:10.1016/j.peccs.2018.10.001.
- Wang, Q., Li, Y., Zhang, J., and Chen, X. (2023) *The use of alternative fuels for maritime decarbonization: Special marine environmental risks and solutions from an international law perspective, Frontiers in Marine Science*, 9. doi:10.3389/fmars.2022.1082453.
- Yao, S., Zhang, L., Li, X., and Chen, Z. (2022) *Integrated design and optimization research of LNG cold energy and main engine exhaust heat utilization for LNG powered ships, Case Studies in Thermal Engineering*, 33, 101976,. doi:10.1016/j.csite.2022.101976.



## Performance of direct air capture process in honeycomb channel configuration: A CFD study

Majid Nejadseifi\*. Shervin Karimkashi\*\*. Tero Tynjälä\*. Payman Jalali\*

\*Laboratory of Thermodynamics, School of Energy Systems, Lappeenranta-Lahti University of Technology, Lappeenranta, 53850, Finland (Tel: 358-415750813; e-mail: [majid.nejadseifi@lut.fi](mailto:majid.nejadseifi@lut.fi), [tero.tynjala@lut.fi](mailto:tero.tynjala@lut.fi), [payman.jalali@lut.fi](mailto:payman.jalali@lut.fi)).

\*\*Laboratory of Energy conversion and systems, Department of Mechanical Engineering, School of Engineering Sciences, Aalto university, Espoo, 02150, Finland (Tel: +358505602614; e-mail: [shervin.karimkashiarani@aalto.fi](mailto:shervin.karimkashiarani@aalto.fi))

**Abstract:** This study presents a kinetic reaction modeling method for direct air capture (DAC) process of CO<sub>2</sub> adsorption using computational fluid dynamics (CFD). Here, CO<sub>2</sub> is adsorbed by amine coated air-surface contact area. The Langmuir model is employed to represent the kinetics of CO<sub>2</sub> adsorption. Despite neglecting the diffusive phase of the adsorption, which is dominant only in the later stages of adsorption, the surface reaction model gives a satisfactory representation of the adsorption for a major part of the process. Honeycomb reactors with coated adsorbent may yield a better control of reaction rate and pressure drop compared to commonly used packed bed adsorption columns. Their enhanced performance in distributing the flow homogeneously between and within channels creates unique features for the reactor. In this study, we have analyzed mechanical and electrical energy demand for adsorbing CO<sub>2</sub> per unit mass of adsorbed CO<sub>2</sub> as a function of air flow rate. Adsorption performance of honeycomb structure is anticipated to significantly improve in comparison to the packed beds.

**Keywords:** CFD, Hexagonal channel, adsorption, Direct Air Capture, surface reaction

### 1. INTRODUCTION

From scientific viewpoints, excess amount of CO<sub>2</sub> in atmosphere causes climate change with making oceans more acidic (Wang et al., 2011). As the solution for this problem, we need to replace fossil primary energy sources with CO<sub>2</sub> neutral energy sources. However, there is still a significant release of CO<sub>2</sub> (33 billion tons per year in 2021) which indicates great challenges ahead of fixing global warming (Zhongming et al., 2021). The negative consequences of CO<sub>2</sub> accumulation in the atmosphere will exist for thousands of years, even if emissions stop today (Solomon et al., 2009). So, there will be also needed to remove CO<sub>2</sub> out of the atmospheric air, which is addressed as negative emissions. In this context, we should increase annual CO<sub>2</sub> removals from the atmosphere to 10 billion tons of CO<sub>2</sub> by 2050 (McQueen et al., 2021). It is a challenging objective, which makes scientists to explore different ways to find the most efficient method. One promising method is the Direct Air Capture (DAC), which removes CO<sub>2</sub> directly from the air (Sanz-Pérez et al., 2016).

In order to obtain a comprehensive and detailed analysis, computational fluid dynamics (CFD) simulations can be performed. In a series of studies (Deutschmann et al., 2001; Tischer and Deutschmann, 2005; Nejadseifi et al., 2024), a modern multi-dimensional approach to detailed modeling of fluid flow transfer processes and chemical kinetics have been developed. Specifically, there has been critical evaluation of plug flow, boundary layer, and general three-dimensional models for simulating steady-state transport processes and chemistry in a honeycomb channel flow (Klenov et al., 2009).

In any gas-solid contact system, such as in CO<sub>2</sub> capture processes, maximizing the interface between gas and solid surfaces while minimizing the pressure drop is a crucial goal. This enhances CO<sub>2</sub> capture efficiency and reduces the energy required for blowing air in direct air capture (DAC) systems. Ceramic monoliths, due to their high surface area to volume ratio, have emerged as prime candidates for DAC contactors. They serve as excellent supports for CO<sub>2</sub> sorbents, as demonstrated by several studies (Choi et al., 2011; Rodriguez-Mosqueda et al., 2018; Thakkar et al., 2016). Monoliths are also recognized for their minimal pressure drop (Thakkar et al., 2016). Recently, Fu and Davis (2023) demonstrated that employing monoliths as air contactors in DAC significantly reduces energy consumption compared to fixed beds. Verougstraete et al. (2020) have proposed the use of a carbon monolith for DAC to facilitate rapid heating and cooling, thereby achieving shorter adsorption-desorption cycles. Monoliths are comprised of straight channels with various cross-section shapes. Depending on their intended application, monoliths can adopt cylindrical, cubic, or hexagonal shapes. Sorbents are applied to the walls of these channels. In the case of CO<sub>2</sub> capture, gaseous CO<sub>2</sub> diffuses from the air stream to the walls of the monolith channels, where it is subsequently adsorbed by the sorbents. Improving the transfer of CO<sub>2</sub> to the sorbent surface means a higher process efficiency and energy saving (Jiang et al., 2023).

In this paper, simulations are carried out for monolithic contactor reactor with conventional straight channels. By changing the hexagonal channel flow rate, we promote the CO<sub>2</sub> transport to the sorbent-gas interface and investigate its

effect on CO<sub>2</sub> capture rate. Simulations are conducted under atmospheric pressure and isothermal conditions at 25 C. The computational CFD package of ANSYS-Fluent, V. 2021 R2 has been used for the simulations. Upon solving the governing equations, steady flow profiles across the channel, local mass conversion between channel surface and gas, variation of CO<sub>2</sub> concentration at the channel outlet, pressure drop across the channel and power consumption per absorbed CO<sub>2</sub> are investigated.

## 2. METHODOLOGY

### 2.1 Sorbent coating and reactor model

In our simulations, a solid sorbent coat on an aluminum support of hexagonal straight channels is constructed. Typical samples of such hexagonal monoliths are presented in Fig. 1 and cross section of meshed symmetrical view in Fig. 2. For simplifying simulations, one sixth of the hexagonal cross section is considered (triangular symmetrical channel). The schematic of the single symmetrical hexagonal reactor is shown Fig. 1, where the air flows through. Considering narrow channel and lower flow velocities, the air flow can be assumed confidently laminar.

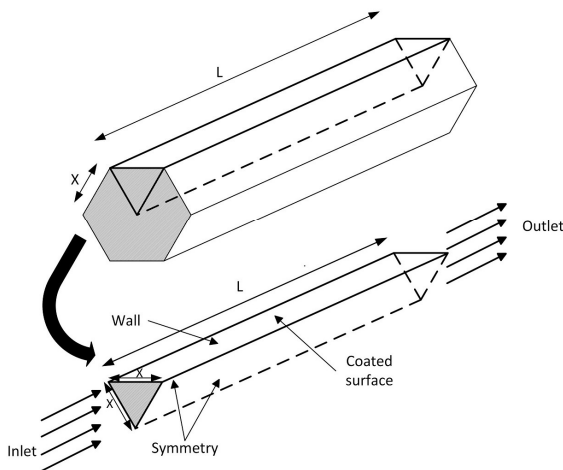


Fig. 1. A. schematical view of the symmetric hexagonal channel.

### 2.2 Mathematical description of Transport Equations:

The steady flow and transient reactive flow of air including CO<sub>2</sub> is simulated in the hexagonal channel in which CO<sub>2</sub> reacts with the sorbent covering the walls of the hexagonal channel. First, steady flow is solved and only the reaction part is transient. Due to small amount of CO<sub>2</sub> in the air flow, it is assumed that the reactions don't have any effect on flow. The Navier–Stokes equations(1,2,3) for incompressible single-phase fluid (air) solved. Mass transport equation (1) considered as time dependent, but momentum equation (2), which is taken as steady state are:

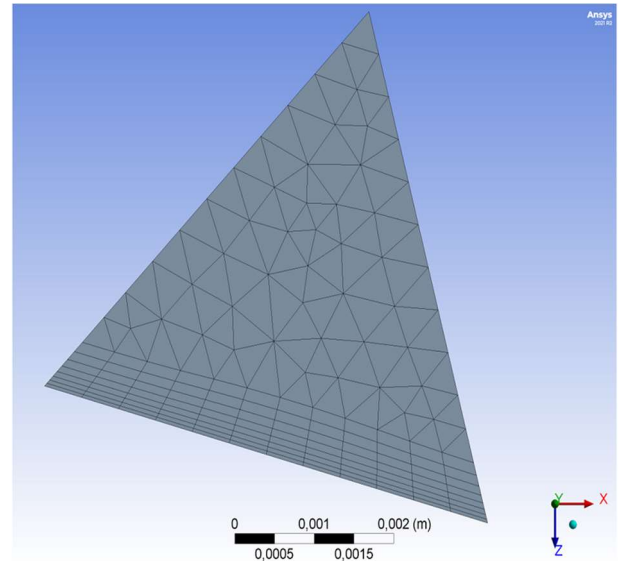


Fig. 2. Cross section mesh view of hexagonal channel cross section prepared for simulations.

$$\nabla \cdot \mathbf{u} = 0 \quad (1)$$

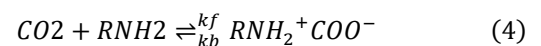
$$(\nabla \cdot \mathbf{u})\mathbf{u} = -\frac{\nabla p}{\rho} + \frac{\mu}{\rho} \nabla^2 \mathbf{u} \quad (2)$$

The inlet and outlet boundary conditions are velocity inlet and pressure outlet, respectively. No-slip wall boundary condition is applied over the sidewalls of the hexagonal channel. Flow is considered isothermal and energy equation is not taken into account, due to neglective share of adsorption heat to overall heat capacity of the flow. After solving steady Navier-Stokes equations for the velocity field,  $\mathbf{u}$ , it is used in the transient CO<sub>2</sub> species transport equation to solve it for CO<sub>2</sub> concentration  $C$ :

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C = D \nabla^2 C \quad (3)$$

Here,  $C$  represents the concentration of CO<sub>2</sub> in the interstitial space, and  $D$  is the diffusion coefficient of CO<sub>2</sub> in air. There is no source term in this equation since CO<sub>2</sub> is not produced or consumed within the flow. However, CO<sub>2</sub> adsorption occurs over the reactive surface of the channel. We apply a surface reaction model over the surface to consider CO<sub>2</sub> adsorption. At the inlet,  $C$  is known and remains constant over time. The outlet is governed by a zero diffusive flux condition for CO<sub>2</sub>. No mass flux is allowed through the sidewall except that made via reaction model.

For reaction, only simple one-step reaction is considered. The reaction rate is calculated at the wall surface for the reaction of CO<sub>2</sub> with amine. The following kinetic equations (4, 5) are taken as representative of several complex reactions:



$$r = k_{\text{langmuir}} C_{\text{CO}_2} \quad (5)$$

Equation (4) is the simplified representation of reactions, where capture of CO<sub>2</sub> occurs in the case of one primary amine group reacting with CO<sub>2</sub>. (Choi et al., 2009; Elfving and Sainio, 2021; Sanz-Pérez et al., 2016). In (4) R is the chain of atoms which is not participating in the reaction. In real physics, there is both forward and backward reactions, however, the model assumes that reaction occurs only in forward direction. Note that at moderate temperature of 25°C, assuming the irreversibility of the adsorption process is plausible if the sorbent has a strong affinity for CO<sub>2</sub>. Sorbent material properties used were same as in the adsorption modelling. Gas mixture inside the DAC-unit is estimated to contain only carbon dioxide (Elfving and Sainio, 2021). Equation (5) shows the relation between reaction rate  $r$  and concentration of reaction mixture components CO<sub>2</sub> ( $C_{CO_2}$ ). Also,  $r$  is the reaction rate based on Langmuir kinetic model and  $k_{langmuir}$  is the reaction kinetic constant. The Langmuir equation assumes that the adsorption of the gas to the sorbent is a reversible process and occurs only on a homogeneous surface with a fixed number of adsorption sites without considering the effects of the water vapor (one-step process). Since the purpose is to focus mainly on the adsorption of CO<sub>2</sub> during fast adsorption phase, selected single-step process model is considered to be sufficient for the analysis. Reaction kinetic constant  $k_{langmuir}$  is obtained from fitting simulation results to the experiments (Elfving and Sainio, 2021). The fitting results are specific to a given reaction and depend on temperature, humidity, and other reaction conditions. The unit of  $k$  depends on the order of the reaction, which can be determined experimentally. For the case of CO<sub>2</sub> adsorption, the rate of reaction is directly proportional to the concentrations of both CO<sub>2</sub> and the adsorbent. This means that increasing the concentration of sorbent or CO<sub>2</sub> will increase the rate of reaction.

### 3. RESULTS AND DISCUSSION

#### 3.1 Fluid flow, pressure drop, concentration.

The calculation of pressure drop is performed in various mesh sizes to verify the mesh independence. As seen from Fig. 2, we have used 12 meshes per reactive side of the hexagon(wall). In near wall, there is boundary layer and also surface reaction occurrence; therefore, it's necessary that meshes near wall is in layered shape and structured (hexahedral), while in the other regions only unstructured tetrahedral bigger meshes are enough. Table 1 shows the dependence of pressure drop on mesh resolution. For the coarse, fine and finer meshes, the number of nodes are 136421, 431597, 647372, respectively. The results for the pressure drop for different mesh sizes are shown in Table 1, which reveal the discrepancy in the pressure drop. The case with 12 mesh has the minimum difference with the neighboring number of meshes, and it can be chosen sufficient for the simulation. Flow is solved at steady state with 5000 iterations to ensure the convergence while the species transport model is off.

Six different inlet velocities, 0.004 m/s, 0.007 m/s, 0.01 m/s, 0.02 m/s, 0.1 m/s, 0.3 m/s are set for fixed hexagonal channel. The dependence of pressure drop-velocity is shown in Fig 3.

The results demonstrate a monotonic increase of pressure drop with velocity. Higher pressure drops mean a higher fan power requirement. In this work, length is kept constant although downsizing (shortening the channel length) can help with reducing the pressure drop.

Table 1. Mesh independence study.

Mesh per hexagon side(X)	Number of Nodes	Pressure drops per length (Pa/m)
8	136421(coarse)	0.5289
12	431597(fine)	0.5309
15	647372(finest)	0.5316

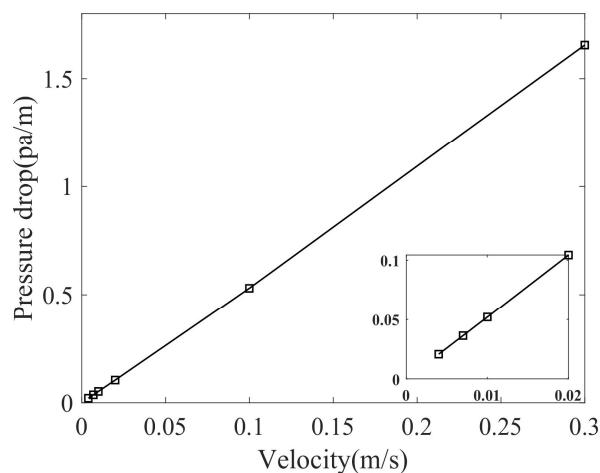


Fig. 3. Variation of pressure drop per unit length (Pa/m) with velocity(m/s). Hexagonal channel is 1 m. Air flow temperature is 25 °C and humidity is 2 %.

To understand the adsorption phenomenon inside hexagonal channel and its efficiency, it is essential to compare the results with other geometries. In previous works, adsorption performance of honeycomb structure is compared to the packed beds for different sorbents and gases and also for CO<sub>2</sub> adsorption. (Jänchen et al., 2015; Querejeta et al., 2022; Sakwa-Novak et al., 2016; Wajima et al., 2011). To this end, a simulation has been conducted also for the cylindrical packed bed system by the length of 1.77 cm and the diameter of 0.9 cm. This packed bed is consisted of particles with the diameter of 0.6 mm and particle volume fraction of 0.61. The working fluid temperature and humidity are taken as 25°C and 2%. The hexagonal channel length is 1.77cm, but its side length is allowed to have different values. Figure 4 represents the time variation of CO<sub>2</sub> concentration at the outlets of hexagonal unit and the packed bed during the adsorption process for the inlet velocity being as 0.13 m/s. Note that CO<sub>2</sub> adsorption on an

amine-based sorbent is fast. A key benefit of using hexagonal channels in a DAC contactor is the lower pressure drop, in comparison to the packed beds or the channels with other shapes of cross sections, while keeping the reactive area still high. Surface reactions serve as the main transport mechanism at this study. For qualitative comparison between hexagonal channels and packed beds, several hexagonal channels with various side lengths are considered in this study. Variation of the side length of hexagons affects the rate of adsorption and as seen in Fig. 4, for hexagon side  $X=1.1$  mm, the results for hexagonal channel and packed bed channel are pretty close. In future works, we will optimize the channel based on geometry and compare the system with similar-capacity packed bed systems. Figure 4 indicates that in an adsorption system whether open channels like the hexagonal one or a packed bed, the sorbent saturates after some time. For instance, for the hexagonal channel with hexagon side of  $X=1.1$  mm, as time advances to around 5000 seconds, the adsorption site starts to saturate which is similar to what happens in our studied packed bed (Nejadseifi et al., under preparation). Following this comparative study, we will demonstrate a wider study on hexagonal units having the length of 1 m, where we investigate the effects of velocity on adsorption process of CO<sub>2</sub>. The purpose is to investigate the performance of hexagonal reactor for various flow rates. At some point, since there is no experimental data for hexagonal cases, but for the packed bed with small size (length = 1.77 cm) experiments results were available and CFD simulation for packed bed fitted to them, while deviation was below 5% (Nejadseifi et al., under preparation; Elfving and Sainio, 2021). Therefore, for small scale hexagon with length of  $L=1.77$  cm, result of CFD simulations qualitatively compared with packed bed. And at the following, we have used the same reaction coefficients for big size hexagonal channel.

Figure 5 shows the CO<sub>2</sub> outlet concentration variation with time for large hexagonal channels with the length of 1 m for the hexagonal side of 0.58 cm at different velocities. As it is observed, increasing the velocity shortens the time to reach saturation. For the velocity of  $V=0.1$  m/s, the outlet CO<sub>2</sub> concentration begins to rise between the time  $0.5 \times 10^4$  and

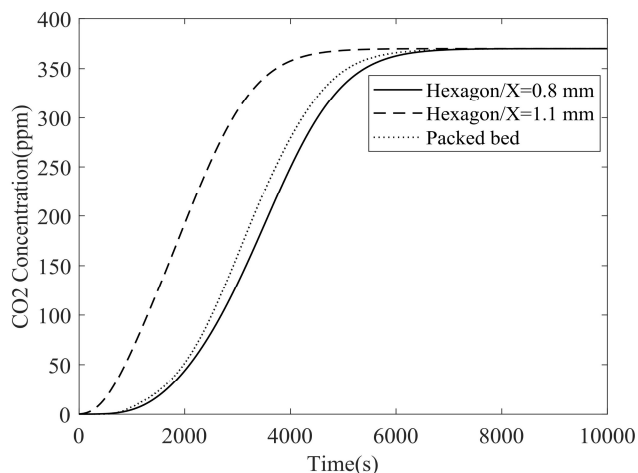


Fig. 4. CO<sub>2</sub> Concentration(ppm) variation at the outlet of hexagonal channel by time. Hexagonal channel and packed bed length is 1.77 cm. Air flow temperature is 25 °C and humidity is 2 vol-%. Velocity=0.13m/s.

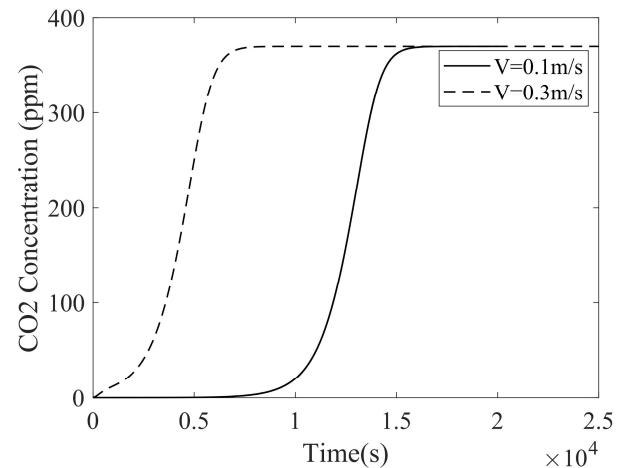


Fig. 5. CO<sub>2</sub> Concentration(ppm) change at the outlet of hexagonal channel by time. Hexagonal channel length is 1 m. Air flow temperature is 25 °C and humidity is 2 vol-%.

$1.0 \times 10^4$  s. This is the time where system moves toward full saturation which is achieved around  $\sim 1.0 \times 10^4$  s. For  $V=0.3$  m/s, the outlet concentration rises almost from the beginning of the adsorption until it reaches full saturation around  $t \sim 0.5 \times 10^4$  s.

### 3.2 Techno-economics of CO<sub>2</sub> capture in hexagonal channel

In designing a DAC system comprised of adsorption-desorption cycle in a monolithic contactor, air flow rate, as well as adsorption and desorption CO<sub>2</sub> concentration cut-off values (target CO<sub>2</sub> capture) play prominent roles in the final CO<sub>2</sub> capture cost. It should be mentioned that the focus of the current work is only on the adsorption stage. As the flow rate increases, more CO<sub>2</sub> is captured. However, higher flow rates are also associated with larger pressure drops that impose higher cost for consumed blowing power. Figures 6-8 present capture rates, demanded blowing (fan) power, and electricity consumption per ton of captured CO<sub>2</sub> for hexagonal channel contactor per unit length of channel at the side length of the hexagon as 0.58 cm. Air velocity ranges from 0.004 m/s to 0.3 m/s.

Figure 6 shows the variation of adsorption completion time (time for outlet concentration to reach 300 ppm) for a range of velocities between 0.004 and 0.3 m/s. Smaller panel inside Figs. 6-8 showed detailed values within small ranges of velocities. As it can be seen, increasing the velocity of air decreases the time of adsorption. However, experimental validations are required to confirm the performance of CO<sub>2</sub> adsorption at higher velocities.

In this study, the sensitivity analysis is aimed to obtain how the energy consumption related to air blow varies with air velocity as part of adsorption process, where the desorption stage is not considered. It can be noticed that in very small velocities, the saturation time increases exponentially as velocity decreases. Normally, very large saturation time



characterizes the adsorption process of DAC as inefficient or expensive from the energy consumption point of view.

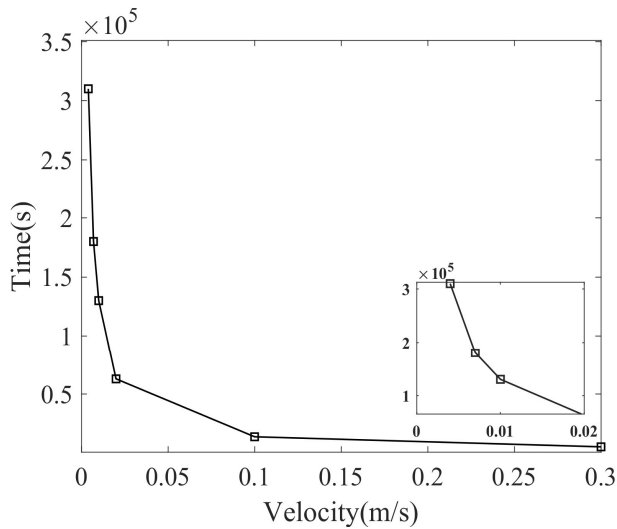


Fig. 6. Variation of adsorption reaction time (Time to reach 300 ppm) by velocity. Hexagonal channel length is 1 m. Air flow temperature is 25 °C and humidity is 2 vol-%.

The consumed blowing power cost per ton of captured CO<sub>2</sub> is estimated based on the assumption that the cost is proportional to the energy required in adsorption process. Based on the design of our DAC system, we assume that the energy is all provided by electricity. In the adsorption cycle, air is blown through the contactor and the contactor is kept at constant temperature (~25 °C). The blowing (fan) power for adsorption

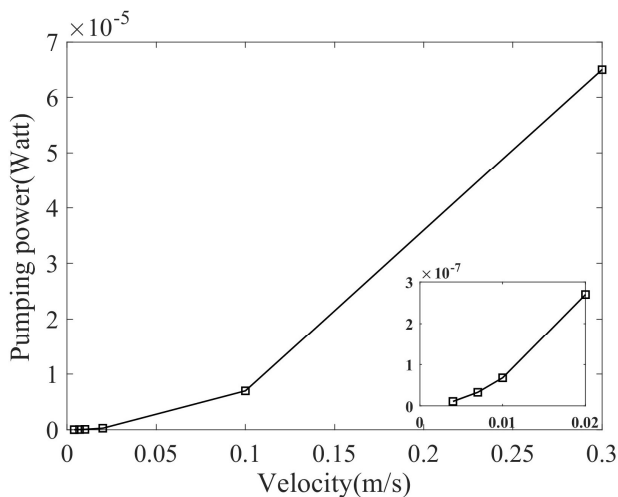


Fig. 7. Consumed blowing (fan) power vs. velocity. Hexagonal channel length is 1 m. Air flow temperature is 25 °C and humidity is 2 vol-%.

is therefore estimated by (6):

$$P = \frac{Q \cdot \Delta p}{\eta} \quad (6)$$

where  $P$  is the electrical power consumption during each adsorption cycle,  $\Delta P$  is the air pressure drop for each hexagon, and  $Q$  is the volumetric flow rate of the air. The efficiency  $\eta$  is

the effectiveness of the system in converting electrical energy into mechanical energy to move the air, which is considered as a fixed value of 0.8. Figure 7 shows the variation of consumed power by velocity. For both higher ranges of velocities  $0.1 \leq V \leq 0.3$  m/s and lower ranges  $0.004 \leq V \leq 0.02$  m/s, we see a non-linear relation between the blower (fan) power and Velocity. Then, the total electricity consumption per ton of adsorbed CO<sub>2</sub> per each cycle is calculated by (7):

$$E = \frac{P \cdot t}{1000 \cdot m_{CO_2}} \quad (7)$$

where  $P$  is the electrical power consumption,  $t$  is the adsorption time of each cycle (the time for CO<sub>2</sub> concentration to reach 300 ppm). Obviously, it is not efficient to run the system after saturation. Also,  $m_{CO_2}$  is the amount of CO<sub>2</sub> adsorbed during each cycle. Figure 8 shows the electricity consumption per unit mass of CO<sub>2</sub> for various velocities, which displays the same trend as in Fig. 7 except for the lower range of velocity that turns to be linear.

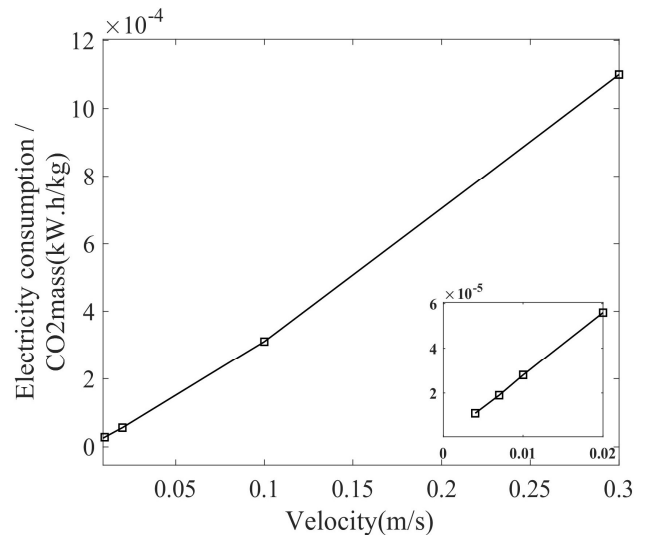


Fig. 8. Electricity consumption per / Adsorbed mass of CO<sub>2</sub>.vs. velocity. Hexagonal channel length is 1 m. Air flow temperature is 25 °C and humidity is 2 vol-%.

#### 4. CONCLUSIONS

Recently, CO<sub>2</sub> removal from air, also known as the Direct Air Capture (DAC), has attracted more attention due to its promise in reducing the greenhouse gas (GHG) emissions. However, costs associated with DAC have been the main obstacle for large-scale commercialization. Here, we presented a novel hexagonal shaped contactor and investigated performance of adsorption under various flow rates. Direct air flow inside straight channel is more commercialized with a higher demand because of its lower pressure drop, which makes it practical for large scale applications. Such a straight channel application is supposed to be more economic than non-straight channels or packed bed reactors. The techno-economic analysis showed how the adsorption stage is dependent of various air velocities. Focus of this work was on lower ranges of velocities ( $V \leq 0.3$



m/s); where the flow regime remained laminar; This study could be repeated for higher ranges of velocities at the ranges of velocities ( $V \geq 2$  m/s) at the future. As future work, we also perform experiments on the hexagonal channel contactor, and examine pressure drop, wall thickness effects, and hexagons side length. We will focus more on novel numerical methods and investigate using novel contactors shapes.

This is a preliminary study to shed light on the issue. Existing set of results are not covering a broad range, and we know that. However, most of the previous works are laboratory-scale experimental study and maneuver on the different aspects including fluid mechanic parts or technoeconomic. However, since types of studied barely done earlier, there is big differences on data reported by different authors, especially when it comes to cost estimation. Importance of these results are somehow can be shown on the cost dependence on pressure drop and air volume flow rate. At the future, if desorption phase is studied beside adsorption phase, this hexagonal geometry can be optimized for minimal consumption of electrical energy for unit of adsorbed CO<sub>2</sub>.

## REFERENCES

- Choi, S., Drese, J. H., Eisenberger, P. M., and Jones, C. W. (2011). Application of amine-tethered solid sorbents for direct CO<sub>2</sub> capture from the ambient air. *Environmental Science and Technology*, 45(6), 2420–2427.
- Cornejo, I., Nikrityuk, P., and Hayes, R. E. (2022). Heat and mass transfer inside of a monolith honeycomb: From channel to full size reactor scale. *Catalysis Today*, 383, 110–122.
- Deutschmann, O., Schwiedemoch, R., Maier, L. I., and Chatterjee, D. (2001). Natural gas conversion in monolithic catalysts: interaction of chemical reactions and transport phenomena. In *Studies in surface science and catalysis* (Vol. 136, pp. 251–258). Elsevier.
- Fu, D., and Davis, M. E. (2023). Toward the feasible direct air capture of carbon dioxide with molecular sieves by water management. *Cell Reports Physical Science*, 4(5).
- Hawthorn, R. D. (1974). Afterburner catalysts-effects of heat and mass transfer between gas and catalyst surface.
- Jiang, L., Liu, W., Wang, R. Q., Gonzalez-Diaz, A., Rojas-Michaga, M. F., Michailos, S., Pourkashanian, M., Zhang, X. J., and Font-Palma, C. (2023). Sorption direct air capture with CO<sub>2</sub> utilization. *Progress in Energy and Combustion Science*, 95, 101069.
- Klenov, O. P., Pokrovskaya, S. A., Chumakova, N. A., Pavlova, S. N., Sadykov, V. A., and Noskov, A. S. (2009). Effect of mass transfer on the reaction rate in a monolithic catalyst with porous walls. *Catalysis Today*, 144(3–4), 258–264.
- Kreutzer, M. T., Du, P., Heiszwolf, J. J., Kapteijn, F., and Moulijn, J. A. (2001). Mass transfer characteristics of three-phase monolith reactors. *Chemical Engineering Science*, 56(21–22), 6015–6023.
- Mazzotti, M., Baciocchi, R., Desmond, M. J., and Socolow, R. H. (2013). Direct air capture of CO<sub>2</sub> with chemicals: optimization of a two-loop hydroxide carbonate system using a countercurrent air-liquid contactor. *Climatic Change*, 118, 119–135.
- McQueen, N., Gomes, K. V., McCormick, C., Blumanthal, K., Pisciotta, M., and Wilcox, J. (2021). A review of direct air capture (DAC): scaling up commercial technologies and innovating for the future. *Progress in Energy*, 3(3), 32001.
- Raja, L. L., Kee, R. J., Deutschmann, O., Warnatz, J., and Schmidt, L. D. (2000). A critical evaluation of Navier–Stokes, boundary-layer, and plug-flow models of the flow and chemistry in a catalytic-combustion monolith. *Catalysis Today*, 59(1–2), 47–60.
- Rodríguez-Mosqueda, R., Bramer, E. A., and Brem, G. (2018). CO<sub>2</sub> capture from ambient air using hydrated Na<sub>2</sub>CO<sub>3</sub> supported on activated carbon honeycombs with application to CO<sub>2</sub> enrichment in greenhouses. *Chemical Engineering Science*, 189, 114–122.
- Sanz-Pérez, E. S., Murdock, C. R., Didas, S. A., and Jones, C. W. (2016). Direct capture of CO<sub>2</sub> from ambient air. *Chemical Reviews*, 116(19), 11840–11876.
- Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6), 1704–1709.
- Thakkar, H., Eastman, S., Hajari, A., Rownaghi, A. A., Knox, J. C., and Rezaei, F. (2016). 3D-printed zeolite monoliths for CO<sub>2</sub> removal from enclosed environments. *ACS Applied Materials and Interfaces*, 8(41), 27753–27761.
- Tischer, S., and Deutschmann, O. (2005). Recent advances in numerical modeling of catalytic monolith reactors. *Catalysis Today*, 105(3–4), 407–413.
- Verougstraete, B., Martin-Calvo, A., der Perre, S., Baron, G., Finsy, V., and Denayer, J. F. M. (2020). A new honeycomb carbon monolith for CO<sub>2</sub> capture by rapid temperature swing adsorption using steam regeneration. *Chemical Engineering Journal*, 383, 123075.
- Wang, W., Wang, S., Ma, X., and Gong, J. (2011). Recent advances in catalytic hydrogenation of carbon dioxide. *Chemical Society Reviews*, 40(7), 3703–3727.
- Zhongming, Z., Linong, L., Xiaona, Y., Wangqiang, Z., Wei, L. (2021). *Global energy review 2021*. IEA: Paris, France.
- Nejadseifi, M., Kiani-Oshtorjani, M., Tynjälä, T., and Jalali, P. (2024). Permeability of partially to fully randomized arrays of square disks: A lattice Boltzmann approach. *Journal of Porous Media*, 27(3), 85–99.
- Nejadseifi, M., Tynjälä, T., Jalali, P., and Jere, Elfving. (under preparation). Adsorption of CO<sub>2</sub> by direct air capture in particle packed beds of sorbent: Direct simulation

approach. Under Preparation for Submission to Journal 'powder Technology'.

Jänchen, J., Herzog, T. H., Gleichmann, K., Unger, B., Brandt, A., Fischer, G., and Richter, H. (2015). Performance of an open thermal adsorption storage system with Linde type A zeolites: Beads versus honeycombs. *Microporous and Mesoporous Materials*, 207, 179–184.

Querejeta, N., Rubiera, F., and Pevida, C. (2022). Experimental study on the kinetics of CO<sub>2</sub> and H<sub>2</sub>O adsorption on honeycomb carbon monoliths under cement flue gas conditions. *ACS Sustainable Chemistry and Engineering*, 10(6), 2107–2124.

Sakwa-Novak, M. A., Yoo, C.-J., Tan, S., Rashidi, F., and Jones, C. W. (2016). Poly (ethylenimine)-functionalized monolithic alumina honeycomb adsorbents for CO<sub>2</sub> capture from air. *ChemSusChem*, 9(14), 1859–1868.

Wajima, T., Munakata, K., Takeishi, T., Hara, K., Wada, K., Katekari, K., Inoue, K., Shinozaki, Y., Mochizuki, K., Tanaka, M., (2011). Adsorption characteristics of water vapor on honeycomb adsorbents. *Journal of Nuclear Materials*, 417(1–3), 1166–1169.

Choi, S., Drese, J. H., and Jones, C. W. (2009). Adsorbent materials for carbon dioxide capture from large anthropogenic point sources. *ChemSusChem: Chemistry and Sustainability Energy and Materials*, 2(9), 796–854.

Elfving, J., and Sainio, T. (2021). Kinetic approach to modelling CO<sub>2</sub> adsorption from humid air using amine-functionalized resin: Equilibrium isotherms and column dynamics. *Chemical Engineering Science*, 246, 116885.

Sanz-Pérez, E. S., Murdock, C. R., Didas, S. A., and Jones, C. W. (2016). Direct capture of CO<sub>2</sub> from ambient air. *Chemical Reviews*, 116(19), 11840–11876.

# Computational analysis of conjugate heat transfer in a 2D rectangular channel with mounted obstacles using lattice Boltzmann method

Majid Nejadseifi\*. Shervin Karimkashi\*\*. Tero Tynjälä\*. Payman Jalali\*

\*Laboratory of Thermodynamics, School of Energy Systems, Lappeenranta-Lahti University of Technology, Lappeenranta, 53850, Finland (Tel: 358-415750813; e-mail: [majid.nejadseifi@lut.fi](mailto:majid.nejadseifi@lut.fi), [tero.tynjala@lut.fi](mailto:tero.tynjala@lut.fi), [payman.jalali@lut.fi](mailto:payman.jalali@lut.fi)).

\*\*Laboratory of Energy conversion and systems, Department of Mechanical Engineering, School of Engineering Sciences, Aalto university, Espoo, 02150, Finland (Tel: +358505602614; e-mail: [shervin.karimkashiarani@aalto.fi](mailto:shervin.karimkashiarani@aalto.fi))

**Abstract:** The objective of this paper is to investigate the fluid flow and conjugate heat transfer in a 2D channel using lattice Boltzmann method (LBM). In this work, fluid flow and heat transfer are studied for the Reynolds numbers varying between 250 and 1000. The working fluid in the simulations is air with the Prandtl number of 0.72. At the Reynolds number of 600, the effect of different conductivity ratio (1, 10, 100, 400) between solid and fluid are investigated. Furthermore, at this Reynolds number, the distance between obstacles for the conductivity ratio of 10 is evaluated. The results show that any increase in Reynolds number leads to a heat transfer improvement. Moreover, increase in the conductivity ratio leads to an isothermal surface and enhanced heat transfer. The more the distance between the obstacles, the better the heat transfer rate. The results obtained from LBM are in good agreement with experimental and conventional computational fluid dynamics methods.

**Keywords:** Lattice Boltzmann Method (LBM), Conjugate heat transfer, Nusselt number, Prandtl number

## 1. INTRODUCTION

In recent years, the computer technology and e-commerce have significantly progressed. Over time, electronic components have become more compact and occupy smaller space. As the components are more compact, the processing speed has increased considerably. Due to the smaller space occupied by the components, temperature resistance is increased against heat transfer and the performance has dropped with increasing temperatures in electronic equipment.

In electronic cooling, forced convection is often the dominant heat transfer mode, with the cooling agent being either a common gas or a heat-transferring liquid. Various techniques have been proposed to enhance the heat transfer rate between the solid electronic devices and the adjacent cooling fluid, and their thermal performance has been evaluated. Among these techniques, the use of solid fins Chen et al. (1997) and blocks Sara et al. (2001) have proven to be efficient. Additionally, Ramesh et al. (2021) has done an extensive review of various numerical and experimental studies that have investigated methods for enhancing heat transfer in cooling devices.

Chikh et al. (1998) conducted a numerical investigation on forced convection heat transfer within a partially heated channel, focusing on the impact of porous obstacles installed on the heated section to enhance the heat transfer rate. Their results demonstrated that using porous obstacles under certain flow and thermal conditions lead to a 90% drop in wall temperature.

There have been many studies on the heat transfer of the extended surfaces in channels. Bhowmik et al. (2009)

examined pressure loss and heat transfer in a channel with two bending blades, and provided connections for a laminar, transient and turbulent flow in a marginal geometry. Improvement of heat transfer for equally spaced plates was done experimentally by Leung et al. (1999). They examined the parameters including the diameter of the slopes, the geometry and height of the cavities, and the Reynolds number. They concluded that for the ducts, heat transfer decreases with increasing the diameter of the duct due to the decrease of the surface. The effect of cubic layout on turbulent flow was experimentally investigated by Meinders et al. (2002). They concluded that, with increasing flow velocity, fluid flow would be affected by the distance between extended surfaces. Nazari et al. (2013) investigated a heat transfer problem in a closed compartment with a vertical or horizontal porous layer using LBM.

Numerical simulations that use conventional computational fluid dynamic (CFD) techniques, such as finite volume and finite difference methods, are difficult to apply in complex boundary conditions, for instance in heat transfer problems at the interface between the fluid and solid instead of a constant temperature or heat flux condition. Using such a conjugate boundary condition in CFD methods will increase the computational cost. Therefore, it is required to use a less expensive method. One of the most useful methods for this kind of problems is the Lattice Boltzmann Method (LBM).

LBM is a reliable approach for studying numerous fluid and heat transfer issues, including multi-component and multiphase flows, microflows, turbulent flows, fluid-solid interactions, and both forced and natural convection in

complex geometries. Based on the kinetic theory of gases with a mesoscopic approach, LBM is widely used due to its effectiveness in handling physical problems with intricate geometries and boundary conditions. Consequently, it is frequently employed for solving fluid flow problems in porous media. Ataei-Dadavi et al. (2019) studied fluid flow and heat transfer in a porous cavity with side heating. Results showed that heat transfer decreased with porous media compared to a non-porous cavity. They developed a new method to predict Nusselt numbers for such cavities. Nejadseifi et al. (2024) used the lattice Boltzmann method to computationally study porous media composed of monodisperse square obstacles within the Darcy regime. Mirahsani et al. (2023) investigated heat transfer enhancement in a partially heated channel with porous obstacles. Using a lattice Boltzmann method, the flow and temperature fields were analyzed. Optimal design parameters, including the height, pitch, and permeability of the porous blocks, were determined. Results showed that the optimal block height and pitch depend on the Darcy number (Da). The impact of obstacle geometry and porous material distribution on thermal performance was also examined. Matsuda et al. (2024) enhanced enthalpic lattice Boltzmann method (LBM) to simulate conjugate heat transfer in non-homogeneous media with time-dependent thermal properties. Their findings highlighted the potential of the modified LBM for simulating complex heat transfer in a non-homogeneous media and in optimizing heat exchanger designs. Paknahad et al. (2023) investigated pore-scale flow and conjugate heat transfer in high-porosity open-cell metal foams. They showed that lower porosity foams transition from the Darcy to non-Darcy flow regime at lower Reynolds numbers, while pore density has minimal impact on this transition. Heat transfer results indicated that metal foams cool rapidly at high flow velocities, with minimal temperature rise in the fluid.

Although LBM has been successfully applied to simulating fluid flows in small-scale channels, there are few reports of using LBM to simulate fluid-solid coupling heat transfer in such channels. In this study, the fluid flow and heat transfer within the channel with extended surfaces is investigated and lattice Boltzmann equations are used for simulations. The simulation setup is a channel with insulated internal surfaces, wherein heat transfer occurs through extended surfaces; In this setup, effects of changing parameters such as Reynolds number, conduction coefficient ratio and obstacles distance are investigated. It is assumed that there is a constant heat flux at the bottom of the obstacles. Validation of the results shows a good agreement with other works.

## 2. METHODOLOGY

### 2.1 Lattice Boltzmann Method

In recent years, LBM has been developed as a powerful simulation method to simulate many fluid mechanic problems. This method, based on the kinetics of gases theory, has been considered as a powerful numerical technique for simulating fluid flow and heat transfer (Paknahad et al., 2023, Ramesh et al., 2021) Compared to the conventional CFD methods, this method has several advantages. The conventional CFD methods discretize mass, momentum, and energy equations, and solve them based on macroscopic quantities such as

velocity and pressure. Unlike CFD, LBM uses the number of finite speeds created on a regular lattice to solve problems (Noble et al., 1996).

### 2.2 Mathematical description of fluid flow and heat Transport Equations:

In contrast with the conventional macroscopic Navier–Stokes approach, the lattice Boltzmann method utilizes a mesoscopic simulation model to represent fluid flow (Suss et al., 2023). This technique is based on modeling the movement of fluid particles in different directions to derive macroscopic fluid characteristics, such as velocity and pressure. In this method, the fluid domain is divided into uniform cells, each containing a set number of Distribution Functions that represent fluid particles movement in specific discrete directions. Various LBM models are available depending on the dimensionality(2D-3D) and the number of velocity directions. This study focuses on two-dimensional 2D flow using a 2D square lattice with nine velocities, known as the D2Q9 model. The velocity vectors of the D2Q9 model are labeled as  $c_0$  to  $c_8$ . The discrete velocities of the D2Q9 model are:

$$e_i = \begin{cases} (0,0) & i = 0 \\ (\pm 1,0)c, (0, \pm 1)c & i = 1, 2, 3, 4 \\ (\pm 1, \pm 1)c & i = 5, 6, 7, 8 \end{cases} \quad (1)$$

Under equilibrium conditions, macroscopic quantities can be measured before allowing the particles to move and collide in the next time step. In standard single relaxation time (SRT) lattice Boltzmann (LB) models, the BGK (Bhatnagar-Gross-Krook) model (Zou and He, 1997) is used to represent collisions, as described in Eq. (2). In this equation,  $f_i$  denotes the density distribution function in direction  $i$ , and  $e_i$  represents the discretization direction. Additionally,  $\Delta t$  is the LBM time step, and  $\tau$  is the dimensionless relaxation time provided in Eq. (3).

$$f_i(x + e_i \Delta t, t + \Delta t) = f_i(x, t) - \frac{\Delta t}{\tau} [f_i(x, t) - f_i^{eq}(x, t)] \quad (2)$$

where:

$$\tau = 3\nu + \frac{1}{2} \quad (3)$$

where  $c = \frac{\Delta x}{\Delta t}$  and  $i$  is an index for directions. The equilibrium distribution function  $f_i^{eq}(x, t)$  is calculated as (4):

$$f_i^{eq}(x, t) = \omega_i \rho \left( 1 + \frac{u \cdot e_i}{c_s^2} + \frac{(u \cdot e_i)^2}{2c_s^4} - \frac{u^2}{2c_s^2} \right) \quad (4)$$

In the chosen D2Q9 model, the speed of sound  $c_s$  is equal to  $\sqrt{3}$  in the specified value. Additionally,  $\omega_i$  is the weight function represented by (5) as:

$$\rho = \sum_{i=0}^8 f_i, \quad V = \frac{1}{\rho} \sum_{i=0}^8 f_i e_i \quad (5)$$

$$\omega_i = \begin{cases} 4/9 & i = 0 \\ 1/9 & i = 1, 2, 3, 4 \\ 1/36 & i = 5, 6, 7, 8 \end{cases} \quad (6)$$

In summary, the SRT-LBM process includes three main steps: (1) calculating the distribution function within the simulation domain, (2) computing collisions at each time step, and (3) streaming, or transferring the distribution function to neighboring nodes (Nejadseifi et al.,2024).

The distribution function for the thermal LBM can be represented as:

$$g_i(x + e_i \Delta t, t + \Delta t) = g_i(x, t) - \frac{\Delta t}{\tau} [g_i(x, t) - g_i^{eq}(x, t)] \quad (7)$$

The corresponding equilibrium distribution functions for fluid and solid are defined as follows (Mohammad, 2007):

$$g_i^{eq}(x, t) = \omega_i T \left( 1 + \frac{u \cdot e_i}{c_s^2} \right) \quad (8)$$

$$g_i^{eq}(x, t) = \omega_i T \quad (9)$$

Finally, for both solid and fluid, the temperature field is computed as:

$$T = \sum_{i=0}^8 g_i \quad (10)$$

### 3. RESULTS AND DISCUSSION

The physical geometry used for calculations is shown in Fig. 1. The fluid flow is assumed laminar, incompressible, viscous, and Newtonian. The length of the entrance to the channel ( $L_{in}$ ) before the obstacles is 6 times the obstacle’s height. The length of the outer part of the channel, which is extended after the second obstacle ( $L_{out}$ ), has been selected to be long enough for vortices observation. For this reason, the length of the outlet portion of the channel, which is located after the second obstacle, is  $8H$  where  $H$  is the height of the channel.

The bounce back scheme is used for representing the wall boundary condition. In the bounce back scheme, it is assumed that the particles moving towards solid boundaries are returned into the fluid. Therefore, the distribution functions that are toward the wall bounced back in the opposite direction. The fully developed inlet velocity profile is set as  $u_0(y) = 4u_{max}(Hy - y^2)/H^2$ , and a negligible compressible flow is assumed (the maximum velocity is 0.1 in the input current). The unknown distribution functions in the input and output sections of the channel are obtained using the Zou and He.

(1997) boundary function. The adiabatic boundary condition is employed on the walls. However, for the parts of the walls which include obstacles, non-zero temperature gradient and constant heat flux boundary conditions govern at the bottom part of the obstacle ( $q=1$ ). On the boundaries between solid obstacles and fluid, the equilibrium distribution function, as well as the conductivity values for the fluid and solid, is changed. At the interface of the solid obstacles and fluid, the continuity of heat flux and temperature is naturally satisfied.

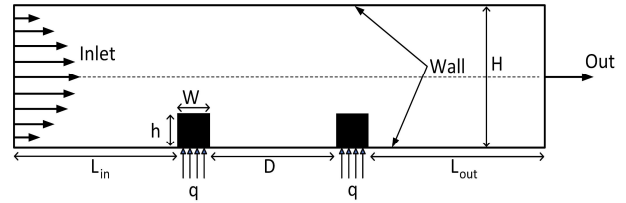


Fig 1. Schematic diagram of the problem geometry.

#### 3.1 Grid independence test and code validation

The numerical simulations are carried out by a FORTRAN in-house code. To ensure the accuracy of the written code, a comparison is made between different studies to ensure a dimensionless reattachment study ( $Xr/L$ ). These results are compared with the results of CFD(ANSYS-Fluent 2021 R2) software, (Korichi and Oufer, 2005; Pirouz et al., 2011). The fluid flow at the channel entrance has parabolic state and Reynolds number, which varies between 250 and 1000. Fig. 2 shows the dimensionless reattachment length for Reynolds numbers below 1000. The results show good agreement with other studies.

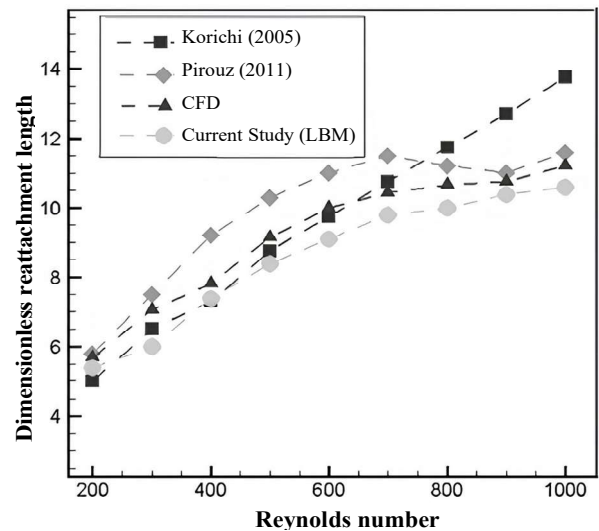


Fig. 2. Dimensionless reattachment length comparison for various Reynolds Number for a single obstacle.

The second validation is for grid independency. Validation is done for Reynolds numbers 500 and 750. The number of lattices required for code convergence is considered. (The channel width is 96, 128, and 160 lattice). The mean Nusselt number on the two obstacles for these lattices are compared with each other. Table 1 shows that for the range of  $Re$  that



have been studied, with 160 points across the channel, acceptable results are achieved. As  $Re$  increases, the lattice size needs to be finer: for example, when the  $Re$  reaches 500, the difference in the results for 96 lattice approaches 2% and this lattice is no longer suitable for such  $Re$  value, and that is why the number of grids needs to be higher (160 lattice).

Table 1. Effect of resolution on the mean Nusselt Number at different Reynolds numbers

Obstacle	Re	Mean Nusselt number		
		96 Grid	128 Grid	160 Grid
1	500	9.82 (1.44%)	9.91(0.6%)	9.97
2	500	8.08 (1.92%)	8.17(0.8%)	8.24
1	750	11.25 (1.83%)	11.33 (1.16%)	11.5
2	750	10.14 (2.42%)	10.23 (1.52%)	10.40

The third validation is associated with the thermal part of the written code. The flow inside the channel with 3 obstacles within it is considered and compared with the results of Korichi and Oufer. (2005). These results are obtained for  $Re=400$ . The lattice considered is square and the channel's height is divided into 160 cells. Here, the Nusselt curve has been calculated and plotted on extended surfaces. As shown in Fig. 3, the results are in good agreement with Korichi's work.

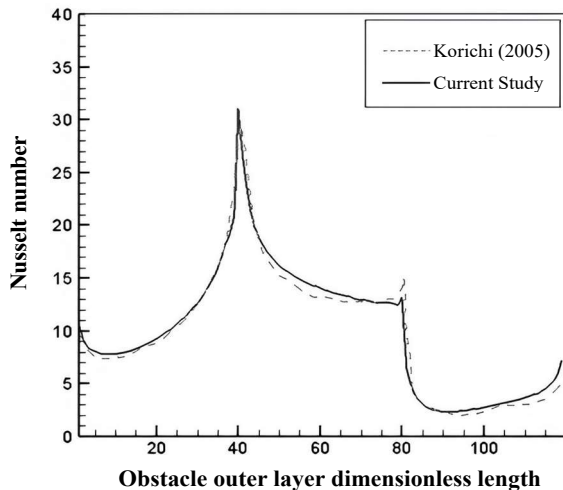


Fig. 3. Validation of the local Nusselt number distribution on obstacle Surface ( $Re=400$ ).

### 3.2 Fluid flow and heat transfer analysis

Conjugate heat transfer, especially in complex and small sized geometries, strongly depends on the flow in the desired location. In such cases, low  $Re$  is expected because of low velocity and small dimensions of the channel. In the small

sized channels where laminar flow regime exists, the viscosity leads to flow separation and circulation.

In this study, the effect of flow velocity and other parameters on heat transfer have been investigated; the intensity of heat transfer depends on the local  $Nu$  number based on the length of the obstacle. The coefficient of heat transfer between solid and fluid is equal to 1,10,100,400  $W/m^2K$ . The flow velocity is proportional to  $Re$  number, which is based on the channel height and varies between 250 and 1000. The fluid flow is in the laminar regime. The inlet fluid velocity varies from 0.3 to 5  $m/s$  which maintains forced convective heat transfer (Anderson and Moffat, 1992).

#### 3.2.1 Effect of $Re$ number

The  $Re$  number is intended to be 250, 500,750 and 1000. At a low  $Re$ , flow is stable, and the velocity fluctuations decrease over a certain period of time to reach a constant velocity. In Fig. 4(a), the velocity fluctuations are plotted over time at  $y = 0.5$  and  $x = 9$ . When  $Re$  increases, the fluid enters a transient state. When this happens, the fluctuation that occurs in the fluid decreases over time to reach a stable value. The velocity variations are shown in Fig. 4(b) for two values of  $Re$ .

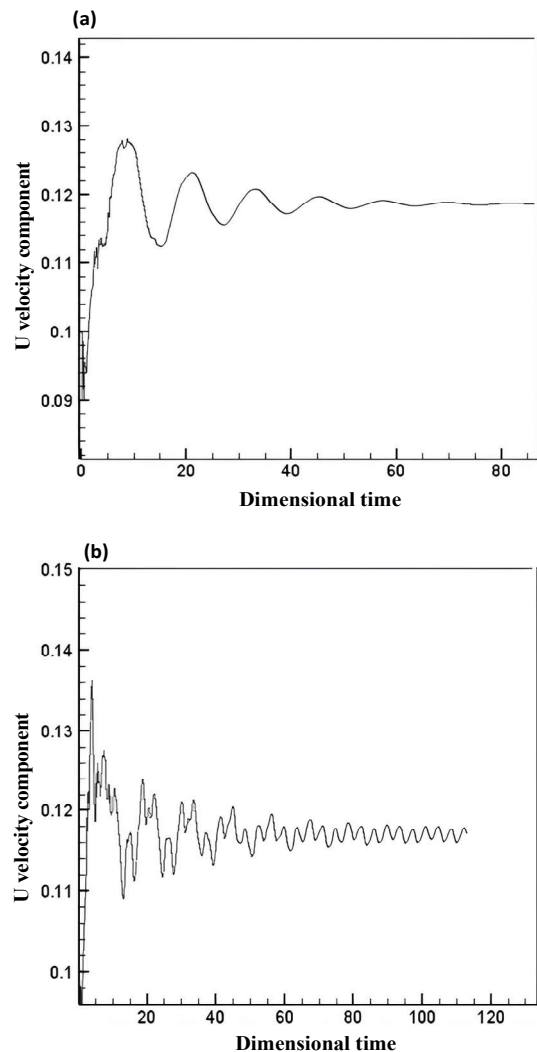


Fig. 4. Changes in the Horizontal component of velocity by time at  $y = 0.5$  and  $x = 9$  for a)  $Re = 250$  and b)  $Re = 1000$ .

As fluid flows in the channel, vortices are generated behind the obstacles. At lower Reynolds number, these generated vortices remain fixed behind the obstacles. When Reynolds number exceeds a certain value, due to the increase of the cross-section (which occurs after the second obstacle) at the downstream and due to a sudden expansion, a large vortex is generated and moves forward. In this case, the flow regime changes from laminar to transient. In transient flows, these small vortices that form behind the obstacles originate, grow, and move forward in the flow direction. The order of magnitude for the velocity inside these formed vortices is 2 or 3 times less than the flow in the center of the channel. These streamlines are shown in Fig. 5. Isotherm lines are presented for various Reynolds number in Fig. 5. The dimensions of the computational domain are very large in comparison with dimensions of the obstacles. For this reason, in order to see the isothermal lines and streamlines around the obstacles, only part of the simulation domain which includes obstacles is shown in Fig. 5.

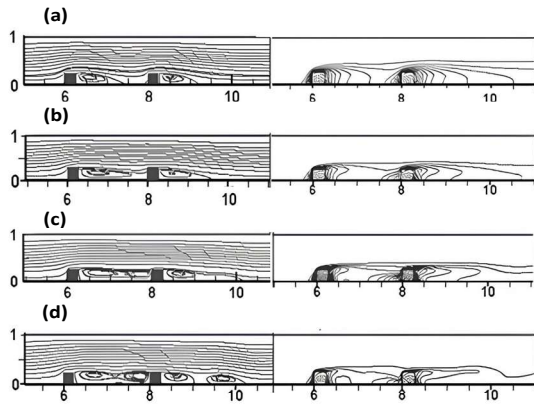


Fig. 5. Snapshot of variations of Streamlines and isothermal lines in terms of Reynolds number for  $k_s / k_f = 10$  and Reynolds (a) 250. (b) 500 (c) 750, (d) 1000.

The heat transfer rate is characterized by the local Nusselt number based on the obstacle's outer layer dimensionless length. Time-averaged dimensionless temperature and velocity components are calculated over the simulation domain. Local and mean Nusselt numbers and dimensionless temperature are formulated as (11) and (12):

$$Nu_x = \frac{-1}{\theta_m} \frac{\partial \theta}{\partial n} \tag{11}$$

$$\theta = \frac{T - T_0}{qH/k_f} \tag{12}$$

Fig. 6 shows the variation of Nusselt number against the length of the outer surface of the obstacle for different Reynolds numbers. As seen in these figures, the Nusselt increases with increasing Re. In the lower left corner, due to the rotation of the flow and the proximity of the inlet heat flux, the Nusselt number has a relative maximal value. Near the upper left corner on the obstacle, the increases in the temperature gradient of the fluid ( $|\partial \theta / \partial n|$ ) is due to the increase in

momentum. For the obstacle left face ( $0.25 > \text{obstacle outer surface length} > 0$ ), and before the sudden increase of the Nusselt in the upper left corner, Nusselt has a minimum relative value. For the upper face of the obstacle ( $0.5 > \text{obstacle outer surface length} > 0.25$ ), the temperature gradient value is higher than the obstacle left and right face. This is due to the higher flux and, in turn, the reduction of the thickness of the boundary layer. For the obstacle right face ( $0.75 > \text{obstacle outer surface length} > 0.5$ ), the Nusselt increases slightly with the increase of the Reynolds number, and this increase is negligible compared to the increase in the Nusselt in the upper and left face. It is important to mention that the trend of Nusselt variation versus the length of the outer surface of the obstacle remains unchanged with Re within its range between 250 and 1000.

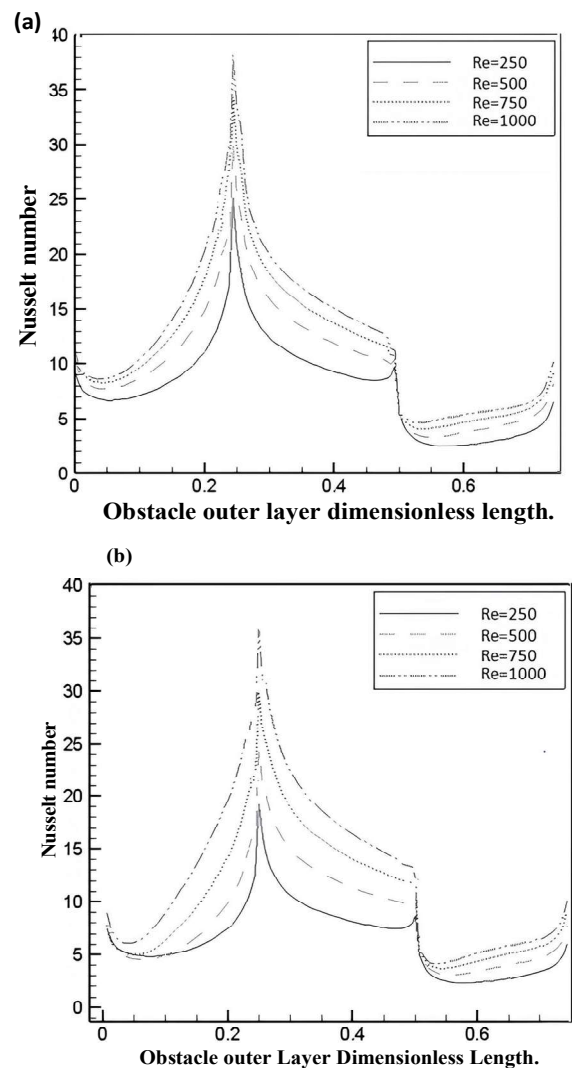


Fig. 6. Time-averaged variation of the Nusselt number in terms of the outer surface of the obstacle for  $k_s / k_f = 10$ . a) first obstacle b) second obstacle

### 3.2.2 Effect of conductivity ratio

One of the parameters that plays an important role in heat transfer is the conductivity ratio between the solid and the fluid

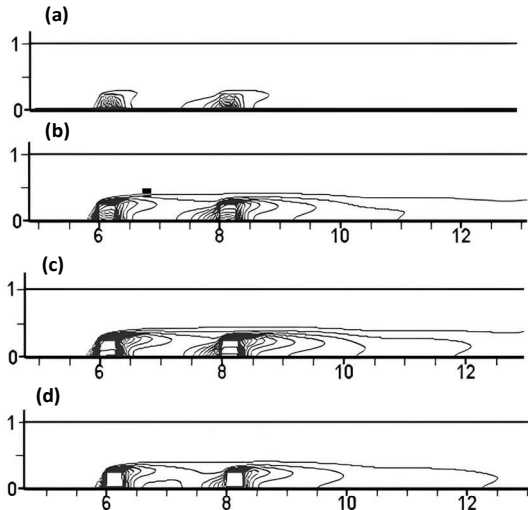


Fig 7. Variation of time-averaged isotherms for different Solid-Fluid thermal conductivity ratio for  $Re = 600$ . a)  $k_s / k_f = 1$  b)  $k_s / k_f = 10$  c)  $k_s / k_f = 100$  d)  $k_s / k_f = 400$

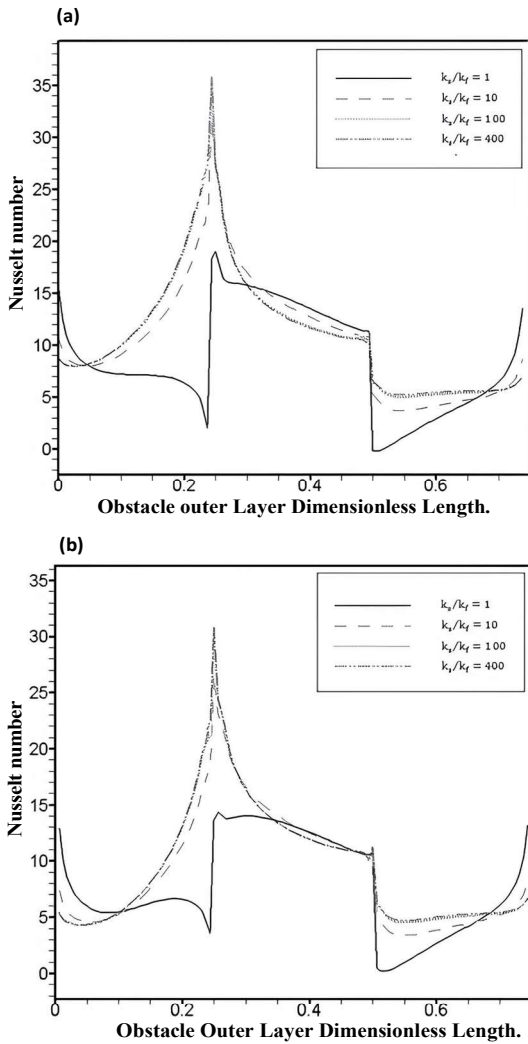


Fig. 8. Time-averaged variation of Nusselt in terms of the outer surface of the obstacle for the ratio of different Solid-Fluid thermal conductivity ratio for  $Re = 600$ . a) first obstacle b) second obstacle

( $k_s / k_f$ ). By increasing the conductivity ratio, the internal resistance against thermal flux decreases, which also reduces the maximum temperature and reduces the internal temperature gradient. When the conductivity order of magnitude for solid obstacles is 2 or 3 times larger than the conductivity of fluid, the obstacle behaves as an isothermal object. Fig. 7 shows the dimensionless isothermal curves. Here,  $Re$  is fixed and equal to 600. The thermal conductivity ratio between solid and fuel ( $k_s / k_f$ ) is set to 1, 10, 100, and 400. For ( $k_s / k_f$ ) = 1, the density of the isothermal curves for both the solid and fluid is identical. When this ratio increases, gradually the temperature of the obstacle decreases, and the obstacle itself behaves as an isothermal object.

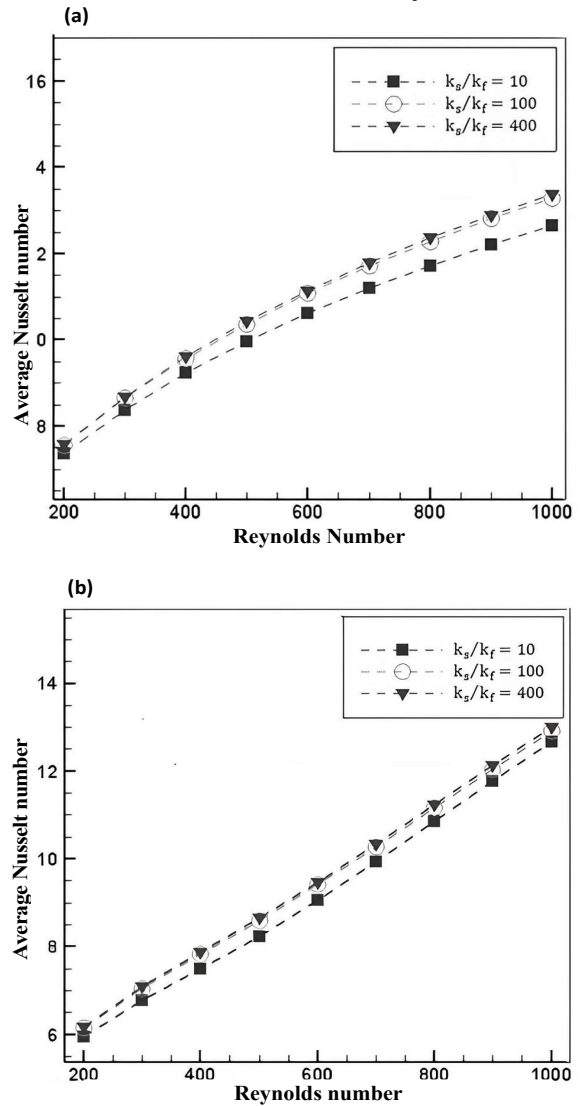


Fig. 9. Time-averaged variations of the average Nusselt number with variation of the Reynolds number for  $k_s / k_f = 10, 100, 400$ . a) first obstacle b) second obstacle

Fig. 8 shows the distribution of local Nusselt on extended surfaces. When the ratio  $k_s / k_f = 1$ , the behavior of Nusselt is completely different from when this ratio is higher; and in this case, the obstacle behaves such as heat insulator. When the thermal conductivity ratio of the solid and the fluid increases,

the difference in the surface temperature of the obstacle is almost negligible. For a situation in which this ratio is greater than 100, the distribution of Nusselt does not change, and solid obstacles behave as an isothermal object.

The mean time-averaged variation of Nusselt in terms of Re for different conductivity ratios of solids and fluids is shown in Fig. 9. As Re increases, the magnitude of heat transfer from the obstacle surface increases. Also, increasing thermal conductivity ratio between solid and fluid causes reduction of the heat flux resistance in the solid obstacle.

#### 4. CONCLUSIONS

In this paper, fluid flow and heat transfer in a two-dimensional channel with two obstacles located on its lower wall were investigated. A two-dimensional Lattice Boltzmann model was used to study the conjugate heat transfer between solid and fluid and the effect of many parameters on heat transfer was investigated. Compared to the commercial CFD methods, the use of the Lattice Boltzmann method has many advantages. One of the benefits of this is the simple application of the Boltzmann's method for complex geometries and the simple calculation process. In order to demonstrate the flexibility of this method, many parameters such as Reynolds number, thermal conductivity ratio between solid and fluid, and also distance between obstacles were investigated. The obstacle was assumed to heat up with constant heat flux, which actually simulates the heat transfer in electronic equipment. The results showed that with the increase of Reynolds number, the total heat transfer rate increases, and also the highest heat transfer occurs at the corners of the obstacles. The solid conductivity coefficient plays an important role in heat transfer. As conductivity ratio between solid and fluid increases, the resistance to heat transfer decreases and this means more heat transfer for extended surfaces.

#### REFERENCES

- Nejadseifi, M., Kiani-Oshtorjani, M., Tynjälä, T., and Jalali, P. (2024). Permeability of partially to fully randomized arrays of square disks: A lattice Boltzmann approach. *Journal of Porous Media*, 27(3), 85-99.
- Sara, O. N., Pekdemir, T., Yapici, S. I., and Yilmaz, M. (2001). Heat-transfer enhancement in a channel flow with perforated rectangular blocks. *International Journal of Heat and Fluid Flow*, 22(5), 509–518.
- Suss, A., Mary, I., Le Garrec, T., and Marié, S. (2023). Comprehensive comparison between the lattice Boltzmann and Navier–Stokes methods for aerodynamic and aeroacoustic applications. *Computers and Fluids*, 257, 105881.
- Ataei-Dadavi, I., Rounaghi, N., Chakkingal, M., Kenjeres, S., Kleijn, C. R., and Tummers, M. J. (2019). An experimental study of flow and heat transfer in a differentially side heated cavity filled with coarse porous media. *International Journal of Heat and Mass Transfer*, 143, 118591.
- Bhowmik, H., and Lee, K.-S. (2009). Analysis of heat transfer and pressure drop characteristics in an offset strip fin heat exchanger. *International Communications in Heat and Mass Transfer*, 36(3), 259–263.
- Chen, Z., Li, Q., Meier, D., and Warnecke, H.-J. (1997). Convective heat transfer and pressure loss in rectangular ducts with drop-shaped pin fins. *Heat and Mass Transfer*, 33, 219–224.
- Chikh, S., Boumedien, A., Bouhadef, K., and Lauriat, G. (1998). Analysis of fluid flow and heat transfer in a channel with intermittent heated porous blocks. *Heat and Mass Transfer*, 33(5), 405–413.
- Leung, C. W., Chan, T. L., Probert, S. D., and Kang, H. J. (1999). Forced convection from a horizontal ribbed rectangular base-plate penetrated by arrays of holes. *Applied Energy*, 62(2), 81–95.
- Matsuda, V. A., Martins, I. T., Moreira, D. C., Cabezas-Gómez, L., and Bandarra Filho, E. P. (2024). A Modified Enthalpic Lattice Boltzmann Method for Simulating Conjugate Heat Transfer Problems in Non-homogeneous Media. *Inventions*, 9(3), 57.
- Meinders, E. R., and Hanjalić, K. (2002). Experimental study of the convective heat transfer from in-line and staggered configurations of two wall-mounted cubes. *International Journal of Heat and Mass Transfer*, 45(3), 465–482.
- Mirahsani, S., Ahmadpour, A., and Hajmohammadi, M. R. (2023). Optimal design of an array of porous obstacles in a partially heated channel using lattice Boltzmann method for the heat transfer enhancement. *International Communications in Heat and Mass Transfer*, 143, 106737.
- Nazari, M., Kayhani, M. H., and Bagheri, A. A. H. (2013). Comparison of heat transfer in a cavity between vertical and horizontal porous layers using LBM. *Modares Mechanical Engineering*, 13(8), 93–107.
- Paknahad, R., Siavashi, M., and Hosseini, M. (2023). Pore-scale fluid flow and conjugate heat transfer study in high porosity Voronoi metal foams using multi-relaxation-time regularized lattice Boltzmann (MRT-RLB) method. *International Communications in Heat and Mass Transfer*, 141, 106607.
- Ramesh, K. N., Sharma, T. K., and Rao, G. A. P. (2021). Latest advancements in heat transfer enhancement in the micro-channel heat sinks: a review. *Archives of Computational Methods in Engineering*, 28, 3135–3165.
- Anderson, A. M., and Moffat, R. J. (1992). The adiabatic heat transfer coefficient and the superposition kernel function: Part 1—Data for arrays of flatpucks for different flow conditions.

- Mohammad, A. A. (2007). Applied lattice Boltzmann method for Transport phenomena, momentum, heat and mass transfer.
- Korichi, A., and Oufar, L. (2005). Numerical heat transfer in a rectangular channel with mounted obstacles on upper and lower walls. *International Journal of Thermal Sciences*, 44(7), 644–655.
- Pirouz, M. M., Farhadi, M., Sedighi, K., Nemati, H., and Fattahi, E. (2011). Lattice Boltzmann simulation of conjugate heat transfer in a rectangular channel with wall-mounted obstacles. *Scientia Iranica*, 18(2), 213–221.
- Zou, Q., and He, X. (1997). On pressure and velocity boundary conditions for the lattice Boltzmann BGK model. *Physics of Fluids*, 9(6), 1591–1598.
- Noble, D. R., Georgiadis, J. G., and Buckius, R. O. (1996). Comparison of accuracy and performance for lattice Boltzmann and finite difference simulations of steady viscous flow. *International Journal for Numerical Methods in Fluids*, 23(1), 1–18.



## Experimental and computational studies to investigate flow dynamics of Geldart A and Geldart B particles in a Circulating Fluidized Bed, CFB

Subham Kandel\*, Niroj Koirala\*, Nirajan Raut\*\*, Rajan Jaiswal\*\*, Sunil Prasad Lohani\*, Britt M.E. Moldestad\*\*

\*Department of Mechanical Engineering, Kathmandu University, Nepal

(\*nepal.subham, ern.koirala@gmail.com, \*splohani@ku.edu.np

\*\*Department of Process, Energy and Environmental Technology, University College of Southeast Norway, Norway

\*\*nirajannraut@gmail.com, (\*\*rajan.jaiswal, \*\*britt.moldestad@usn.no

**Abstract:** Circulating fluidized beds is one of the emerging technologies to convert waste to energy and an attractive method on a large scale. Key components such as the loop seal, gas distributor and cyclone separator play pivotal roles in facilitating solid recirculation and heat transfer within the system. This study focuses on the design and optimization of a CFB reactor using data derived from Barracuda Virtual Reactor software (CPFD). Initially, data from a small scale CFB reactor with main dimensions of 84 mm diameter and a loop seal diameter of 34 mm was utilized for simulation validation. By comparing simulation results with experimental data, the accuracy and reliability of the computational model were ensured. Subsequently, different reactor models were constructed and analyzed to explore various configurations and operating conditions. The results obtained from simulation based design and optimization provided valuable insights into achieving the optimal performance of the CFB system. By refining geometry, efficiency was increased by 32%. Overall, this study contributes to advancing the understanding, application and design modification of CFB technology in waste to energy conversion and large-scale industrial processes.

**Keywords:** Circulating Fluidized Bed, Fluidization, Minimum fluidization velocity, CFB, simulation, CPFD, Grid Size

### 1. INTRODUCTION

In the contemporary energy landscape, the escalating demand for sustainable energy solutions has accelerated efforts to harness energy from renewable sources. Among various technologies, fluidized beds stand out for their efficiency in biomass gasification and combustion, attributed to superior mixing, enhanced heat transfer and uniform temperature distribution (Moradi *et al.*, 2020).

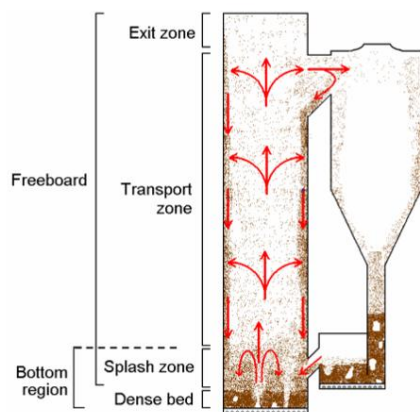


Fig. 1. Schematics of CFB (Pallarès, 2008).

Fluidization is a process wherein solid particles in a loosely packed bed exhibit fluid-like behavior when gas is blown upwards through them. In gas-solid systems, gas is introduced at the bottom of a column containing particles, causing them to vibrate and spread out to balance the drag force from the gas. As the gas velocity increases, this drag force equals the weight of the particles, resulting in a fluidized bed. Various flow patterns can emerge depending on the gas velocity, as illustrated in Fig. 2.

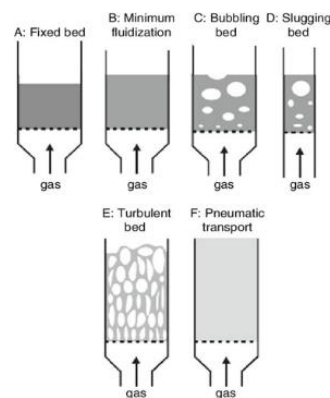


Fig. 2. Schematic of fluidized bed in different regimes (Soomro *et al.*, 2012).

These regimes include fixed, bubbling, slugging, turbulent and pneumatic conveying. The transition from a fixed bed to a fluidized bed occurs at the minimum fluidization velocity, the gas velocity needed to suspend the solids in the gas stream. At very high gas flow rates, a substantial portion of solid is ejected into the space above the bed, leading to significant particle loss through elutriation and entrainment (Huang, 2006).

To mitigate this issue, gas-solid separators are employed to capture and return most particles to the bed (Soomro *et al.*, 2012). Maintaining stable and continuous solid circulation in a gas-solid system has led to the development of Circulating Fluidized Beds (CFB's) as shown in Fig. 1, for large-scale processes, particularly in the petroleum and power generation industries (Huang, 2006). CFB's function similarly to bubbling beds but with substantially higher fluid flow velocities, resulting in more intense mixing and enhanced gas-solid contact. The high relative velocity between gas and solid particles leads to exceptionally high rates of heat and mass transfer. However, the increased gas velocities and recirculation of solids make CFB systems more costly in terms of power consumption and investment compared to conventional fluidized bed reactors (Moradi *et al.*, 2020).

The efficiency of CFBs is highly dependent on flow behavior making it crucial to understand this behavior for scaling, designing and optimization. Over the past decades, Computational Fluid Dynamics (CFD) has emerged as a valuable tool for predicting flow behavior in fluidized bed processes. However, further model development and validation are necessary (Pallarès, 2008).

The experiment utilized sand particles as the bed material and air as the fluidizing fluid. Sand is commonly used in fluidized bed reactors to enhance the mixing of fuel with the fluidizing gas, improving mass and energy transfer. In a biomass CFB reactor, biomass rapidly reacts with the gas to produce synthesis gas and char. The unreacted char circulates with the sand particles through the system and back to the reactor, with sand primarily controlling the circulation behavior rather than the char (Niven, 2002). Various design modifications were implemented on the cyclone, adjustments to the recirculating pipe inlet angle and the height of the recirculating pipe. Among the different design variants, the one exhibiting the maximum particle recirculation rate was selected.

This work underscores the importance of precise modeling and experimentation in optimizing fluidized bed reactors for biomass gasification and combustion, paving the way for more efficient and sustainable energy production.

## 2. MP PIC MODEL DESCRIPTION

The gas phase mass and momentum conservation can be modelled by the volume averaged Navier-Stokes equation

and are used as a continuum on a Eulerian grid (Chutima Dechsiri, 2004).

$$\frac{\partial(\theta_f \rho_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f u_f) = 0 \quad (1)$$

$$\begin{aligned} \frac{\partial(\theta_f \rho_f u_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f u_f) \\ = \nabla p - F + \theta_f \rho_f g \\ + \nabla \cdot (\theta_f \tau_f) \end{aligned} \quad (2)$$

$$F = \iint f m_s \left[ D_s (u_f - u_s) - \frac{1}{\rho_s} \nabla p \right] dm_s du_s \quad (3)$$

where  $\theta_f$ ,  $\rho_f$ ,  $u_f$ ,  $\tau_f$  are fluid phase volume fraction, density, velocity and stress tensor and  $m_s$ ,  $u_s$  are the mass and velocity of the particle.  $F$  is the total momentum exchange with particle phase per volume,  $g$  is the acceleration due to gravity and  $p$  is the pressure.

The solid phase can be modelled by a particle distribution function given by equation 4 (Snider, (2001)). Considering the time rate of change of above equation the Liouville equation is obtained. This equation assumes that there are no direct collisions or particle breakup.

$$f(x, m_s, u_s, t) dm_s du_s$$

$$\frac{\partial f}{\partial t} + \nabla_x \cdot (f u_s) + \nabla_{u_s} \cdot (f A) = 0 \quad (4)$$

The particle acceleration,  $A$  as a function of aerodynamics drag, buoyancy, gravity and interparticle normal stresses can be expressed as,

$$A = D_s (u_f - u_s) - \frac{1}{\rho_s} \nabla p + g - \frac{1}{\theta_s \rho_s} \nabla \tau_s \quad (5)$$

The particle volume fraction,  $\theta_s$  and the particle stress  $\tau_s$ , which are used to calculate the interparticle collisions and are expressed as

$$\theta_s = \iint f \frac{m_s}{\rho_s} dm_s du_s \quad (6)$$

$$\tau_s = \frac{10 P_s \theta_s^\beta}{\max[(\theta_{cp} - \theta_s), \epsilon(1 - \theta_s)]} \quad (7)$$

Here,  $P_s$ ,  $\beta$ ,  $\theta_{cp}$  are the constant term related with pressure, is a constant, particle volume fraction equals the close pack volume (Andrews *et al.*, 1996).

### 3. METHOD AND COMPUTATIONAL MODEL

#### 3.1 Experimental Model

The experiment was conducted at ambient temperature using sand particles with diameter ranging from 63 to 200  $\mu\text{m}$  and density of 2650  $\text{kg/m}^3$ . Particle size distribution analysis was performed prior to the experiment revealed a mean particle size diameter of 116  $\mu\text{m}$ . The gas flow rate was varied from 0 to 650 SLPM in increments of 50 SLPM. Pressure transducers were installed at various locations, and data acquisition was facilitated using LabVIEW software.

#### 3.2 Computational Model:

Following the measurement of the dimensions of the CFB at the University of South-Eastern Norway, a CAD geometry was created using SOLIDWORKS 2020. Various simulation models with different grid sizes were analyzed before finalizing the computational model. Experimental pressure transducer data were primarily used to validate the model (Figs. 4 and 5). The gas inlet of the riser and the loop seal were configured as flow boundary conditions, while the top of the cyclone was set up as a pressure boundary condition. The simulation time step was set to 0.0005 seconds (Bandara *et al.*, 2018), with a total simulation duration of 45 seconds. The maximum momentum from particle collision redirection was assumed to be 40%, and the default values of 0.85 were used for normal and tangential wall collisions.

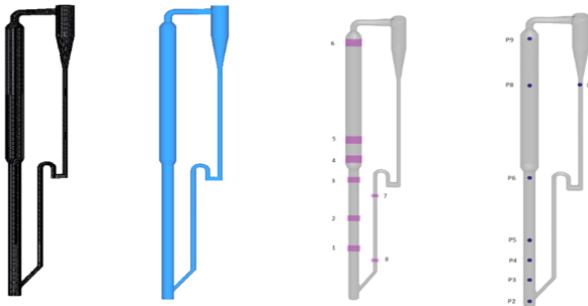


Fig. 3. (a) Grid (b) CAD Geometry (c) Flux Plane (d) Pressure reading Points.

### 4. RESULT AND DISCUSSION.

#### 4.1 Model Validation:

After conducting multiple simulations with varying grid sizes and drag models, the Wen-Yu Ergun drag model was selected. As illustrated in Figs. 4 and 5 (*experiment with trim means first 180 sec data has been excluded on an experiment of 3600 sec*), the minimum deviation was observed for a grid size of 80,000 cells. Increasing the grid size beyond this point resulted in deviations, as the grid size became smaller than the particle size. The relationship between cells to computational particles is shown in Table 1.

Table 1. Computational Particle to Cell Ratio

Grid	Cells	Computational Particles	Computational Particle to Cell Ratio
01	40000	35640	0.891
02	60000	56400	0.940
03	80000	77792	0.972
04	80000	77800	0.973
05	100000	96768	0.968
06	120000	111384	0.928

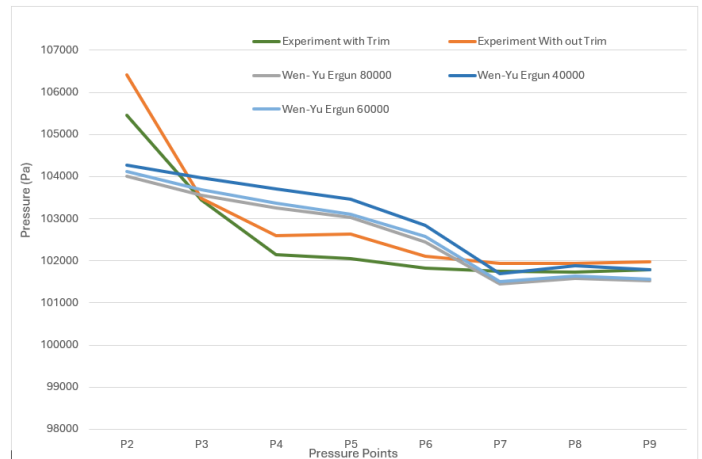


Fig. 4. Pressure Variation with Pressure point (Different grid size).

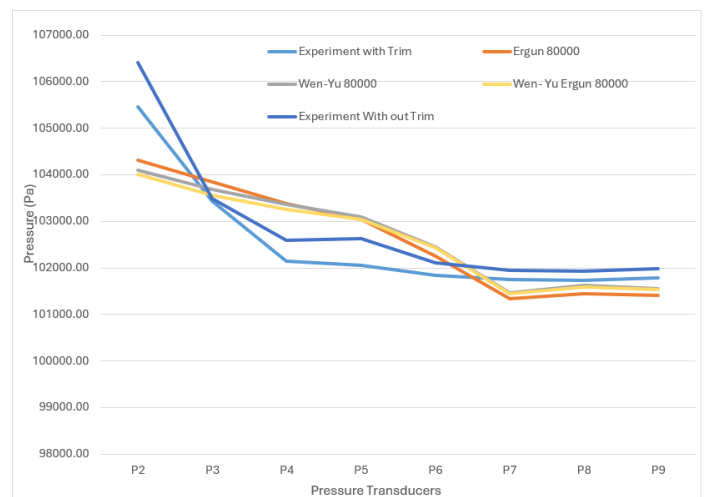


Fig. 5. Pressure Variation with Pressure point (Different drag model).

4.2 Impact of changing the diameter of Cyclone Keeping height constant

The diameter of the cyclone was varied while keeping the height constant to investigate its impact on particle circulation rate. Adjusting the diameter of a cyclone separator in a circulating fluidized bed (CFB) yielded various outcomes in particle circulation rate. In the design, with a height to diameter (H/D) ratio of 2.78, it was observed a 1.6% increase in particle circulation rate as shown in Figs. 7 and 8 and Table 2. The CAD models varying diameter is shown in Fig. 6.

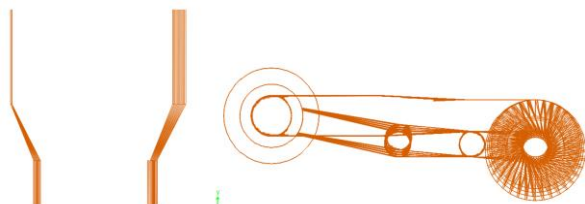


Fig. 6. Variation in diameter keeping height constant.

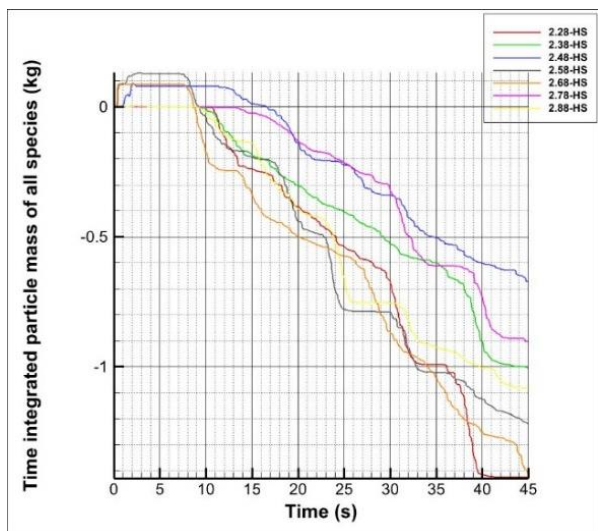


Fig. 7. Time Integrated particle mass of all Species (Seventh Plane).

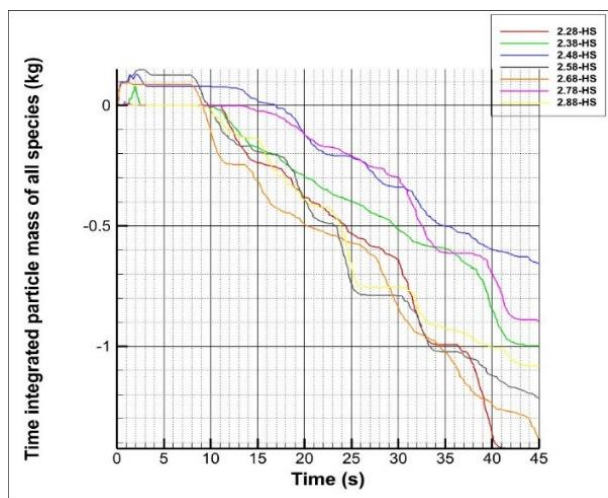


Fig. 8. Time Integrated particle mass of all Species (Eighth Plane)

Table 2. Change in circulation rate keeping height constant (305 mm)

H/D	Diameter (mm)	Circulation Rate (kg)	Change in Circulation Rate (%)
2.28	133.77	1.02	-16.4
2.38	128.15	1.17	-4.1
2.48	122.98	0.80	-34.4
2.58	118.21	1.22	0.0
2.68	113.80	0.79	-35.2
2.78	109.71	1.24	1.60
2.88	105.90	0.75	-38.5

4.3 Impact of changing the height of cyclone keeping diameter constant:

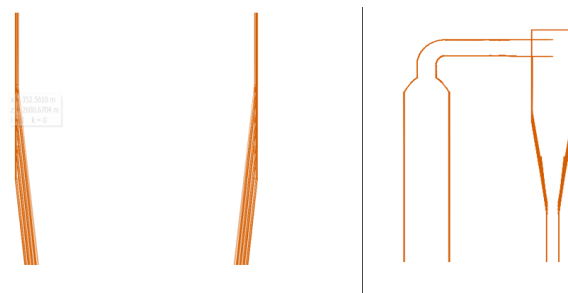


Fig. 9. Variation in height keeping diameter constant.

Changing the height of a cyclone can have a direct impact on both gas and particle residence times. The longer the residence time, the more opportunity particles have to be separated from the gas stream by centrifugal force. A taller cyclone can enhance both the separation efficiency and the rate of circulation, as particles have more time to move towards the cyclone wall under the influence of centrifugal force.

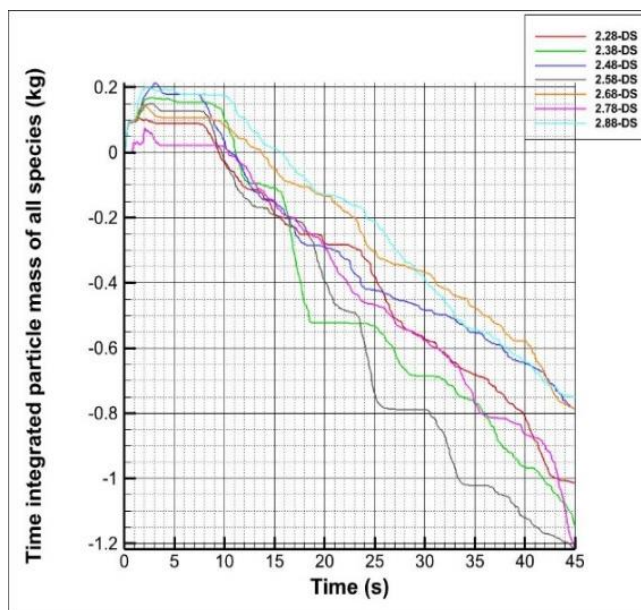


Fig. 10. Time Integrated particle mass of all species (Seventh Plane).



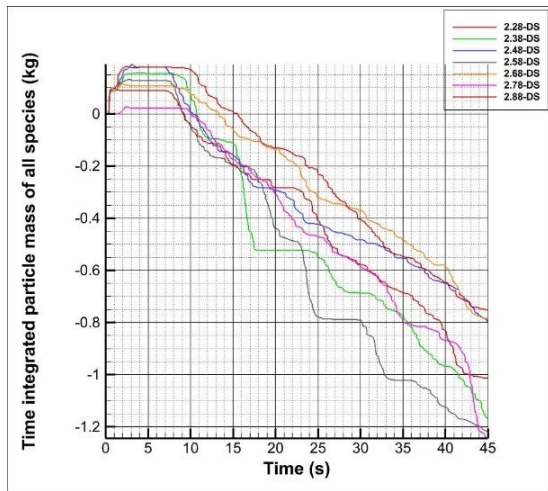


Fig. 11. Time Integrated particle mass of all species (Eighth Plane).

Table 3. Change in circulation rate keeping diameter constant (118 mm)

H/D	Height (mm)	Circulation Rate (kg)	Change in Circulation Rate (%)
2.28	269.04	1.40	14.75
2.38	280.84	1.00	-18.03
2.48	292.64	0.65	-46.72
2.58	304.44	1.22	0.00
2.68	316.24	1.42	16.39
2.78	328.04	0.89	-27.05
2.88	339.84	1.08	-11.48

In the design modification, with an H/D ratio of 2.68, it was observed a 16.39% increase in the particle circulation rate. Figs. 10 and 11 shows particle circulation rate for 45 sec on flux plane 7 and 8. *Several factors influence the particle circulation rate, and a design optimized for a particular particle size may not be effective for a wide range of particle sizes.*

4.4 Impact of changing the angle of recirculation pipe (Return Leg /Downcomer)

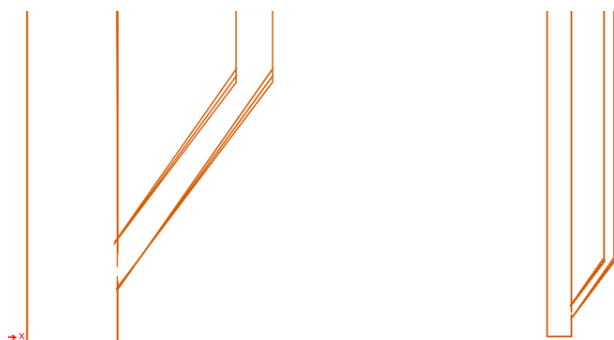


Fig. 12. Variation in inlet angle of return leg.

The angle of the recirculation pipe significantly affects particle movement to the riser in a Circulating Fluidized Bed (CFB). Typically, increasing the angle of the downcomer enhances gravitational forces, aiding particle movement. This improvement results in a higher particle recirculation rate, as particles flow more smoothly and quickly to the riser.

Additionally, the risk of blockage decreases with steeper angles, as particles are more likely to settle and move smoothly within the downcomer.

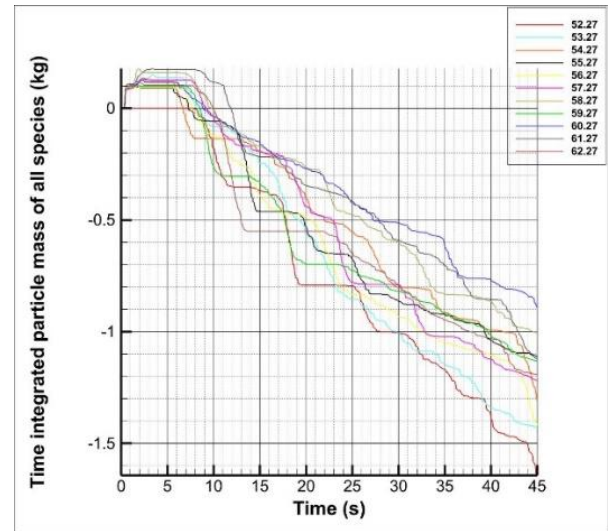


Fig. 13. Time Integrated particle mass of all species (Seventh Plane).

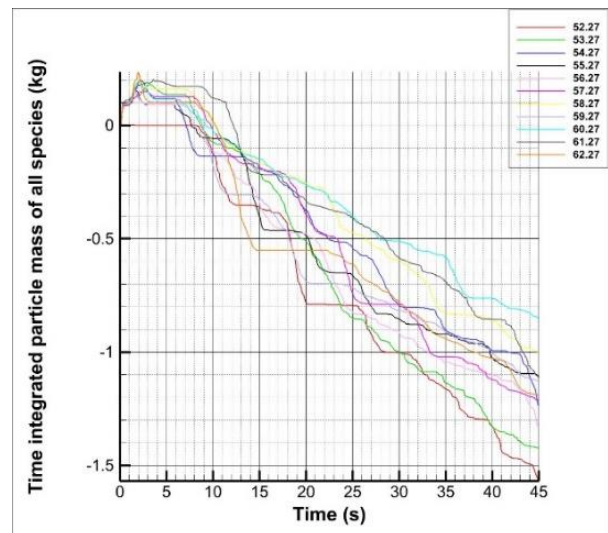


Fig. 14. Time Integrated particle mass of all species (Eighth Plane).

To assess the impact on particle recirculation rate, 11 design variations were implemented, altering the angle from 52.27° to 62.27°. The angle of the return pipe used in the experiment was 57.27°. In the observation, the maximum circulation rate occurred at an angle of 52.27°. Although this result appears unusual, it can be attributed to the backflow of particles from the riser to the loop seal, causing blockages and pushing particles backward towards the standpipe rather than the downcomer. As long as there is backpressure from particles inside the riser, the circulation rate will decrease. Therefore, the optimal angle was found to be 52.27 degrees, despite it being a shallower return angle. The circulation rate varying return leg angle is shown in Figs. 13 and 14.



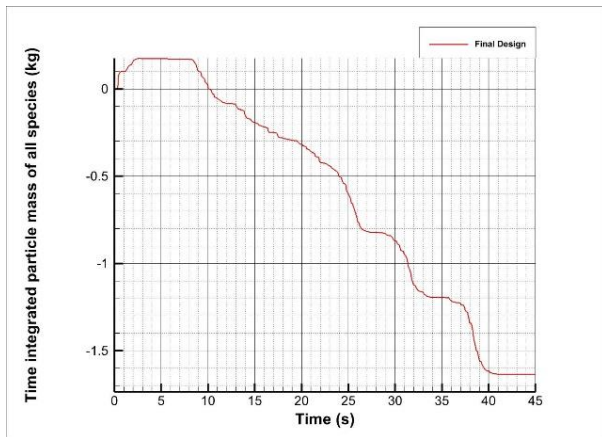


Fig. 15. Final Design.

In the final design, each individual component was optimized to maximize circulation rate, resulting in a 32% increase in particle circulation rate as shown in Fig. 15.

## 5. CONCLUSIONS

This study examines the influence of design parameters on the particle circulation rate in a circulating fluidized bed with a mixture of Geldart A and B particles. Key design modifications included varying the cyclone diameter while keeping the height constant, varying the height while keeping the diameter constant and adjusting the angle of the recirculating pipe. An initial H/D ratio of 2.58 was optimized to 2.78 (height same, diameter 109.7 mm) and 2.68 (diameter same, height 316.2 mm) for improved particle circulation. The optimal recirculation angle was found to be 52.27 degrees. The results are influenced by multiple factors, including particle size distribution, particle to particle and particle to wall collisions, particle breakdown after collisions and velocity profile irregularities. Consequently, a design suitable for one particle size may not be valid for a wide range of particles.

## REFERENCES

- Andrews, M. J. and P. J. O'Rourke (1996). "The multiphase particle-in-cell (MP-PIC) method for dense particulate flows." *International Journal of Multiphase Flow* 22(2): 379-402.
- Bandara, J. C., Nielsen, H. K., Moldestad, B. M. E., and Eikeland, M. S. (n.d.). Sensitivity Analysis and Effect of Simulation parameters of CPFD Simulation in Fluidized Beds (2018).
- Chutima Dechsiri (2004) Particle Transport in Fluidized Beds: Experiments and Stochastic Models. <https://pure.rug.nl/ws/portalfiles/portal/9807341/thesis.pdf>
- Huang, Y. (2006). Dynamic model of circulating fluidized bed. <https://researchrepository.wvu.edu/etd/2749>
- Moradi, A., Samani, N. A., Mojarrad, M., Sharfuddin, M., Bandara, J. C., and Moldestad, B. M. E. (2020). Experimental and computational studies of circulating fluidized bed.
- Pallarès, David., and Chalmers tekniska högskola. Department of Energy and Environment. Division of Energy Technology. (2008). Fluidized bed combustion: modeling and mixing. Chalmers University of Technology.
- R. K. Niven, Physical insight into the Ergun and Wen and Yu equations for fluid flow in packed and fluidised beds, *Chemical Engineering Science*, vol. 57, no. 3, pp. 527–534, Feb. 2002.
- Soomro, A., Samo, S. R., and Hussain, A. (2012). Fluidization in cold flow circulating fluidized bed system. In *Energy, Environment and Sustainable Development*, pp. 161–173. Springer-Verlag Vienna.
- Snider, D. M. (2001). An Incompressible Three-Dimensional Multiphase Particle-in-Cell Model for Dense Particle Flows. *Journal of Computational Physics* 170(2), 523-549.

# Impact of grid sensitivity and drag model along with the height of recirculating pipe on a cold flow circulating fluidized bed

Subham Kandel\*, Nirajan Raut\*\*, Niroj Koirala\*, Sunil Prasad Lohani\*, Britt M.E. Moldestad\*\*

\*Department of Mechanical Engineering, Kathmandu University, Nepal

(nepal.subham, ern.koirala@gmail.com, splohani@ku.edu.np)

\*\*Department of Process, Energy and Environmental Technology, University College of Southeast Norway, Norway

\*\*nirajannraut@gmail.com, (\*rajan.jaiswal, \*\*britt.moldestad@usn.no)

**Abstract:** Fluidized bed technology known for its efficient heat and mass transfer and controlled material handling, is widely used across industries. However, CFD simulation of fluidized beds presents challenges that require extensive validation. This study leverages the Multiphase Particle-In-Cell (MP-PIC) method, a recent Lagrangian modeling technique to improve computational efficiency and accuracy. The CAD model was developed using SolidWorks 2020 and simulation was carried out in the commercial CFD package Barracuda VR 21.1.0. The sensitivity of grid size, drag models and the impact of recirculating pipe height after loop seal was examined. Sand particles 63-200  $\mu\text{m}$  and air were used as bed material and fluidization gas respectively achieving full flow circulation at 650 SL/min and 12 SL/min aeration in the riser and loop seal. A total of 19 different simulations were conducted, varying grid size and drag models each for a duration of 45 seconds with a time step of 0.0005 seconds. Pressure transducers along the CFB walls provided validation data. The Wen-Yu Ergun drag model showed a minimal error margin of 0.60%, followed by the Wen-Yu 80000 model at 0.62%, demonstrating high predictive accuracy.

**Keywords:** Circulating Fluidized bed (CFB), Minimum fluidization velocity, CFD simulations, Time step, drag model, Multiphase particle-in-cell method, Grid size

## 1. INTRODUCTION

Gas-solid fluidized bed technology is widely utilized in energy generation, pharmaceutical, chemical, petrochemical, electronic and metallurgical processing industries due to its distinct advantages of high heat and mass transfer and controlled material handling (Moradi *et al.*, 2020). Computational fluid dynamic (CFD) modeling has been identified as an excellent tool to produce information during the scaling up of pilot scale circulating fluidized beds to industrial scale. Further, it is a fast and cost-effective method for system optimization (Jaiswal *et al.*, 2022). CFD solves the conservation equations for mass, momentum, energy and species and this technique has been critically validated for accurate performance in gas or liquid single-phase flows. However, challenges remain related to interface coupling, solid-phase modeling and scale differences in gas-solid multiphase flow systems (Bandara *et al.*, 2016).

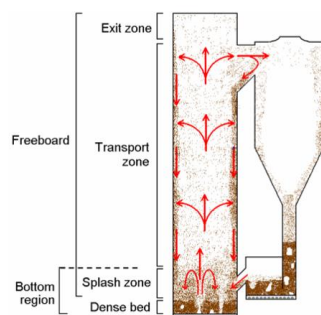


Fig. 1. Schematics of CFB (Pallarès, 2008).

Eulerian-Eulerian and Eulerian-Lagrangian are the two basic approaches for CFD modeling of multiphase flows. Multiphase Particle-In-Cell (MP-PIC) modeling is a development of Eulerian-Lagrangian modeling and aims to reduce the computational cost in discrete modeling of the particle phase (Andrews and O'Rourke, 1996). Instead of tracking individual particles, it considers packets containing a certain number of particles with similar properties. The packets are modeled in the discrete phase while the particle phase interactions are modeled in an Eulerian frame (Snider *et al.*, 2001). Therefore, particle properties are calculated in both Eulerian and Lagrangian frames which are correlated via interpolation functions.

Validated CFD models can be used to analyze circulating fluidized beds in terms of particle circulation velocity, particle mixing and segregation (Bandara *et al.*, 2018). The conservation equations of mass species, momentum and energy are in partial differential form so the simulation geometry is divided into small cells referred to as the computational grid. The conservation equations are then discretized in space and time to form a set of algebraic equations (Bandara *et al.*, 2018). Finite difference, finite element and finite volume are the main techniques used with the finite volume method being most common for 3D systems involving mass, momentum and energy (Andrews and O'Rourke, 1996).

Errors and uncertainties are integrated from the modeling stage to the final computer simulations. The use of empirical equations and model simplification leads to deviations during

model development. Therefore, it is necessary to identify ways to reduce errors in simulations with minimal computational cost. This includes selecting the optimal drag model in gas-solid multiphase flow systems and conducting mesh sensitivity analysis to develop a grid independent model (Bandara *et al.*, 2018).

This paper investigates the impact of grid sensitivity and drag models alongside the height of the recirculating pipe after loop seal on the performance of cold flow circulating fluidized beds. Barracuda VR 21.1.0 was used to compare pressure data, varying grid size and drag model along with SolidWorks for CAD design. Through comprehensive CFD analysis, this study aims to enhance the accuracy and efficiency of circulating fluidized bed modeling for industrial applications.

## 2. MP PIC MODEL DESCRIPTION

The gas phase mass and momentum conservation can be modelled by the volume averaged Navier-Stokes equation and are used as a continuum on a Eulerian grid (Snider, 2001).

$$\frac{\partial(\theta_f \rho_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f u_f) = 0 \tag{1}$$

$$\begin{aligned} \frac{\partial(\theta_f \rho_f u_f)}{\partial t} + \nabla \cdot (\theta_f \rho_f u_f) \\ = \nabla p - F + \theta_f \rho_f g \\ + \nabla \cdot (\theta_f \tau_f) \end{aligned} \tag{2}$$

$$\begin{aligned} F = \iint f m_s \left[ D_s (u_f - u_s) \right. \\ \left. - \frac{1}{\rho_s} \nabla p \right] dm_s du_s \end{aligned} \tag{3}$$

where  $\theta_f$ ,  $\rho_f$ ,  $u_f$ ,  $\tau_f$  are fluid phase volume fraction, density, velocity and stress tensor and  $m_s$ ,  $u_s$  are the mass and velocity of the particle. F is the total momentum exchange with particle phase per volume, g is the acceleration due to gravity and p is the pressure.

The solid phase can be modelled by a particle distribution function given by equation 4 (O'Rourke *et al.*, 2014). Considering the time rate of change of above equation the Liouville equation is obtained. This equation assumes that there are no direct collisions or particle breakup.

$$\begin{aligned} f(x, m_s, u_s, t) dm_s du_s \\ \frac{\partial f}{\partial t} + \nabla_x \cdot (f u_s) + \nabla_{u_s} \cdot (f A) = 0 \end{aligned} \tag{4}$$

The particle acceleration, A as a function of aerodynamics drag, buoyancy, gravity and interparticle normal stresses can be expressed as,

$$A = D_s(u_f - u_s) - \frac{1}{\rho_s} \nabla p + g - \frac{1}{\theta_s \rho_s} \nabla \tau_s \tag{5}$$

The particle volume fraction,  $\theta_s$  and the particle stress  $\tau_s$ , which are used to calculate the interparticle collisions and are expressed as (Rourke *et al.*, 2014)

$$\theta_s = \iint f \frac{m_s}{\rho_s} dm_s du_s \tag{6}$$

$$\tau_s = \frac{10 P_s \theta_s^\beta}{\max[(\theta_{cp} - \theta_s), \epsilon(1 - \theta_s)]} \tag{7}$$

Here,  $P_s$ ,  $\beta$ ,  $\theta_{cp}$  are the constant term related with pressure and is a constant, particle volume fraction equals the close pack volume.

## 3. METHOD AND COMPUTATIONAL MODEL

### 3.1 Experimental Model

The experiment was carried out at ambient temperature using sand particles with diameters ranging from 63 to 200  $\mu\text{m}$  and a density of 2650  $\text{kg/m}^3$ . Prior to the experiment, particle size distribution analysis revealed a mean particle size of 116  $\mu\text{m}$ . The gas flow rate in the riser was varied from 0 to 650 SLPM in increments of 50 SLPM and in the loop seal from 0 to 12 SLPM in increments of 2 SLPM. Pressure transducers were installed at various locations, with data acquisition managed through LabVIEW.

### 3.2 Computational Model:

Following the measurement of the dimensions of the CFB at the University of South-Eastern Norway, a CAD geometry was created using SOLIDWORKS 2020. The gas inlet of the riser and the loop seal were configured as flow boundary conditions, while the top of the cyclone was set up as a pressure boundary condition. The simulation time step was set to 0.0005 seconds (Bandara *et al.*, 2018), with a total simulation duration of 45 seconds. The maximum momentum from particle collision redirection was assumed to be 40%, and the default values of 0.85 were used for normal and tangential wall collisions.

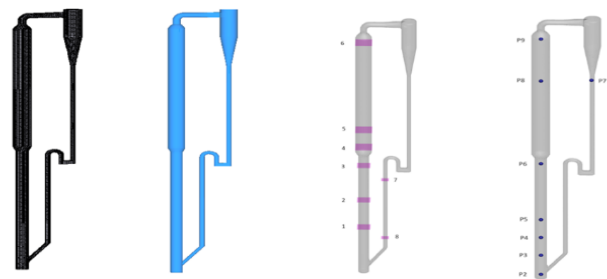


Fig. 2. (a) Grid (b) CAD Geometry (c) Flux Planes (d) Pressure reading Points.

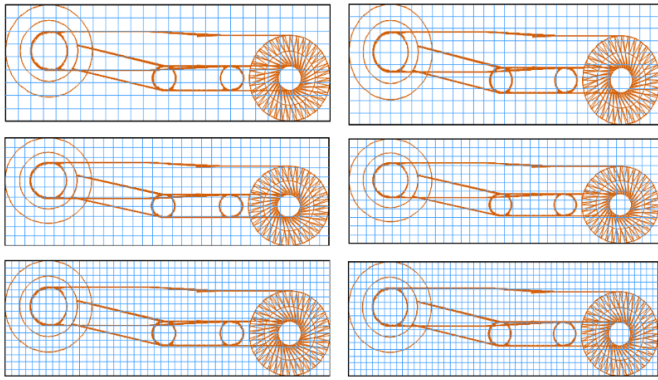


Fig. 3. Grid size with 40000, 60000, 80000, 120000, 240000, 300000 uniform grid size (Top left to bottom right).

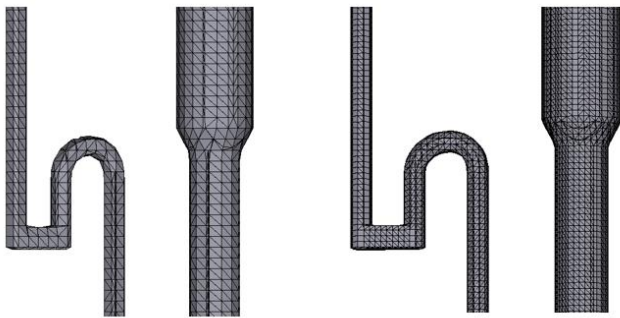


Fig. 4. CAD model with 40000 and 30000 uniform grid size.

#### 4. RESULT AND DISCUSSION

##### 4.1 Drag Model

Three different drag models Ergun, Wen-Yu and a combined Wen-Yu Ergun model were evaluated to accurately compare the computational results with the experimental data.

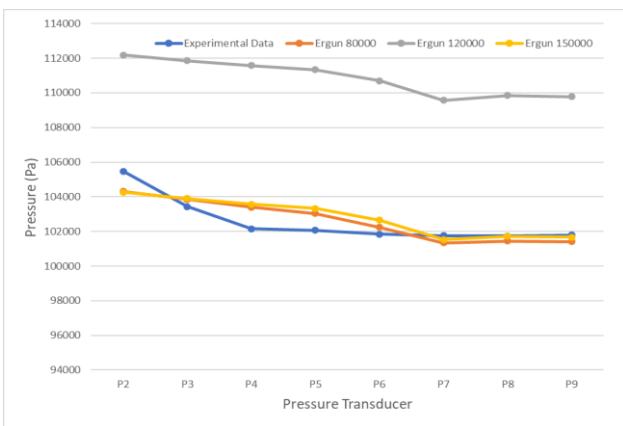


Fig. 5. Ergun drag model with different grid size

In the above study, the Ergun model was analyzed using different grid sizes. When increasing the grid size from 80,000 to 150,000 cells it was observed that a grid size of 120,000 cells resulted in greater deviation, whereas grid sizes of 80,000 and 150,000 cells closely matched the experimental data. The average deviations for grid sizes of 80,000, 120,000 and 150,000 cells were 0.69%, 8.12% and 0.67% respectively as

shown in Fig. 5. These results clearly indicate that the Ergun model with grid sizes of 80,000 and 150,000 cells provides accurate predictions for our experimental model.

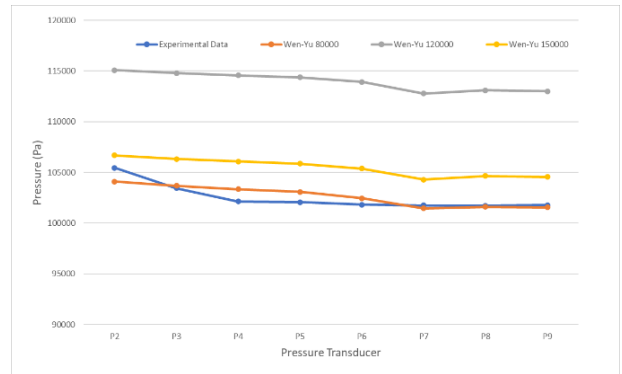


Fig. 6. Wen-Yu drag model with different grid size.

In the above analysis, the Wen-Yu model was evaluated using different grid sizes. As the grid size was increased from 80,000 to 150,000 cells, it was observed that the 120,000 cell exhibited the highest deviation followed by the 150,000 cell whereas the 80,000 cell provided accurate predictions compared to the experimental data. The average deviations for the 80,000, 120,000 and 150,000 grid cells were 0.62%, 11.15% and 2.88% respectively as shown in Fig. 6. These results clearly indicate that the Wen-Yu model with a grid size of 80,000 cells offers the best prediction accuracy for our experimental model.

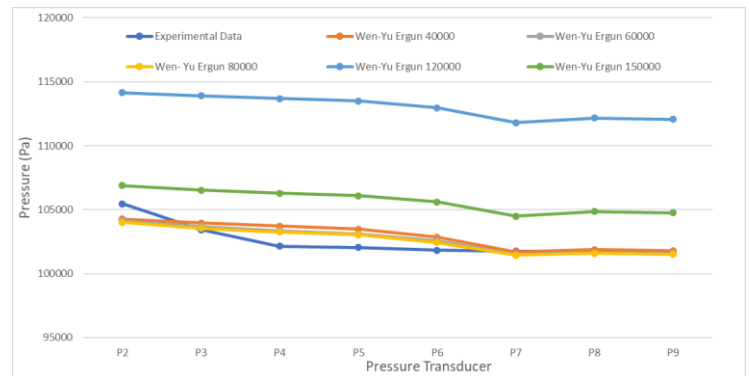


Fig. 7. Wen-Yu Ergun drag model with different grid size.

In the above study, the Wen-Yu Ergun model was evaluated using various grid sizes. Increasing the grid size from 40,000 to 150,000 cells revealed that grids of 120,000 and 150,000 cells exhibited greater deviation, while grids of 40,000, 60,000 and 80,000 cells closely matched the experimental data. The average deviations for grid sizes of 40,000, 60,000, 80,000, 120,000 and 150,000 cells were 0.71%, 0.62%, 0.60%, 10.25% and 3.09% respectively as shown in Fig. 7. These result clearly indicates that the Wen-Yu Ergun model with a grid size of 80,000 cells provides the most accurate predictions for our experimental model.

##### 4.2 Grid Size

Different grid sizes ranging from 40,000 to 300,000 were tested for grid independence test with three different drag models and the results are presented below.



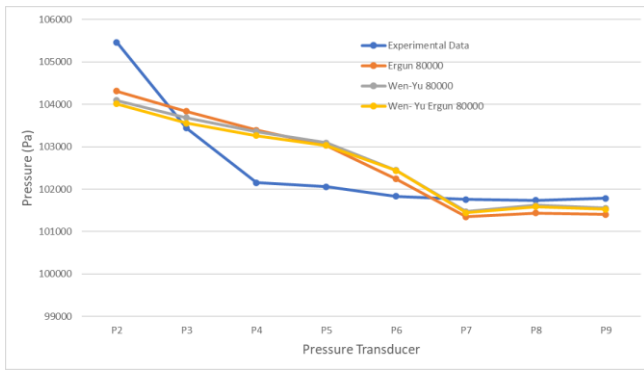


Fig. 8. Different drag model with 80000 grid size.

In this study, the Wen-Yu Ergun drag model with an 80,000 grid cells accurately predicted the experimental setup. The errors for the Ergun, Wen-Yu, and Wen-Yu Ergun models were 0.63%, 0.62% and 0.60%, respectively as shown in Fig. 8.

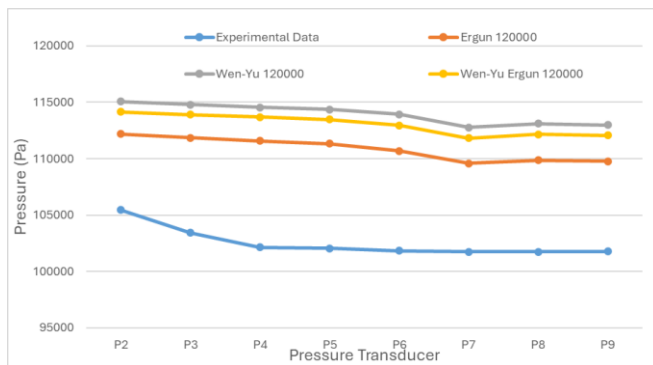


Fig. 9. Different drag model with 150000 grid size.

In the above study, increasing the grid size resulted in greater deviation. The deviation is smallest for the Ergun followed by the Wen-Yu Ergun and is highest for the Wen-Yu model. The errors for the Ergun, Wen-Yu and Wen-Yu Ergun model are 8.12%, 11.15% and 10.25% respectively as shown in Fig. 9.

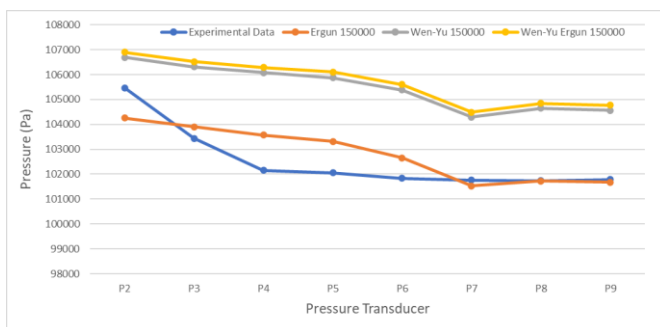


Fig. 10. Different drag model with 300000 grid size.

In the above study, increasing the grid size led to greater deviation. The Ergun model exhibited the least deviation followed by the Wen-Yu and Wen-Yu Ergun models. The deviation for these drag models with the specified grid sizes are 0.67%, 2.88% and 3.09% respectively as shown in Fig. 10.

### 4.3 Impact of height of recirculating pipe after loop seal on recirculation rate

Increasing the height of the recirculating pipe after the loop seal in a circulating fluidized bed (CFB) system significantly impacts particle recirculation rate and system performance. The height affects gravitational forces, particle velocity and settling behavior. A taller recirculating pipe increases pressure drop and alters particle flow dynamics potentially reducing velocities if not compensated for the increased gravitational and pressure effects.

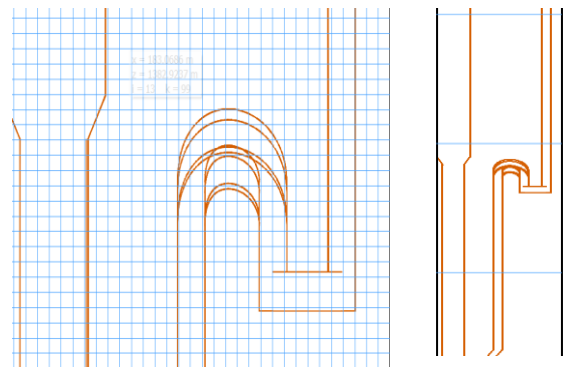


Fig. 11. CAD model with varying recirculating pipe height.

Length and Height of Recirculating Pipe		
Height (mm)	Circulation Rate (kg)	Change in Circulation Rate
(950-43)	1.31	7.38
(955-48)	1.16	-4.92
(960-53)	1.2	-1.64
(965-58)	1.26	3.28
(970-63)	1.22	0.00
(975-68)	1.15	-5.74
(980-73)	1.08	-11.48
(985-78)	1.1	-9.84
(990-83)	1.08	-11.48

In the design, nine modifications were made varying the recirculating height from 950 mm to 990 mm, with the original height being 970 mm. These changes also affect the height just after the loop seal. A greater height requires more pressure for sand transfer to the riser while a lower height facilitates particle flow from the riser to the loop seal.

In nine simulations, the configuration with a 965 mm height and 58 mm height just after the loop seal achieved the highest recirculation rate as shown in Figs. 12 and 13. Initially the 950 mm configuration had the highest rate up to 45 seconds, but the (965-58) mm configuration provided the best circulation rate over a 300 second simulation.



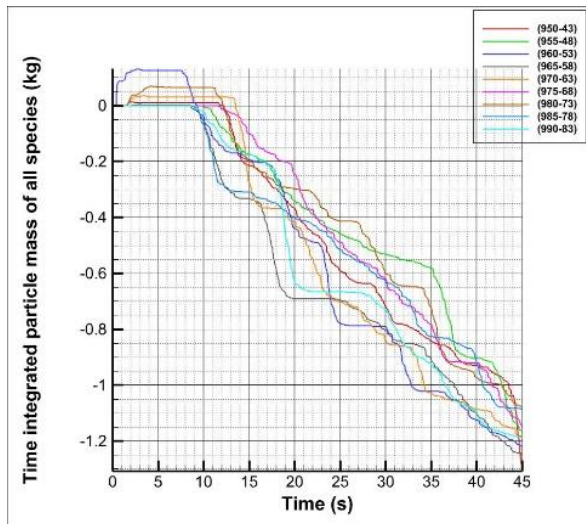


Fig. 12. Particle circulation rate with different height of recirculating pipe (45 seconds).

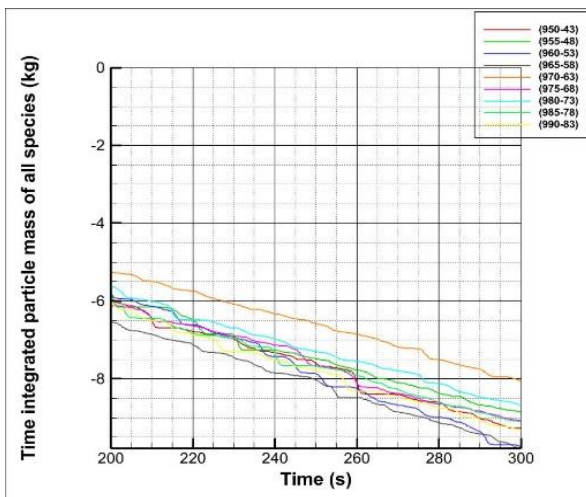


Fig 13. Particle circulation rate with different height of recirculating pipe (300 seconds).

## 5. CONCLUSION

This study investigates the impact of drag model and grid size on the accuracy of experimental and computational models in a cold flow circulating fluidized bed (CFB) system. Sand particles ranging from 63-200  $\mu\text{m}$  were used as the bed material with air as the fluidizing agent. The minimum airflow for particle circulation was determined experimentally and was found to be 650 SLPM in the riser and 12 SLPM in the loop seal.

Analysis reveals that the Wen-Yu Ergun model with a grid size of 80,000 accurately represents the experimental model computationally. Accuracy improves progressively from 40,000 to 80,000 grid size peaking at 80,000 before declining beyond this threshold. Conversely, when the same model is analyzed with a higher grid size of 300,000, the deviation increases to 11.17%. This deviation is attributed to particle size becoming smaller than the grid size beyond 80,000.

Furthermore, the circulation rate is observed to be highest for the (965-58) mm configuration compared to the (970-63) mm model used in the experiment. This is because greater height necessitates more pressure for sand transfer to the riser, while lesser height facilitates more particle flow from the riser to the loop seal. Insufficient height leads to suboptimal particle circulation due to pressure pushing particles from the loop seal towards the standpipe.

In conclusion, drag model and grid size significantly impacts computational accuracy, with no one particular model applicable across all CFB models. Additionally, particle characteristics play a crucial role in model validation with uniform particle size essential for accurate predictions. Moreover, the height of the recirculating pipe profoundly influences particle circulation rate, emphasizing the importance of designing an optimal height for efficient operation of circulating fluidized beds.

## REFERENCES

- Andrews, M. J., and O'Rourke, P. J. (1996). THE MULTIPHASE PARTICLE-IN-CELL (MP-PIC) METHOD FOR DENSE PARTICULATE FLOWS. Pergamon Int. J. Multiphase Flow, 22(2).
- Bandara, J. C., Nielsen, H. K., Moldestad, B. M. E., and Eikeland, M. S. (2018). Sensitivity Analysis and Effect of Simulation parameters of CDF Simulation in Fluidized Beds. Proceedings of The 59th Conference on Simulation and Modelling (SIMS 59), 26-28 September 2018, Oslo Metropolitan University, Norway, 153, 334-341.
- Bandara, J. C., Thapa, R. K., Moldestad, B. M. E., and Eikeland, M. S. (2016). Simulation of Particle Segregation in Fluidized Beds. Proceedings of The 57th EUROSIM Congress on Modelling and Simulation, EUROSIM 2016, The 57th SIMS Conference on Simulation and Modelling SIMS 2016, 142, 991-997.
- Huang, Y. (2024). Dynamic model of circulating fluidized bed. <https://researchrepository.wvu.edu/etd/2749>

- Jaiswal, R., Eikeland, M. S., Moldestad, B. M. E., and Thapa, R. K. (2022). Influence on the fluidization pattern of a freely bubbling fluidized bed with different modes of air supply. Proceedings of the 63rd International Conference of Scandinavian Simulation Society, SIMS 2022, Trondheim, Norway, September 20-21, 2022, 192, 291–296.
- Moradi, A., Samani, N. A., Mojarrad, M., Sharfuddin, M., Bandara, J. C., and Moldestad, B. M. E. (2020). Experimental and computational studies of circulating fluidized bed. *International Journal of Energy Production and Management*, 5(4), 302–313.
- Pallarès, David., and Chalmers tekniska h ögskola. Department of Energy and Environment. Division of Energy Technology. (2008). Fluidized bed combustion: modeling and mixing. Chalmers University of Technology.
- O'Rourke, P. J., and Snider, D. M. (2014). A new blended acceleration model for the particle contact forces induced by an interstitial fluid in dense particle/fluid flows, *Powder Technology*, 256, 39–51,
- Snider, D. M. (2001). An Incompressible Three-Dimensional Multiphase Particle-in-Cell Model for Dense Particle Flows, *Journal of Computational Physics*, 170(2), 523–549.

# The Application and Advantages of a Generic Component-Based SI/CI Engine Model with VVA Compatibility<sup>\*</sup>

Oskar Lind Jonsson<sup>\*,\*\*</sup> Lars Eriksson<sup>\*</sup>

<sup>\*</sup> *Vehicular Systems, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden*

<sup>\*\*</sup> *oskar.lind.jonsson@liu.se*

---

**Abstract:** Engine models developed for control purposes are often developed with a time schedule in mind under the pressure of deadlines without reusability in mind and end up being hard-coded with a single engine type usage in mind. This approach can lead to more work when a new engine is created and a model is to be developed, as it usually takes less time than changing or modifying the old one. To facilitate a more rapid development process, there is a desire to have control-oriented models that can be adapted to new types of hardware with ease while at the same time providing fundamental insights into the physics of the engine that limit the control performance. The main idea is to use a component-based structure where the components are designed to be reused for similar processes; when combined, it constitutes a generic engine model with parametrization and compatibility with VVT/VVA and SI/CI combustion. An open-source mean value engine model was created in MATLAB/Simulink to meet the objectives. The engine model describes components such as the air filter, intercooler, and exhaust system components as incompressible flow restrictions. Bypass, throttle, intake/exhaust valves, and wastegate are modeled as compressible flow restrictions. Adiabatic control volumes are placed between each component to keep track of masses, pressures, and temperatures. The remaining components are modeled separately, with unique functions for each model, while integrated into the component-based structure. To demonstrate the concept and the generality of the approach, two engines, a 6-cylinder 12.7-liter Scania diesel engine, and a 4-cylinder 2.0-liter Volvo petrol engine, are used as case studies. The generic simulation platform is parameterized and validated against experimental data for both engines.

*Keywords:* Engine modeling, Parametrization, Mean value model, Component-based structure

---

## 1. INTRODUCTION

Engine models are often developed with a specific control purpose, sometimes even hard-coded for a single engine type. This approach presents a challenge when new engines are developed, as creating a new model is usually quicker than modifying an existing one. Control-oriented models that can easily adapt to new hardware types are needed to expedite development. These models should also provide fundamental insights into the engine physics that limits control performance. An open-source mean value MATLAB/Simulink model is created to meet these objectives. This generic engine model is parameterized and compatible with VVT/VVA and SI/CI combustion.

The model is built on a component-based structure designed for reusability. It includes components for the air filter, intercooler, and exhaust system, modeled as incompressible restrictions. Bypass, throttle, valves, and wastegate are modeled as compressible flow restrictions. Adiabatic control volumes are placed between each component

to monitor pressures and temperatures. The additional components are modeled separately with unique functions. To demonstrate the generality of this approach, we've used two engines as case studies: a 6-cylinder 12.7-liter Scania diesel engine and a 4-cylinder 2.0-liter Volvo petrol engine. The generic simulation platform was parameterized and validated against experimental data for both engines. This paper aims to create a generic engine model that captures the dynamics in both SI and CI engines with VVA. The goal is a model with a parameterizing structure that makes it easy to fit for different engines. The model will simulate the engine's behavior from the air inlet to the exhaust system.

This objective is then divided into three goals.

- Implement a model that can handle both SI and CI combustion with VVA.
- Utilize the ability to change model equations simulated easily.
- Structurally present equations, parameters, and validation.

---

<sup>\*</sup> This work was performed within the Competence Center SEDDIT-Sensor Informatics and Decision making for the Digital Transformation, supported by Sweden's Innovation Agency within the research and innovation program Advanced digitalization.

### 1.1 Reason for Research

The primary motivation for developing a more generic engine model is to facilitate the simulation of various engine configurations and dependencies, eliminating the need to create a new model each time. If a generic engine model can demonstrate a strong correlation with real engines across various engines and displacements. In that case, it can be inferred that the model applies to various engine configurations.

Such a model can be utilized in both engine simulations and controller development. It could also be employed in real-time applications, allowing for direct selection of control parameters in the Engine Control Unit (ECU).

To validate this model, it will be tested using two distinct engines: a 2.0-liter turbocharged petrol engine from Volvo and a 12.7-liter turbocharged diesel engine from Scania.

### 1.2 Related Research

In engine modeling, various models are utilized, including transfer function models, cycle mean value models (MVEM), zero- or one-dimensional models, and computational fluid dynamic models Theotokatos et al. (2018). MVEM compromises the simpler transfer function models and the more detailed zero- or one-dimensional models. These are particularly useful in the design of the engine control system, where rapid simulation times are crucial Theotokatos et al. (2018). This makes MVEM suitable for Hardware-in-the-Loop (HIL) simulations, as the model needs to capture engine behavior accurately Maroteaux and Saad (2015).

The MVEM approach is favored in modeling due to its low computational power requirements, stemming from using nonlinear ordinary differential equations Eriksson et al. (2002). This allows for faster-than-real-time simulations Llamas and Eriksson (2019), making it ideal for controller development and tuning. In 2001, a component-based MVEM was developed in Modelica. This included a detailed explanation of the model structures, the Modelica Language, and the models' validations. This methodology provides a deeper understanding of the potential of the component-based modeling approach Brugård et al. (2001). Component-based modeling involves utilizing derived models from literature and building your model one component at a time Eriksson (2003). In engine modeling, components can include the air filter, throttle, intake manifold, and cylinder. These models must then be validated to capture the system dynamics accurately.

Different parts of the engine model assume different types of flows, which can generally be categorized as compressible or incompressible. Most gas flows in engine pipes are incompressible and turbulent. However, a compressible flow should be assumed over most types of valves. This is particularly important for components like the throttle and the opening of an intake/exhaust valve Eriksson and Nielsen (2014).

Several open-source engine models are currently available. One such model is a modified MVEM for Spark Ignition (SI) engines with Exhaust Gas Recirculation (EGR) using the Simulink environment. The modification involved re-

placing the usual isothermal models for manifolds with a model for temperature dependency. The results found that neglecting the instances that change the input regime in the manifold air temperature is approximately constant Mostofi et al. (2006).

MVEM can also be used to validate models for control purposes, such as the validation of a control model for an electronic throttle Chaing et al. (2007).

LiU-Diesel is a MATLAB/Simulink model simulating a Diesel Engine with EGR and Variable Geometry Turbocharger (VGT) Wahlström and Eriksson (2011). Further development of the LiU-Diesel, called LiU-Diesel2, adds compatibility of Throttle and Turbocharger with Wastegate Ekberg et al. (2018).

### 1.3 Description of Modelling Concept

The MVEM is utilized to analyze and design control and diagnostic systems. This is due to the MVEM's ability to describe variations that occur faster than an engine cycle. MVEM models are sometimes called control-oriented models or filling and emptying models.

The process of averaging over one or several cycles in MVEM modeling necessitates a thorough investigation of physical sensor and actuator dynamics. These components are crucial as they actuate and measure the actual engine Eriksson and Nielsen (2014).

The MVEM will employ the 'filling and emptying' method for the control volumes, including the intake and exhaust manifold. This method divides volumes into sections, as the manifolds are represented as finite volumes Heywood (1988).

### 1.4 Model Limitations

Items that are outside the scope of this paper are:

- Fuel spray models
- Thermal stress and solid mechanics that affect geometries
- Validation of EGR model
- Warm-up process of the engine
- Formation of emissions
- Implementation of control system
- Variable compression by varying stroke length
- Start-stop technology and functionality

## 2. ENGINE COMPONENTS

The engine consists of multiple components. The components can be seen in Fig. 1, and the flow is positive when flowing to the right. The engine components include the air filter, compressor, intercooler, throttle, intake and exhaust, EGR, WG, turbine, after-treatment, and the rest of the exhaust system.

### 2.1 Parameter Estimation

There are three different types of measurement data

- Stationary (also called the engine map)
- Dynamic

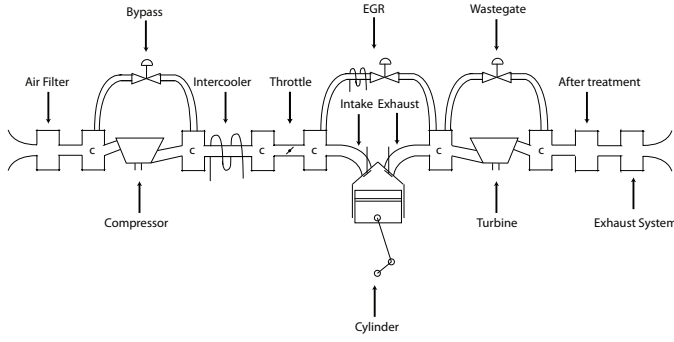


Fig. 1. The engine components include air filter, compressor, intercooler, throttle, intake and exhaust, EGR, WG, Turbine, after treatment, and the rest of the exhaust system. In this paper, the after-treatment and exhaust system is modeled as the "Exhaust system."

- Crankshaft based.

From that data, all the parameters can be found mainly using the least-squares method. The least-squares method can be divided into two categories: linear and non-linear. For a more detailed description of the least-squares method, see Björck (1996).

## 2.2 Incompressible Flow Model

There are different kinds of flow restrictions among the engine components. The flow can be assumed incompressible for the air filter, intercooler, and exhaust system (including the catalyst and muffler). This means that the throttle and cylinder flows will be modeled differently.

The model selected for the incompressible flow restrictions is found in Eriksson and Nielsen (2014)

$$\dot{m} = \begin{cases} C_{tu} \sqrt{\frac{p_{us}}{RT_{us}}} \sqrt{p_{us} - p_{ds}}, & \text{if } p_{us} - p_{ds} \geq \Delta p_{lin} \\ C_{tu} \sqrt{\frac{p_{us}}{RT_{us}}} \frac{p_{us} - p_{ds}}{\Delta p_{lin}}, & \text{otherwise} \end{cases} \quad (1)$$

$$T_{flow} = T_{us} \quad (2)$$

The parameters for this model are  $C_{tu}$  and  $\Delta p_{lin}$ .

If  $p_{us} = p_{ds}$  the Lipschitz condition is not fulfilled as the derivative goes towards infinity, which causes problems for the ODE solver. In this paper, ODE15s were used.  $p_{li}$  is the linear pressure region, and it is assumed  $p_{li} \geq 1000$  pa.

## 2.3 Control Volume Model

Control volumes can be divided into two different models. One is the isothermal model, where it can be assumed that the temperature is constant for the whole control volume. The other one is the adiabatic model, which means the heat transfer is often set to zero. All control volumes in the implemented model are adiabatic also to see temperature variations. The model implemented has pressure or mass and temperature as states,  $p$ ,  $m$ , and  $T$ .

$$\frac{dT}{dt} = \frac{RT}{pVc_v} \left[ \dot{m}_{in} c_v (T_{in} - T) + R(T_{in} \dot{m}_{in} - T \dot{m}_{out}) - \dot{Q} \right] \quad (3)$$

$$\frac{dp}{dt} = \frac{RT}{V} (\dot{m}_{in} - \dot{m}_{out}) + \frac{p}{T} \frac{dT}{dt} \quad (4)$$

$$\frac{dm}{dt} = \sum_i \dot{m}_i \quad (5)$$

The parameter for this model is  $V$ , the volume of each control volume.  $\dot{m}_{in}$  is the sum of all positive flows into the volume, and  $\dot{m}_{out}$  is the sum of all negative flows into the volume.  $T_{in}$  is the temperature of the gas flowing into the volume and is modeled as a mean value for all the flow flowing into the volume, assuming the same  $c_v$  for all flowing fluids in the volume. This means

$$T_{in} = \begin{cases} \frac{\dot{m}_1 \cdot T_1}{\dot{m}_{in}} + \frac{\dot{m}_2 \cdot T_2}{\dot{m}_{in}} + \dots + \frac{\dot{m}_n \cdot T_n}{\dot{m}_{in}} & \dot{m}_{in} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $n$  is the flow with the corresponding temperature for inflows into the control volume.

when the mass state is used, pressure is determined using the ideal gas law,

$$p = \frac{mRT}{V}. \quad (7)$$

If the control volume is placed in the intake or exhaust manifold, the gas composition,  $x$ , is modeled as the fractional content of the specimens in the total gas mixture,

$$\frac{dx}{dt} = \frac{1}{m} \sum_j (x_j - x) \dot{m}_j \quad (8)$$

where the index  $j$  indicates the contents of the specimen in the respective flow.

For the isothermal model, the incompressible flow restriction is

$$\frac{dp}{dt} = \frac{RT}{V} (\dot{m}_{in} - \dot{m}_{out}) \quad (9)$$

$$T = T_{in} = T_{out} \quad (10)$$

where the only parameter is  $V$ , the volume of each control volume.

## 2.4 Compressible Flow Model

The flow can be assumed to be compressible over the throttle, bypass, wastegate, and valves. Validation for the throttle can be found in Lind Jonsson (2021). The flow model is based on a compressible flow approach Eriksson and Nielsen (2014).

$$\Pi = \max \left( \frac{p_{ds}}{p_{us}}, \left( \frac{2}{\gamma + 1} \right)^{\frac{\gamma}{\gamma - 1}} \right), \quad (11)$$

$$\dot{m} = \frac{p_{us}}{\sqrt{RT_{us}}} A_{eff} \Psi_{li}(\Pi), \quad (12)$$

$$\Psi_0 = \sqrt{\frac{2\gamma}{\gamma - 1} \left( \Pi^{\frac{2}{\gamma}} - \Pi^{\frac{\gamma+1}{\gamma}} \right)}, \quad (13)$$

$$\Psi_{li} = \begin{cases} \Psi_0 & \text{if } \Pi \leq \Pi_{li} \\ \Psi_0 \frac{1 - \Pi}{1 - \Pi_{li}} & \text{otherwise,} \end{cases} \quad (14)$$



where the parameters are  $\Pi_{li}$  that is defining the linear region.

The effective area is

$$A_{eff} = A_0 + A_1\alpha_{th} + A_2\alpha_{th}^2 \quad \text{for throttle,} \quad (15)$$

$$= C_D \frac{D_v^2 \pi}{4} f([0..1]) \quad \text{for bypass/wastegate,} \quad (16)$$

$$= C_{D,in/ex} \frac{D_v^2 \pi}{4} \quad \text{for intake/exhaust valves.} \quad (17)$$

where  $C_D$  for the intake and exhaust valves are

$$C_{D,in} = f_{lookup}(L_{in}(\theta)) \quad (18)$$

$$C_{D,ex} = f_{lookup}(L_{ex}(\theta), Pr) \quad Pr = \frac{p_{cyl}}{p_{em}}. \quad (19)$$

## 2.5 Turbo

The compressor and turbine model is presented and validated in Lind Jonsson (2021) for the Volvo engine. The ellipse model tuned and validated in LiU CPGui Llamas and Eriksson (2018) was used for the Scania engine.

## 2.6 Turbo Dynamics Model

The turboshaft friction is modeled according to

$$T_{q_{tc},fric} = c_{tc}\omega_{tc} \quad (20)$$

where the friction is assumed to be low since the shaft is oil lubricated, meaning  $c_{tc} = 1 \cdot 10^{-6}$  Nm/(rad/s). Newton's second law of rotation is then used to model the turbo shaft speed

$$\frac{d\omega_{tc}}{dt} = \frac{1}{J_{tc}} (T_{q_t} - T_{q_c} - T_{q_{tc},fric}) \quad (21)$$

The parameters for this model are  $J_{tc}$ , which is the inertia of the shaft.

## 2.7 Intercooler Model

The Intercooler is modeled as an incompressible flow. What differentiates the intercooler from the incompressible flow model is the model for the temperature Eriksson and Nielsen (2014).

$$T_o = T_i - \epsilon (\dot{m}_{cool}, \dot{m}_{ic}, \dots) (T_i - T_{cool}) \quad (22)$$

$$\epsilon = a_0 + a_1 \left( \frac{T_i + T_{cool}}{2} \right) + a_2 \dot{m}_{air} \quad (23)$$

where  $T_o$  is the temperature of the gas flowing out of the intercooler,  $T_i$  is the temperature of the gas flowing into the intercooler,  $T_{cool}$  is the temperature of the air surrounding the intercooler, which in this model is set to  $T_{amb}$ ,  $\dot{m}_{air}$  is the mass flow of the gas flowing into the intercooler and  $\dot{m}_{cool}$  is the mass flow of the gas cooling the intercooler. The parameters for this model are  $a_0$ ,  $a_1$  and  $a_2$ .

## 2.8 Cylinder Model

To fully capture the dynamics of the oxygen concentration, cylinder- temperature, and pressure using VVA, each cylinder has to be simulated separately with full-cycle cylinder states Johansson (2019).

*Cylinder Flow* Four different flows—intake, exhaust, CRB, and blowby—are modeled as the compressible flow restriction.

*Combustion Modeling* Combustion is affected by the amount of available oxygen. Therefore, the oxygen concentration  $X_{O,cyl}$  is a state. This also gives the  $X_{burned} = 1 - X_{O,cyl} - X_{fuel}$ , which is the amount of oxygen that has reacted in the combustion. A perfect stoichiometric combustion is assumed, meaning particle formation has been neglected. Combustion is assumed to be a reaction between air and hydrocarbons.

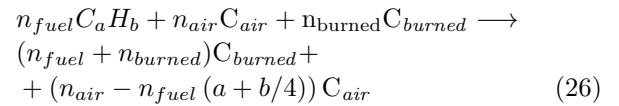
As the most common element in the air is nitrogen, all elements except oxygen are clumped together, creating a common assumption Eriksson and Nielsen (2014)

$$C_{air} = O_2 + 3.773N_2 \quad (24)$$

Exhaust composition depends on what fuel is assumed to be burning during combustion. The fuel used is cetane for CI combustion and isooctane for SI combustion. The burned composition is.

$$C_{burned} = aCO_2 + b/2H_2O + 3.773(a + b/4N_2) \quad (25)$$

The chemical reaction in the combustion is



where the fuel composition is  $C_aH_b$ .

The number of moles in the reaction is calculated at IVC.

$$n_{tot,air} = 1 + 3.773 \quad n_{tot,burned} = a + \frac{b}{2} + 3.773 \quad (27)$$

$$n_{air} = \frac{m_{IVC}X_{O,cyl}}{M_{air}n_{tot,air}} \quad n_{fuel} = \frac{m_{fuel}}{M_{fuel}} \quad (28)$$

$$n_{burned} = \frac{(1 - X_{O,cyl})m_{IVC}}{M_{burned}n_{tot,burned}} \quad (29)$$

where  $m_{IVC}$  is the mass in the cylinder at IVC. The mass in the cylinder is calculated by

$$m_{cyl} = \frac{p_{cyl}V_d}{R_{cyl}T_{cyl}} \quad (30)$$

The  $X_{O,cyl}$  state is modelled similarly as the fractions  $x$  in the manifolds, with the stoichiometric combustion assumption

$$\frac{dx}{dt} = \frac{1}{m_{cyl}} \sum_i (x_i - x) \dot{m}_i - S_{AF_s} C \frac{dx_b}{dt} \quad (31)$$

where the Vibe function in equation (33) gives  $\frac{dx_b}{dt} = \frac{dx_b}{d\theta} \omega_e$ ,  $S_{AF_s}$  scales the composition with regards to the fraction burned, assuming  $(AF_s - 1)$  used air per 1 used fuel and  $C$  is the scaling of the Vibe function to make the Vibe function scale properly to equation (26).

$$x_b(\theta) = \begin{cases} 0, & \theta < \theta_{SOC} \\ 1 - e^{-a \left( \frac{\theta - \theta_{SOC}}{\Delta\theta} \right)^{m+1}}, & \theta \geq \theta_{SOC} \end{cases} \quad (32)$$

$x_b$  determines the fraction burned, from 0 to 1.  $\Delta\theta$  and  $a$  are related to combustion duration,  $m$  affects the shape and  $\theta_{SOC}$  denotes the start of combustion.

For heat release calculations, the Vibe function is derived

$$\frac{dx_b\theta}{d\theta} = \frac{a(m+1)}{\Delta\theta} \left( \frac{\theta - \theta_{SOC}}{\Delta\theta} \right)^m e^{-a \left( \frac{\theta - \theta_{SOC}}{\Delta\theta} \right)^{m+1}} \quad (33)$$

The molar fraction  $x_{air}$  is the amount of free air that has not yet reacted in the combustion. The amount of free air is directly correlated to the amount of oxygen according to the air assumption in equation (24). The scaling factor  $C$  is modeled as the amount of air at IVC and air left after combustion.

$$n_{comb} = n_{air} - n_{fuel} (a + b/4) \quad (34)$$

$$\tilde{x}_{air} = \frac{n_{comb} n_{tot,air}}{n_{comb} n_{tot,air} + n_{comb} n_{tot,Burned}} \quad (35)$$

$$x_{air} = \frac{\tilde{x}_{air} M_{air}}{(1 - \tilde{x}_{air}) M_{Burned} + \tilde{x}_{air} M_{air}} \quad (36)$$

This gives the scaling of the Vibe function

$$C = x_{air,IVC} - x_{air,aC} \quad (37)$$

Since the gas constant and heat capacity change with oxygen concentration and temperature, NASA polynomials are used to decide the heat capacities  $c_{p,Burned}$  and  $c_{p,air}$  according to equation (38).

The specific heat ratio depends on the temperature. NASA polynomial is a database for how different chemical species' specific heat ratio changes with temperature. The NASA polynomial can be read more in McBride (2002).

$$\frac{\tilde{c}_p(T)}{\tilde{R}} = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4 \quad (38)$$

The gas mass-specific constant and specific heat is calculated as

$$R = (1 - X_{o,cyl}) R_{Burned} + X_{o,cyl} R_{air} \quad (39)$$

$$c_p = (1 - X_{o,cyl}) c_{p,Burned} + X_{o,cyl} c_{p,air} \quad (40)$$

The ideal gas assumption is used to get the specific heat constant  $c_v$

$$c_v = c_p - R \quad (41)$$

The parameters for the IVC event use intake temperature from the engine map, but the rest of the measurements are based on crankshaft measurements. The parameters for the Vibe function use pressures and temperatures from the crankshaft-based measurements.

*Energy equations* The energy flows in the cylinder are heat release, heat transfer, the work carried out, and internal energy. The heat release is calculated

$$\frac{dQ_{HR}(\theta)}{d\theta} = m_f Q_{LHV} \eta_{co} \frac{dx_b(\theta)}{d\theta} \quad (42)$$

The heat transfer is calculated using the Woschini method. However, the model for the heat transfer coefficient is taken from Eriksson and Nielsen (2014)

$$h = C_0 B^{-0.2} p^{0.8} w^{0.8} T^{-0.53} \quad (43)$$

where  $C_0 = 1.30 \cdot 10^{-2}$ ,  $B$  is the cylinder bore,  $p$  is the pressure in the cylinder,  $T$  is the temperature in the cylinder and  $w$  is the characteristic velocity equation given by

$$w = C_1 \bar{S}_p + C_2 \frac{VT_{IVC}}{V_{IVC} p_{IVC}} (p - p_m) \quad (44)$$

where  $\bar{S}_p = \frac{2aN_e}{60}$  is the mean piston speed,  $p_m$  is the motored pressure, and the constants  $C_1$  and  $C_2$  are dependent

on what stroke the engine has, and can be seen in Table 1.

Table 1. Constants used in Woschini's model for the heat transfer coefficient.

	Gas Exchange	Compression Compression	Combustion and Expansion
$C_1$	6.18	2.28	2.28
$C_2$	0	0	0.00324

The motored pressure is modeled with a polytope, where  $\kappa$  is the polytropic exponent.

$$p_m(\theta) = \begin{cases} p(\theta), & \text{if } \theta \leq \theta_{SOC} \\ p(\theta_{SOC}) \left( \frac{V(\theta_{SOC})}{V(\theta)} \right)^\kappa & \text{if } \theta > \theta_{SOC} \end{cases} \quad (45)$$

where the polytropic exponent  $\kappa$  is set to a constant value, optimized from a motored cycle. Another way to get the motored pressure is to use the pressure from a measurement in an engine test cell.

The power is calculated using Newton's second law of motion.

$$\dot{W} = p_{cyl} \frac{dV}{dt} \quad (46)$$

and the internal energy is modeled

$$u = c_v T_{cyl} \quad (47)$$

*State equations* The temperature and pressure states are  $T_{cyl}$  and  $p_{cyl}$ .

$$h = c_p T_{us} \quad (48)$$

$$\dot{T}_{cyl} = \frac{R_{cyl} T_{cyl}}{p_{cyl} V c_{v,cyl}} \left( \dot{m}_{intake} (h_{intake} - u) - \sum_i \dot{m}_i (h_i - u) + \frac{dQ_{HR}(\theta)}{dt} - \dot{Q}_{HT} - W \right) \quad (49)$$

$$\dot{p}_{cyl} = \frac{R_{cyl} T_{cyl}}{V} \left( \dot{m}_{intake} - \sum_i \dot{m}_i \right) + \frac{p_{cyl}}{T_{cyl}} \dot{T}_{cyl} - \frac{W}{V} \quad (50)$$

where  $i$  is the flow for exhaust, CRB, and blowby.

*Generated Engine Torque* A simple instantaneous torque model that neglects friction is used in this paper Eriksson and Nielsen (2014)

$$M_{e,i}(\theta) = \sum_{j=1}^{n_{cyl}} (p_{cyl,j} (\theta - \theta_j^{\text{offset}}) - p_{amb}) A L (\theta - \theta_j^{\text{offset}}) \quad (51)$$

where  $\theta_j^{\text{offset}}$  is the crank angle offset for cylinder  $j$ ,  $A$  is the area and  $L(\theta)$  is the crank lever.

$$A L(\theta) = \frac{dV(\theta)}{d\theta} = \frac{dV(\theta)}{dt} \frac{1}{\omega_e} \quad (52)$$

For example, the average torque is of interest so that it does not exceed any loads for the driveline. The average torque is then calculated over four strokes Eriksson and Nielsen (2014).

where  $M_f$  denotes the friction, neglected in this paper.

### 2.9 Complete Model

The complete Simulink model can be seen in Fig. 2 and is run by first running the `init.m` file setting up the necessary model input and loading the parameter struct.

## 3. RESULTS

### 3.1 Validation Points

Stationary measurements are done after the engine has stabilized, meaning all dynamic behavior has had time to subside. Typically, measurements are taken over a brief period, usually a few seconds, and then an average value for that specific operating point is calculated and recorded. The stationary measurements are systematically conducted in a series, moving from one operating point to the next, ensuring comprehensive data collection. An example of a complete engine map can be seen in Fig. 3. The middle point is presented in this paper, and the other operating points are given in Lind Jonsson (2021).

During the validation, pressure is used instead of mass, as it is easier to relate to pressures than masses. Some components were not present on the Volvo and Scania engines and were not simulated. This meant some inputs were kept the same for each iteration. The inputs for the Scania engine were kept the same: throttle angle, wastegate, bypass, and EGR actuation, and they were set to zero. The inputs for the Volvo engine that were kept the same were bypass, EGR, and CRB actuation, which were set to zero. Setting inputs to zero has the same effect as if they were non-existent.

The middle operating point is presented here to validate the model, while all are covered in Lind Jonsson (2021). In Tables 3 to 4 below the parameters are pressure, temperature, turbo shaft speed, engine torque, mass flows and oxygen concentration. The parameters for pressure are after the compressor,  $p_c$ , in the intake manifold,  $p_{im}$ , and after the exhaust manifold,  $p_{em}$ . The temperature parameters are after the compressor,  $T_c$ , in the intake manifold,  $T_{im}$ , after the exhaust manifold,  $T_{em}$ , and after the turbine,  $T_t$ . The parameter for the turbo shaft speed is  $w_{tc}$ , and the parameter for the generated engine torque is  $T_{qe}$ . The parameter for oxygen concentration is after the exhaust manifold,  $Xo_{em}$ , but a reference for this measurement is only available for the Scania engine. The units used to validate the model are *bar* for pressure,  $^{\circ}C$  for temperature, *kRPM* for turbo speed, *Nm* for generated torque, *g/s* for mass flows and *%* for oxygen concentration.

Results are presented when the model has reached a steady value. The simulation time was set to 10 seconds, ensuring that all dynamics were gone at the end of the simulation. Measured data was investigated on a cycle-to-cycle basis. The simulation reached a steady state after about 50 cycles. The results presented below are taken at the 9.5-second mark for the cycle data and a mean of the last 50 simulated cycles for the Volvo engine and the last 25 simulated cycles for the Scania engine for mean and max error. This is due to the difference in engine speed between the engines. The cycle data is over 720 degrees, a complete cycle for the four-stroke engine.

### 3.2 Middle Operating Point

The non-constant parameters needed to run this operational point for the different engines are presented in Table 2. In Table 3 and 4, mean value data is presented for different parameters, and in Figs. 6 and 7 cycle data is presented for the Scania and Volvo engine respectively.

Table 2. Operating points run for Scania and Volvo engines of the middle operating point.

Parameter	Scania	Volvo
Engine Speed [RPM]	1300	2250
Throttle Angle [Deg]	-	8.42
Fuel Flow [mg/stroke]	126.49	17.97
SOI [Deg]	-8.04	-8.00
Crank Angle Intake Offset [Deg]	15.06	-48.00
Crank Angle Exhaust Offset [Deg]	-15.00	30.00

Table 3. Mean error and max error for different parameters for the Scania engine on the mean operating point.

Parameter	Modelled value	Measured value	Units	Mean error [%]
$p_c$	1.76	2.02	bar	-13.00
$p_{im}$	1.72	1.98	bar	-13.02
$p_{em}$	1.68	1.80	bar	-7.14
$T_c$	105.43	104.32	C	1.06
$T_{im}$	27.83	26.56	C	4.77
$T_{em}$	353.55	371.50	C	-4.83
$T_t$	310.33	330.88	C	-6.21
$w_{tc}$	65.28	70.50	kRPM	-7.40
$T_{qe}$	857.17	1249.92	Nm	-31.42
Lambda	1.85	2.24	-	-17.20
$W_{af}$	247.16	266.72	g/s	-7.33
$W_c$	247.16	266.72	g/s	-7.33
$W_{ic}$	247.13	266.72	g/s	-7.34
$W_{cyl\_in}$	247.12	266.72	g/s	-7.35
$W_{cyl\_out}$	262.29	348.93	g/s	-24.83
$W_t$	262.76	348.93	g/s	-24.69
$W_{es}$	262.84	348.93	g/s	-24.67
$Xo_{em}$	42.29	12.00	%	30.29

Table 4. Mean error and max error for different parameters for the Volvo engine on the mean operating point.

Parameter	Modelled value	Measured value	Units	Mean error [%]
$p_c$	1.03	1.03	bar	0.00
$p_{im}$	0.77	0.76	bar	0.95
$p_{em}$	1.06	1.03	bar	2.52
$T_c$	37.92	37.79	C	0.35
$T_{im}$	26.55	32.45	C	-18.20
$T_{em}$	883.79	665.21	C	32.86
$T_t$	881.55	637.69	C	38.24
$w_{tc}$	24.19	32.84	kRPM	-26.34
$T_{qe}$	99.78	79.86	Nm	24.95
Lambda	1.04	0.99	-	4.60
$W_{af}$	15.23	18.75	g/s	-18.77
$W_c$	15.23	18.75	g/s	-18.77
$W_{ic}$	15.23	18.75	g/s	-18.77
$W_{th}$	15.23	18.75	g/s	-18.77
$W_{cyl\_in}$	15.23	18.75	g/s	-18.79
$W_{cyl\_out}$	17.80	19.09	g/s	-6.78
$W_t$	17.80	19.09	g/s	-6.77
$W_{es}$	17.70	19.09	g/s	-6.77

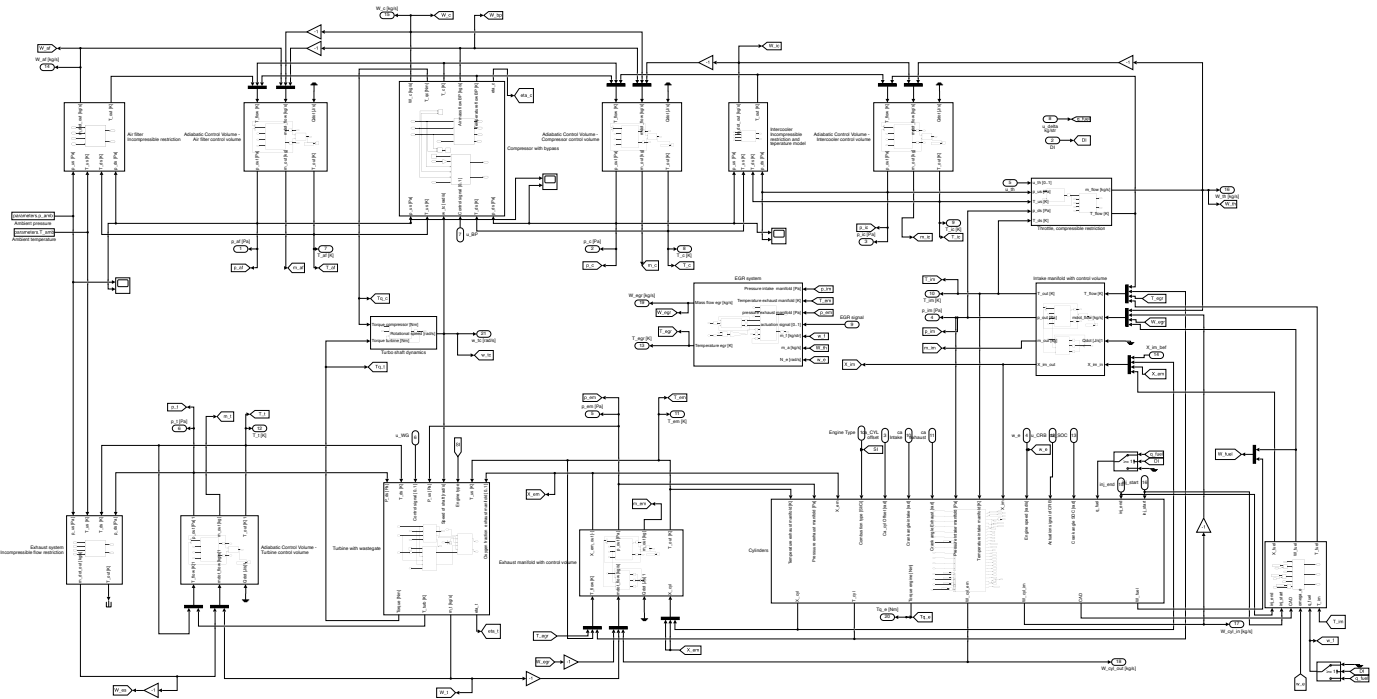


Fig. 2. The layout of the Simulink model

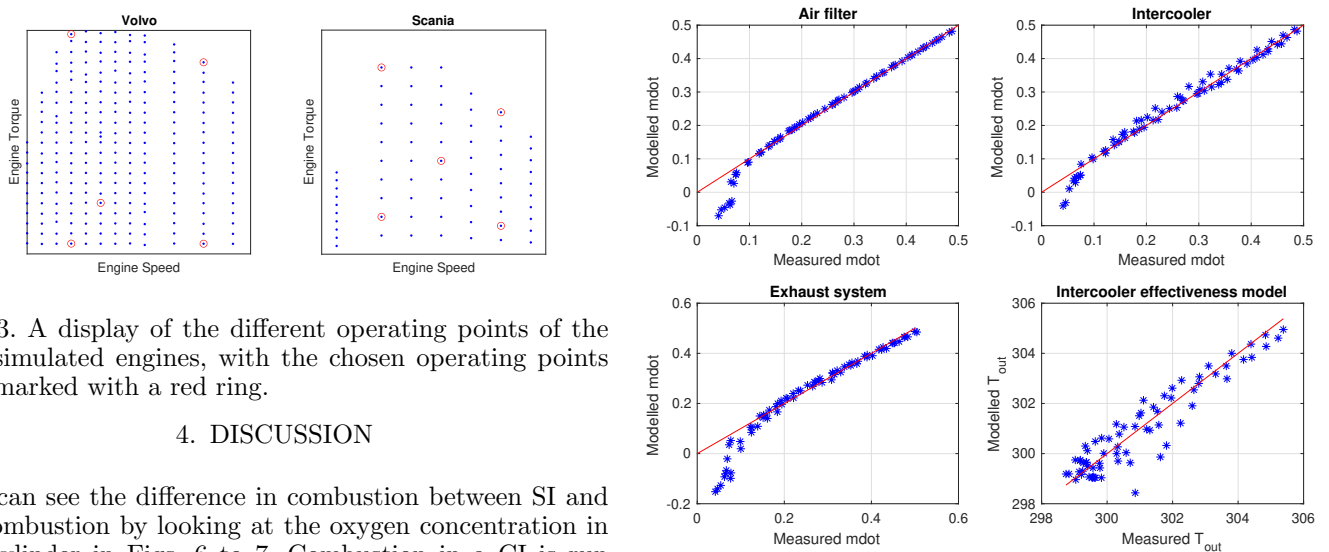


Fig. 3. A display of the different operating points of the simulated engines, with the chosen operating points marked with a red ring.

4. DISCUSSION

One can see the difference in combustion between SI and CI combustion by looking at the oxygen concentration in the cylinder in Figs. 6 to 7. Combustion in a CI is run lean to decrease the chances of creating particles, which also is a reason to use the EGR to reduce the available oxygen Eriksson and Nielsen (2014). In these simulations, the available oxygen is not combusted fully, leading to free air levels not going down to zero after combustion. All oxygen is combusted for SI, meaning the available oxygen is 0% at the combustion's end. The fill-up of available oxygen also follows the effective area. This proves the model can handle fresh air flow and residual gasses.

As seen in the Tables 3 to 4, the simulated pressures are closer for the Volvo model than the Scania model. This is due to the throttle and wastegate regulators regulating the pressures to the measured and desired value used in the Volvo model. Still, temperatures are closer for the Scania engine than the Volvo engine. The temperature deviation is because there is no direct correlation between pressure and temperature, so even if a regulator regulates

Fig. 4. The air filter, intercooler, and exhaust system flow validation, validating  $C_{tu,af}$ ,  $C_{tu,in}$  and  $C_{tu,ex}$  for the Scania engine.

the pressure, it does not mean the temperature deviation will also improve.

Cylinder pressures are the only reference measurement available with cycle-to-cycle variations. In Figs. 6 and 7, they are used to validate combustion parameters. The vibrate parameters are set for one operating point and not changed when the operating point is changed. As combustion varies depending on load, the vibrate parameters should also change. This was, however, a simplification made.

Another aspect differing from measured data is the model for the generated engine torque. This is due to the model not simulating engine friction. Most engine torques are

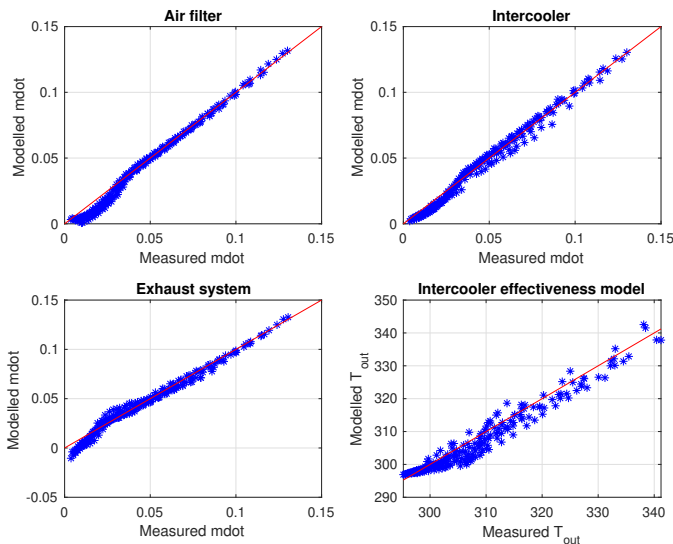


Fig. 5. The air filter, intercooler, and exhaust system flow validation, validating  $C_{tu,af}$ ,  $C_{tu,in}$  and  $C_{tu,ex}$  as well as the intercooler effectiveness model,  $a_0$ ,  $a_1$  and  $a_2$  for the Volvo engine. The red line marks the perfect model, meaning modeled equalling measured.

lower than measured because of how the model has been simulated. As the model uses variable step lengths, the cycle length in the measured data is often less than 720 degrees. However, if the cycle length were increased by one step, it would be longer than 720 degrees. The mass balance is fulfilled if the model is simulated with mass instead of pressure as a state. Currently, the model is tuned, focusing on the individual components. A better agreement between the model and data can be achieved if the total system model is fine-tuned with complete system behavior, as in Ekberg et al. (2018).

## 5. CONCLUSIONS

This paper aims to propose a generic engine model with VVA compatibility. Generic means the possibility of easily changing the equations used for each component and removing some components altogether and the possibility of interchange between CI and SI combustion. This was completed by making a separate MATLAB equation for each component, as each component is represented by one Simulink block. The generality is proven by the fact that the Volvo engine is SI and the Scania engine is CI, as well as by removing and exchanging components for the Scania engine, which used a different compressor and turbine function. That was fulfilled by changing the equations run in MATLAB and rerouting some signals in Simulink. As an added benefit of using this generic model, equations can easily be changed to improve the performance of the models used in this paper.

## ACKNOWLEDGEMENTS

This work was performed within the Competence Center SEDDIT-Sensor Informatics and Decision making for the Digital Transformation, supported by Sweden's Innovation Agency within the research and innovation program Advanced digitalization.

## REFERENCES

- Björck, Å. (1996). *Numerical methods for least squares problems*. SIAM.
- Brugård, J., Eriksson, L., and Nielsen, L. (2001). Mean value engine modeling of a turbo charged spark ignited engine - A principle study. Technical Report LiTH-R-2370, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden.
- Chaing, W.P., Zhu, L., and Patankar, R. (2007). Mean value engine modeling and validation for a 4-stroke, single cylinder gasoline engine. *Trends in Applied Sciences Research*, 2(2), 124–131.
- Ekberg, K., Leek, V., and Eriksson, L. (2018). Modeling and Validation of an Open-Source Mean Value Heavy-Duty Diesel Engine Model. *Simulation Notes Europe (SNE)*, 28(4), 198–203. doi:10.11128/sne.28.4.1044.
- Eriksson, L. (2003). Mean value models for exhaust system temperatures. *SAE 2002 Transactions, Journal of Engines*, 2002-01-0374, 111(3).
- Eriksson, L. and Nielsen, L. (2014). *Modeling and Control of Engines and Drivelines*. John Wiley Sons Inc.
- Eriksson, L., Nielsen, L., Brugård, J., Bergström, J., Petterson, F., and Andersson, P. (2002). Modeling of a turbocharged SI engine. *Annual Reviews in Control*, 26(1), 129–137.
- Heywood, J. (1988). *Internal combustion engine fundamentals*. McGraw-Hill.
- Johansson, L. (2019). *Full Cycle Cylinder State Estimation in DI Engines with VVA*. Master's thesis, Linköping University, Vehicular Systems.
- Lind Jonsson, O. (2021). *A GENERIC Internal Combustion Engine model : LiU-Genie*. Master's thesis, Linköping University, Vehicular Systems.
- Llamas, X. and Eriksson, L. (2018). Liu cpgui: A toolbox for parameterizing compressor models. Technical Report 3102, Linköping University, Vehicular Systems.
- Llamas, X. and Eriksson, L. (2019). Control-oriented modeling of two-stroke diesel engines with exhaust gas recirculation for marine applications. *Journal of Engineering for the Maritime Environment*, 223(2), 551–574.
- Maroteaux, F. and Saad, C. (2015). Combined mean value engine model and crank angle resolved in-cylinder modeling with nox emissions model for real-time diesel engine simulations at high engine speed. *Energy*, 88, 515–527.
- McBride, B.J. (2002). *NASA Glenn coefficients for calculating thermodynamic properties of individual species*. National Aeronautics and Space Administration, John H. Glenn Research Center . . . .
- Mostofi, M., Shamekhi, A., and Gorji-Bandpy, M. (2006). Modified mean value si engine modeling (egr included). *Fuel*, 10, 5.
- Theotokatos, G., Guan, C., Chen, H., and Lazakis, I. (2018). Development of an extended mean value engine model for predicting the marine two-stroke engine operation at varying settings. *Energy*, 143(15).
- Wahlström, J. and Eriksson, L. (2011). Modeling diesel engines with a variable-geometry turbocharger and exhaust gas recirculation by optimization of model parameters for capturing non-linear system dynamics. *Journal of Automobile Engineering*, 225(7).



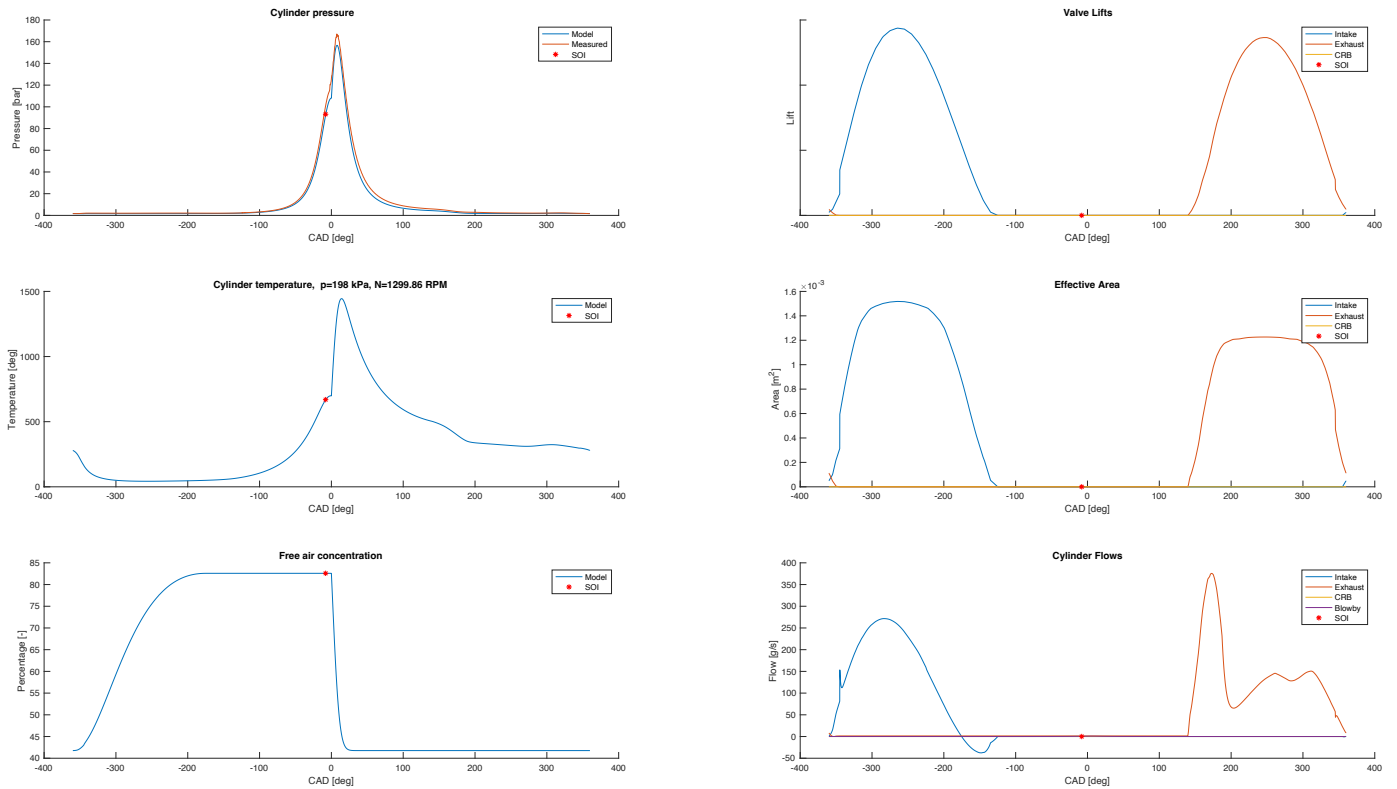


Fig. 6. Cylinder pressure, temperature, oxygen concentration, valve lift, effective area for the valves, and the cylinder flows over one complete cycle on the mean operating point for the Scania engine.

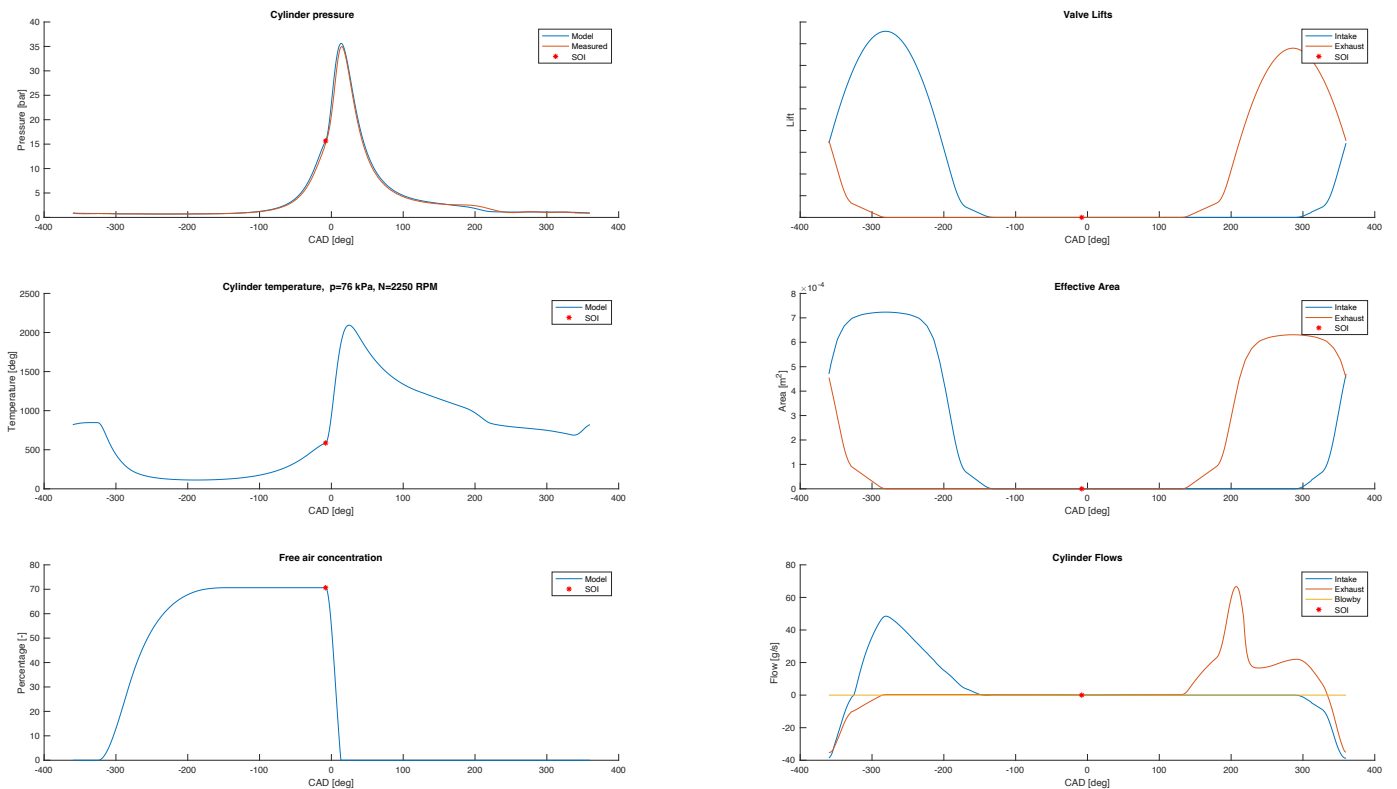


Fig. 7. Cylinder pressure, temperature, oxygen concentration, valve lift, effective area for the valves, and the cylinder flows over one complete cycle on the mean operating point for the Volvo engine.

# Modeling of a tire mounted energy harvester using an inertial and analytical tire deformation model

Mikko Leinonen Jaakko Palosaari Jari Juuti

*Microelectronics research unit, University of Oulu, Finland (e-mail: mikko.leinonen@oulu.fi).*

**Abstract:** In this work, an analytical tire deformation model is created, which can be parameterized using simple measurements. The model consists of three equations which are solved to provide a shape function for the tire.

This model can be used to provide excitation input for energy harvesters embedded inside the tire for example in FEM simulations. Additionally the model can be used in differential equation based simulations for quick parameterized simulations. With this model it is possible to study the effect of tyre inflation state to the energy harvesting performance of the system.

Two different simulation cases are presented in this work. First is a vibration energy harvester simulation using the model with an inertial energy harvester. The second case illustrates an energy harvester using the deformation of the tire as the excitation for the energy harvester as opposed to inertial type harvester.

*Keywords:* piezoelectric, TPMS, energy harvester, tire, automotive, FEM

## 1. INTRODUCTION

Tire inflation sensors or TPMS (Tire Pressure Monitoring System) sensors are currently widely used in the automotive industry. These sensors monitor the pressure inside the tire to improve fuel efficiency and to warn the driver of a tire puncture. Nowadays the pressure sensors employed in the tires use batteries that can last for years, however the batteries add weight and volume to the sensor module and the sensor uses  $0.45mW$  of energy when active Yi et al. (2021). Energy harvesting is a solution to minimize battery usage of the sensor. Several different energy harvesting schemes have been studied. Kinetic energy of the tire movement is converted into electrical energy mainly in two ways - inertially and using strain. Inertial energy harvesters usually couple a seismic mass to a piezoelectric beam and convert the movement of the inertial mass into electricity. In Kubba and Jiang (2014) simple cantilever beams are used to obtain energy from the tire motion. In Leinonen et al. (2017), however, a more general rotational harvester is studied where the inertia of the seismic mass is excited by the rotating gravitational vector. It has to be noted, that inertial energy harvesters are highly sensitive to frequency, as the harvesting is most efficient at the resonant frequency of the piezo beam.

The second type of piezoelectric energy harvesters for the TPMS sensors are the strain type energy harvesters. In these harvesters the deformation of the tire is used to actuate, for example, a piezo fiber patch Lee and Choi (2014). Another example is a flexible PVDF patch used in Maurya et al. (2018). It is also possible to use a more complicated structure such as in Esmaeeli et al. (2019), where a cymbal structure is embedded into a tire. This approach is also studied in this work, with the exception that the cymbal in

this research has symmetrical endcaps whereas in Esmaeeli et al. (2019) and Aliniagerdroudbari et al. (2019) only one of the endcaps is typical cymbal endcap. In Al-Najati et al. (2024) a strain energy harvester was presented and both the tire and harvester were modeled using FEM software. Furthermore, in Al-Najati et al. (2024) it is claimed that double endcap structures cannot be used inside tire, but in this study this is proved inaccurate. In Staaf et al. (2024) tire mounted energy harvester is simulated in FEM by modeling a section of a tire with rotating contact patch.

## 2. THEORY

### 2.1 A simple tire deformation model

A mathematical piecewise model was generated for a tire mounted harvester. A mathematical expression was developed for the shape of the tire.

The model for the local tire radius is as follows (for  $0 \leq \theta \leq 2\pi$ )

$$r(\theta) = \begin{cases} 0 \leq |\theta| \leq \theta_1 & , h/\cos(|\theta|) \\ \theta_1 \leq |\theta| \leq \theta_2 & , D + C|\theta| + B\theta^2 + A|\theta|^3 \\ |\theta| \geq \theta_2 & , 1 \end{cases} \quad (1)$$

where  $\theta$  is the rotation angle and  $A, B, C$  and  $D$  are the coefficients to be determined according to the deformation shape. The loaded tire height  $h$  is the fraction of the radius of the tire. The tire radius is scaled to 1 in this equation. The shape was varied by modifying the tire height reduction under load and the length of the flat area under the tire. The  $\theta_1$  and  $\theta_2$  are the angles between which the transient equation is valid i.e. the space between the flat area and the circular area. These result in a series of equations which result in an analytical expression for

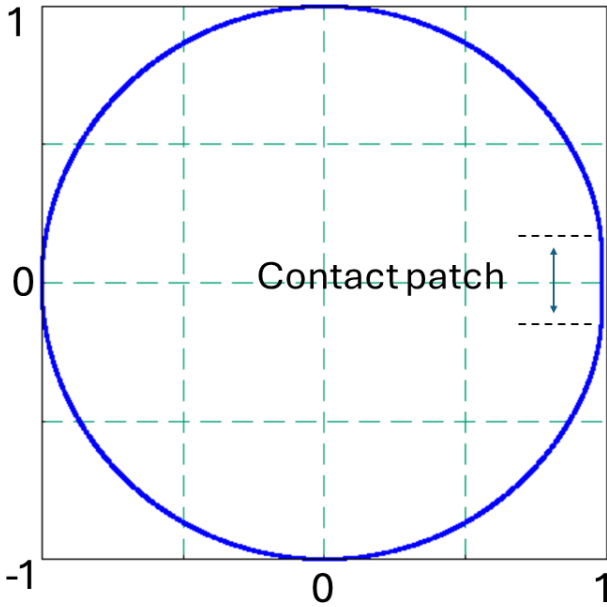


Fig. 1. Tire shape

different cases. An example is shown in Fig. 1, where the tire contact point is on the right ( $\theta=0$ ). The transient equation between  $\theta_1$  and  $\theta_2$  is constructed by forcing the function to be continuous and "smooth" i.e. differentiable over first derivative.

$$\begin{cases} \lim_{\theta \rightarrow \theta_1^-} r(\theta) = \lim_{\theta \rightarrow \theta_1^+} r(\theta) \\ \lim_{\theta \rightarrow \theta_2^-} r(\theta) = \lim_{\theta \rightarrow \theta_2^+} r(\theta) \\ \lim_{\theta \rightarrow \theta_1^-} \dot{r}(\theta) = \lim_{\theta \rightarrow \theta_1^+} \dot{r}(\theta) \\ \lim_{\theta \rightarrow \theta_2^-} \dot{r}(\theta) = \lim_{\theta \rightarrow \theta_2^+} \dot{r}(\theta) \end{cases} \quad (2)$$

When these equations are solved, we get the coefficients  $A, B, C$  and  $D$  for equation 1 as equations 3, 4, 5 and 6. This model provides a continuous and "smooth" shape curve for the tire. However it is not physically accurate, for example, it does not keep the circumference of the tire constant. However, it is possible to use more physically accurate boundary conditions to obtain more accurate functions. In our purpose to have a simple model for quick initial estimates on energy harvester performance on more "macroscopic" aspects of tire dynamics, the model is adequate. This model is also easy to implement on numerical simulations using differential equations for the energy harvester dynamics. This is also important when estimating initial energy harvester performance using FEM modeling. This model is easily incorporated into an FEM model as opposed to more complex models.

The parameters of this model can be estimated using simple measurements for the tire geometry. The length

$$A = \frac{\tan(\theta_1)\theta_2 h + (2 - \theta_1 \tan(\theta_1)) h - 2 \cos(\theta_1)}{\cos(\theta_1)\theta_2^3 - 3\theta_1 \cos(\theta_1)\theta_2^2 + 3\theta_1^2 \cos(\theta_1)\theta_2 - \theta_1^3 \cos(\theta_1)} \quad (3)$$

$$B = -\frac{\theta_2((3 - \theta_1 \tan(\theta_1)) h - 3 \cos(\theta_1)) + 2 \tan(\theta_1)\theta_2^2 h + (3\theta_1 - \theta_1^2 \tan(\theta_1)) h - 3\theta_1 \cos(\theta_1)}{\cos(\theta_1)\theta_2^3 - 3\theta_1 \cos(\theta_1)\theta_2^2 + 3\theta_1^2 \cos(\theta_1)\theta_2 - \theta_1^3 \cos(\theta_1)} \quad (4)$$

$$C = \frac{\theta_2((6\theta_1 - 2\theta_1^2 \tan(\theta_1)) h - 6\theta_1 \cos(\theta_1)) + \tan(\theta_1)\theta_2^3 h + \theta_1 \tan(\theta_1)\theta_2^2 h}{\cos(\theta_1)\theta_2^3 - 3\theta_1 \cos(\theta_1)\theta_2^2 + 3\theta_1^2 \cos(\theta_1)\theta_2 - \theta_1^3 \cos(\theta_1)} \quad (5)$$

$$D = -\frac{(\theta_1 \tan(\theta_1) - 1)\theta_2^3 h + (3\theta_1 - \theta_1^2 \tan(\theta_1))\theta_2^2 h - 3\theta_1^2 \cos(\theta_1)\theta_2 + \theta_1^3 \cos(\theta_1)}{\cos(\theta_1)\theta_2^3 - 3\theta_1 \cos(\theta_1)\theta_2^2 + 3\theta_1^2 \cos(\theta_1)\theta_2 - \theta_1^3 \cos(\theta_1)} \quad (6)$$

of the contact patch can also be measured directly or by using accelerometer or strain measurements. The strain measurements can also be used to measure the length of the transitional zone bounded by  $\theta_1$  and  $\theta_2$ .

## 2.2 An inertial harvester inside a tire

An inertial harvester is shown in Fig. 2. The PZT beam harvester is mounted in the inside of the tire and an endmass is mounted at the tip. The beam is bending by the rotating gravity vector as the tire turns ( $\theta(t) = \omega t$ ) if the tire is completely round. However, in practice the tire has a flat contact area and therefore it is not perfectly round. Furthermore, the harvester is subject to a centripetal force, which in the harvesters own frame of reference acts as a virtual gravity. In a case of a perfectly round tire's case, it's magnitude is

$$a_{vg} = \omega^2 r \quad (7)$$

,where  $\omega$  is the angular velocity and the  $r$  is the radius of the circle.

In the case of a loaded tire, with a flat contact patch, this equation 7, holds for the circular part of the tire, and for the flat part, the virtual gravity is 0, and only the actual gravity is in effect. This is easily proven, as the velocity vector of the harvester stays constant in cartesian coordinates as opposed to polar coordinates in equation 1. As the velocity vector does not change it's direction, the centripetal acceleration is zero. For the polynomial part of the equation 1, the centripetal force could be calculated, but the result is too verbose to reproduce here. Furthermore, as is shown in the measurements a simple linear or step approximation is enough for a simple model.

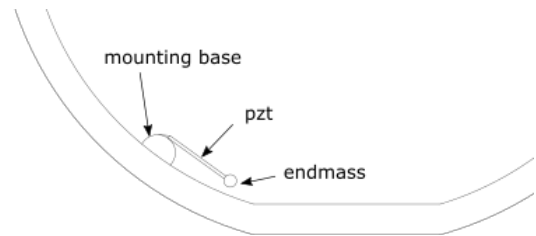


Fig. 2. Piezoelectric energy harvester inside a tire

## 2.3 3 point harvester

In a 3 point energy harvester a piezoelectric beam is connected into the inside of the tire via three points (a, b and c) as in Fig. 3. The displacement between these points deforms the harvester and generates energy. The center

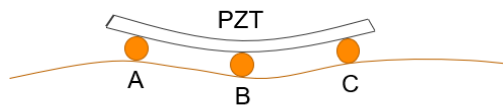
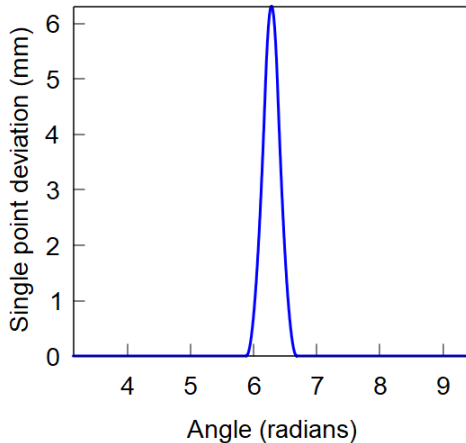
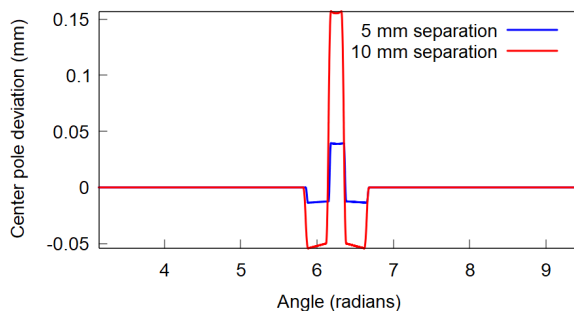


Fig. 3. Harvester connected by 3 points into the tire



(a) single point deviation



(b) center pole relative deviation

Fig. 4. 3 point harvester pole deviations

pole deviation is shown in Fig. 4(b), where the displacement of the point B (Fig. 3) in relation to A and B is shown. A single point deviation (A, B or C) is shown in Fig. 4(a). As can be seen the center point deviation, Fig. 4(b), can be a lot smaller than the single point deviation and furthermore can be tuned by moving the points A and C further from point B. This enables the use of a "stretching" type of harvester, where the material is under stress and strain. It has to be noted, that the center point deviation is similar to longitudinal strain measured in Kim et al. (2012). The 3 point harvester acts as an indirect strain measurement on the inner surface of the tire. In Kim et al. (2015), a strain model for a tire deformation is developed utilizing strain measurements and neural network. The resulting model in Kim et al. (2015) produces similar shaped strain curves as the model presented in this work. Where the models differ most are the transition zones between the contact patch and the circular section of the tire. This is due to the simple "smooth" transition between the regions that was introduced to keep the model as simple as possible.

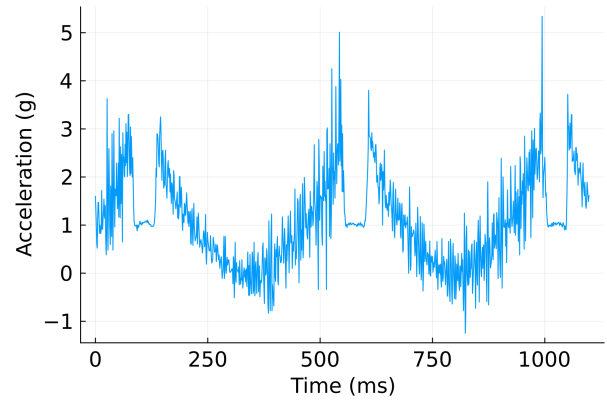


Fig. 5. Acceleration inside the tire

### 3. ACCELERATION MEASUREMENTS

As an experiment, the acceleration of the tire was measured in situ i.e. inside the tire. An accelerometer (kx-134, sparkfun electronics, USA) was used as the sensor. A nut (M6) was glued inside the tire onto which the sensor was mounted (sensor was glued to a bolt). Adafruit Feather nRF52 Bluefruit LE (Adafruit industries, USA) was used as the data logging microcontroller. The I2C bus was used as a datalink between the microcontroller and the sensor. The power and signal wires were routed through the rim and the microcontroller was placed outside the tire.

A mobile phone was used as an interface for the system. A software called "pfodApp V3" (Forward Computing and Control Pty. Ltd. NSW Australia) was used as the development platform. The microcontroller was used as a server, which served a micropage as a user interface. A mobile phone running the pfodApp was used as a client for the server. A bluetooth LE link was used as the radio link between the server and client. With this system it was possible to start the logging of the acceleration data, plot the data and save the data to the phone for post processing.

The sampling rate of the system was 1 kHz and 1000 samples were recorded, which resulted in one second worth of data sampled. The resolution of the samples were 16 bits with  $\pm 8$  g dynamic range.

An example acceleration measurement result is shown in Fig. 5. As can be seen, when the sensor is at the "flat" spot of the tire the acceleration is a constant 1 g i.e. the gravity. This is in agreement with the model. Furthermore, it can be seen that the tire exhibits a lot of high frequency vibration. The transition from a circular motion into the linear motion in the contact area is very rapid and rather linear as can be seen in Fig. 6. This would indicate that the model presented in this work can also be used when studying inertial type energy harvesters. Furthermore, the simpler acceleration model presented in equation 8 can be more suitable in some cases where the displacement input is not needed.

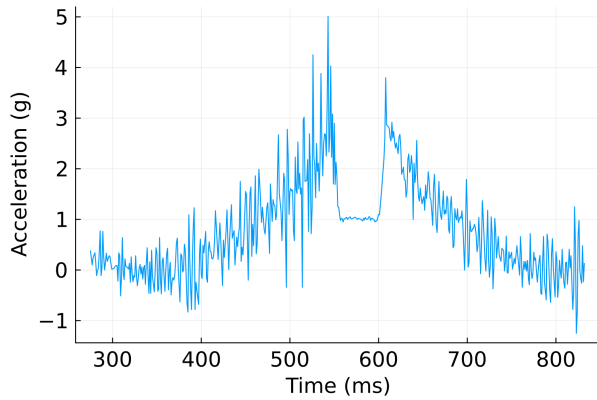


Fig. 6. Acceleration detail

#### 4. SIMULATIONS

##### 4.1 3 point harvester simulations

The simulation was carried out using Comsol Multiphysics 5.5 software. The simulated structure was a cymbal type energy harvester shown in Fig. 7. A similar structure was simulated in Leinonen et al. (2013). The width and length of the harvester was 20 mm. The piezoelectric material was PZT-5H and the endcap material was steel. The thickness of the PZT was 300  $\mu\text{m}$  and the thickness of the steel endcap was 100  $\mu\text{m}$ . At the center of the endcaps were rectangular blocks which acted as the mounting points. The bottom block was fixed and the top one was subject to the prescribed displacement. The top block was assigned to soft material (Youngs modulus of  $10\text{E}4$  Pa) to provide a "cushioning" for the harvester. The bottom block material was steel. The endcap height was 400  $\mu\text{m}$ . The FEM model was a 2D model with solid quadratic elements using plane strain approximation.

The input signal for the simulation was obtained from the Fig. 4(b) by scaling it to go from  $-20$   $\mu\text{m}$  to  $80$   $\mu\text{m}$  i.e. a total of 100  $\mu\text{m}$  stroke. As can be seen from Fig. 4(b), the shape of the curve is preserved for greater separation of the mounting points. This provides flexibility for the harvester design, as proper stroke can be chosen by moving the mounting points. In the case of the cymbal harvester, the mounting through 3 points is not trivial but could be realized by using a suitable casing.

The bottom electrode of the PZT layer was connected to a ground potential. The top layer was connected to a terminal boundary condition, which in turn was coupled to a SPICE circuit consisting of a resistor.

A transient simulation was used to simulate the time response of the system. At first a 1 Hz rotation speed was used as the input and the output voltage of the harvester is shown in Fig. 8. In this simulation the output resistor was  $\infty$ . As can be seen, the voltage of the harvester has enough amplitude for even a simple rectifier, and therefore can be used as an energy harvester.

The load resistance was increased to 1 k $\Omega$ , and the resulting voltage is seen in Fig. 9. This figure is a closeup of the first half of the circle ( $0 \leq \theta \leq \pi$ ). As can be seen, the voltage quickly dissipates due to the load resistance.

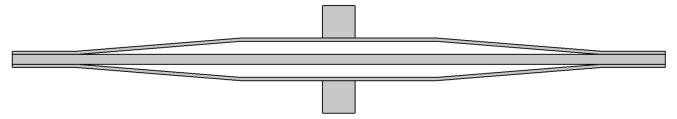


Fig. 7. The simulated cymbal.

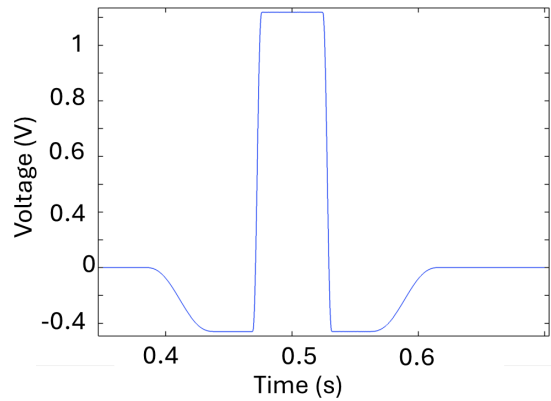
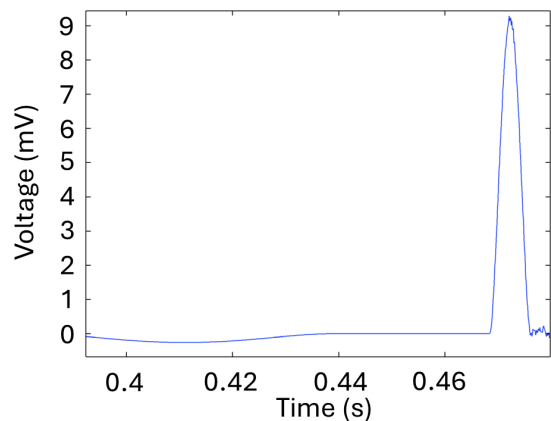


Fig. 8. Voltage of the cymbal harvester at 1 Hz rotation.

Fig. 9. The output voltage of the cymbal with a 1 k $\Omega$ 

The energy generated can be calculated by integrating the power of the signal for one period. The energy for the one cycle was 0.54 nJ at 1 Hz rotation frequency. This corresponds to 2 m/s speed. When extrapolated to 4 m/s the power is 1 nW and at 28 m/s (100 km/h) it is 7 nW. However, when simulated at 10 Hz rotation which corresponds to 20 m/s speed, 70 nW of power is generated. This is over ten times larger than the extrapolated value. This is due to resonances induced by the higher frequency. This resonating movement can be seen in Fig. 10. The whole structure together with the PZT layer is vibrating. This enhances the power generation of the harvester even though it is not strictly speaking a vibration energy harvester.

The maximum von mises stress in the PZT layer was 1.1 MPa at the 10 Hz simulation, which is well below the tensile strength of the material (114 MPa) Anton et al. (2012). This indicates that the cymbal could be stressed further to provide better energy output. The 2D model was extruded to 3D in order to visualise the geometry better, this is shown with the stresses in Fig. 11.



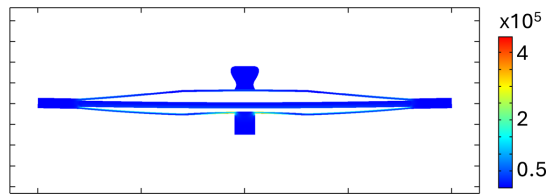


Fig. 10. The shape of the harvester under a resonance, showing the von Mises stress

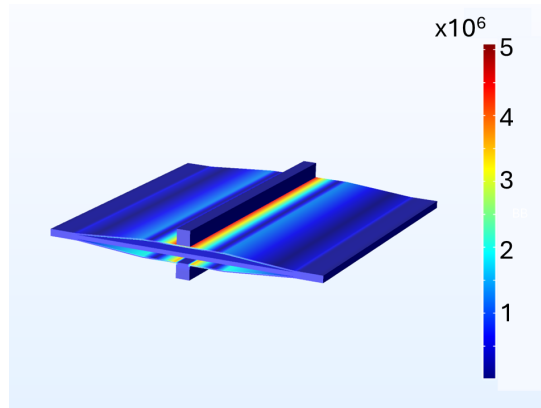


Fig. 11. The extruded 3D model with von mises stresses

#### 4.2 Inertial harvester simulations

The simulations were carried out using Comsol Multiphysics 5.5 software. The simulated geometry was a simple piezoelectric unimorph with an endmass. The length of the beam was 25 mm and the width was 10 mm. The thickness of the PZT layer (PZT-5H) was 250  $\mu\text{m}$  and the passive steel layer was 100  $\mu\text{m}$  thick. The endmass was 0.31 g. The base of the harvester was fixed and the gravity i.e. domain force was the input for the system. The gravity was calculated as a sine wave with a 1 g amplitude and a constant virtual gravity component. The sinusoidal AC component for the force represents the rotating gravity vector and the constant virtual gravity the centrifugal force. To model the contact patch area, a constant 1 g area is introduced to the force. The resulting acceleration equation is

$$a(\theta) = \begin{cases} 0 \leq \theta \leq \theta_1, G \\ \theta_1 < \theta \leq \pi, G \sin(\theta) + G_v \end{cases} \quad (8)$$

,where  $G$  is the gravitation  $9.81 \frac{m}{s^2}$  and  $G_v$  is the virtual gravity. This is a simplified model, since there is no rise time for the acceleration change from the round area to the contact patch area. However, as was seen in Fig. 6, the change in acceleration is very sharp.

For the first simulation, the virtual acceleration was set to 1 G. The rotating frequency was 2 Hz. The resulting endmass displacement is shown in Fig. 12. The power output of the harvester was 17 pW. When the virtual gravity was increased to 5 G, the power output increased to 3.4 nW.

For the 100 G input, the harvester produced 0.7  $\mu\text{W}$  of power. The maximum stress at the PZT layer was 31 MPa, which is well below the tensile strength of the material. The rotation frequency stayed the same for these simulations and only the virtual gravity was

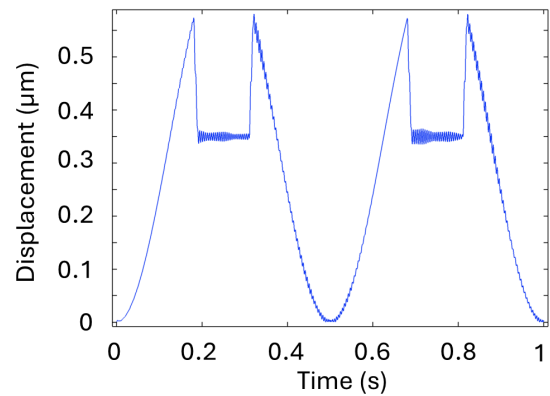


Fig. 12. The displacement of the endmass at 1 G

increased. This equates to using a larger circumference tire rather than rotating the tires faster. The increase in tire rotation speed should also improve the power output further similarly to the 3 point harvester's case.

## 5. CONCLUSIONS

In this work, a minimal model for a rotating tire was developed. The model was used as a tool to create input signals for FEM models.

A 3 point energy harvester was studied, in which the deformation of the tire is translated into a controlled relative motion between the 3 points. A cymbal type energy harvester was studied as a possible candidate for a tire mounted energy harvester. The externally leveraged structure of the cymbal proved advantageous, since the structure can exhibit vibrations even in a displacement driven case. This enhances the power generation, which without the vibrations were considerably lower.

An inertial type harvester was also studied with different centrifugal loading. As the model, and the acceleration measurements showed, the energy harvester is subject to alternating states of either very high virtual gravity or the 1 G of natural gravity. This provides a strong inertial input for the energy harvester. It is vital to design the energy harvester to withstand the high acceleration forces present in the system. For this task, the simple tire model and the FEM modeling combined, provide a good tool for harvester design.

## REFERENCES

- Al-Najati, I.A., Jasim, A.F., Chan, K.W., and Pung, S.Y. (2024). The future of tire energy: a novel one-end cap structure for sustainable energy harvesting. *Materials for Renewable and Sustainable Energy*, 13(2), 181–208.
- Aliniagerdroudbari, H., Esmaeeli, R., Hashemi, S.R., Alhadri, M., Zakri, W., and Farhad, S. (2019). A piezoelectric sandwich structure for harvesting energy from tire strain to power up intelligent tire sensors. In *2019 IEEE Power and Energy Conference at Illinois (PECI)*, 1–7. doi:10.1109/PECI.2019.8698908.
- Anton, S.R., Erturk, A., and Inman, D.J. (2012). Bending strength of piezoelectric ceramics and single crystals for multifunctional load-bearing appli-

- cations. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59(6), 1085–1092. doi:10.1109/TUFFC.2012.2299.
- Esmaeeli, R., Aliniagerdroudbari, H., Hashemi, S.R., Alhadri, M., Zakri, W., Batur, C., and Farhad, S. (2019). Design, modeling, and analysis of a high performance piezoelectric energy harvester for intelligent tires. *International Journal of Energy Research*, 43(10), 5199–5212. doi:10.1002/er.4441.
- Kim, S., Lee, J., Oh, J., and Choi, B. (2012). A self-powering system based on tire deformation during driving. *International Journal of Automotive Technology*, 13, 963–969.
- Kim, S.J., Kim, K.S., and Yoon, Y.S. (2015). Development of a tire model based on an analysis of tire strain obtained by an intelligent tire system. *International Journal of Automotive Technology*, 16(5), 865–875.
- Kubba, A.E. and Jiang, K. (2014). Efficiency enhancement of a cantilever-based vibration energy harvester. *Sensors*, 14(1), 188–211. doi:10.3390/s140100188.
- Lee, J. and Choi, B. (2014). Development of a piezoelectric energy harvesting system for implementing wireless sensors on the tires. *Energy Conversion and Management*, 78, 32–38. doi:10.1016/j.enconman.2013.09.054.
- Leinonen, M., Palosaari, J., Juuti, J., and Jantunen, H. (2013). Combined electrical and electromechanical simulations of a piezoelectric cymbal harvester for energy harvesting from walking. *Journal of Intelligent Material Systems and Structures*, 25, 391–400. doi:10.1177/1045389X13500573.
- Leinonen, M., Palosaari, J., Juuti, J., and Jantunen, H. (2017). Axle mounted piezoelectric energy harvester for continuous energy harvesting from rotation and vibration. In *Proceedings of the First World Congress on Condition Monitoring*, 1064–1074.
- Maurya, D., Kumar, P., Khaleghian, S., Sriramadas, R., Kang, M.G., Kishore, R.A., Kumar, V., Song, H.C., Park, J.M.J., Taheri, S., and Priya, S. (2018). Energy harvesting and strain sensing in smart tire for next generation autonomous vehicles. *Applied Energy*, 232, 312–322. doi:10.1016/j.apenergy.2018.09.183.
- Staaf, H., Matsson, S., Sepheri, S., Köhler, E., Daoud, K., Ahrentorp, F., Jonasson, C., Folkow, P., Ryyänen, L., Penttilä, M., and Rusu, C. (2024). Simulated and measured piezoelectric energy harvesting of dynamic load in tires. *Heliyon*, 10(7), e29043. doi:10.1016/j.heliyon.2024.e29043.
- Yi, Z., Yang, B., Zhang, W., Wu, Y., and Liu, J. (2021). Batteryless tire pressure real-time monitoring system driven by an ultralow frequency piezoelectric rotational energy harvester. *IEEE Transactions on Industrial Electronics*, 68(4), 3192–3201. doi:10.1109/TIE.2020.2978727.

# Interoperability Challenges and Opportunities in Vehicle-in-the-loop Testings: Insights from NUVE Lab's Hybrid Setup <sup>\*</sup>

Sarthak Acharya <sup>\*</sup> Aparajita Tripathy <sup>\*\*</sup> Juho Alatalo <sup>\*\*\*</sup>  
 Pekka Seppänen <sup>\*\*</sup> Aki Lamponen <sup>\*\*</sup> Jukka Säkkinen <sup>\*\*</sup>  
 Tero Päivärinta <sup>\*</sup>

<sup>\*</sup> *M3S Research Unit, ITEE, University of Oulu, Finland (e-mail: firstname.lastname@oulu.fi).*

<sup>\*\*</sup> *Oulu University of Applied Science (OAMK), Oulu, Finland (e-mail: firstname.lastname@oamk.fi)*

<sup>\*\*\*</sup> *IMS Research Unit, FTech, University of Oulu, Finland, (e-mail: firstname.lastname@oulu.fi)*

---

**Abstract:** Research and innovation in Vehicle-in-the-loop (VIL) testing is garnering more attention than ever. Integrating cyber-physical systems (CPS) into the VIL setups further enhances their functionality and hybridises the testing. Setting up any VIL infrastructure involves substantial investments and thus requires critical analysis of the resources to achieve the intended results. This study focuses on such a VIL testing infrastructure development at NUVE-Lab, aiming to provide state-of-the-art facilities for hybrid automotive testing. The facility includes physical components such as a heavy tractor (Valtra), dynamometers, an Actuators power need generation system (APGS) system, and battery emulators (BE), complemented by digital twins (DTs) of each physical machine, process, and environment to automate the testing facilities. This research examines various interoperability challenges within the current VIL framework. Three distinct testing scenarios are created to assess the overall functionalities of the hybrid setup: dynamometer-in-the-loop, APGS-in-the-loop, and BE-in-the-loop. Analyzing individual cases highlighted the need for different modeling and simulation (M/S) tools to develop digital twins. Among the tools, SIMULINK is used to build and refine the models of DTs, whereas MATLAB is used to develop control algorithms. The study also explores the adoption of Functional Mock-up Interface (FMI) standards to facilitate seamless interoperability among modeling and simulation tools. Additionally, the potential integration of the Eclipse Arrowhead framework (EAF), an IoT-edge-based automation tool, is discussed to enhance efficient data management, service interoperability, and the integration of various cyber-physical system components. In conclusion, this paper outlines the interconnection of the digital and physical platforms to evolve a hybrid VIL test laboratory, envisioning the future trajectory of the NUVE-Lab.

*Keywords:* Digital Twin (DT), Interoperability, Vehicle-in-the-loop (VIL), Eclipse Arrowhead Framework (EAF), MATLAB, SIMULINK.

---

## 1. INTRODUCTION

Vehicle-in-the-loop (VIL) is an automotive testing method that combines real machines with virtual simulations rather than conducting purely virtual tests. Advanced levels of testing in the design phase with reasonable costs and enhanced safety features have accelerated special research attention to this field in recent years Cheng et al. (2024). The focus of the traditional VIL setups was to incorporate different hardware and software in the loop for testing facilities. However, the inclusion of industry 4.0/5.0 complaint technologies has introduced hybrid VIL test se-

tups. State-of-the-art testings and validations with hybrid setups are made possible by integrating cyber systems with physical processes, digital twins (DTs), communication (5G/6G), internet of things (IoTs), augmented reality (AR) and Virtual reality (VR) technologies, artificial intelligence (AI), machine learning (ML) algorithms, cloud and edge computing, and others Zhao et al. (2023). However, the addition of diverse technologies increases complexity and interoperability concerns for the interconnected systems Lv et al. (2024). One such cyber-physical-system-based hybrid VIL research infrastructure at the NUVE lab in Oulu, Finland, is discussed in this paper.

Cyber-Physical Systems (CPS) combine computational and engineering fields to develop systems that connect the

---

<sup>\*</sup> The financial support received from the European Union NextGenerationEU instrument, which is funded by the Research Council of Finland under grant number 352726.

physical and digital realms. These systems incorporate intricate networks, control systems, information fusion, and optimization techniques Su et al. (2022). These systems have the potential to revolutionize various domains by enabling advanced functionalities and automation. However, the implementation of CPS is accompanied by several challenges. One significant challenge is the diverse nature of technologies and standards, leading to interoperability issues that hinder seamless communication and integration among different subsystems within CPS Chaudhry et al. (2019). Among others, digital twin technologies (DTTs) have been observed as a core component of CPS for creating the virtual representation of physical objects or systems. The diverse nature of digital twin implementations, with various data structures and interfaces, hinders the interoperability between different digital twins Wang et al. (2023). Additionally, enabling real-time dynamic interactions between the simulation world and the physical world through digital twins introduces complexities in ensuring consistent and reliable data exchange, posing a challenge to achieving seamless interoperability. Therefore, IoT-based automation tools are effective for interoperability in such hybrid setups. Eclipse Arrowhead framework (EAF) is one such potential tool for interoperability solutions and is considered for NUVE's VIL setup.

The creation and optimization of digital twins (DTs) in VIL setups are dependent on various modeling and simulation (M/S) tools. The test-bench-based hybrid testing methods require the simulation of the environment in order to provide realistic inputs for the device under test. These simulation inputs need to be generated in real-time, which narrows down the choice of simulation tools and excludes all the software that needs heavy computing (such as FEM tools). To keep the number of simulation software reasonable, we are going to use Mevea and MATLAB/SIMULINK as the main tools in this research project. Mevea is a software for real-time simulation of mechanics, hydraulics, power transmission, and the operating environment (Mevea (2024)). MATLAB/SIMULINK is a software for modeling and simulating dynamic systems in multiple domains (Mathworks (2024b)).

In this paper, we aim to study the hybrid VIL setup in NUVE-Lab at the Oulu Applied Science University(OAMK), Oulu, Finland. The laboratory setup consists of a Valtra tractor, 4 dynamometers, an APGS system, and a battery emulator as physical components. On the counter, the digital models of the tractor are designed in the Mevea environment for simulation and testing Oulu University of Applied Science (OAMK) (2023). A few digital counterparts are modeled using SIMULINK as well. One of the primary objectives is to study the interoperability challenges across various test scenarios. Next, to analyze the usability of existing Internet of Things (IoT) middleware and Open-source platforms as potential candidates to overcome the hurdles. Thus, two important research questions (RQs) are formulated:

**RQ-1:** *What are the essential components to be considered in setting up a hybrid VIL setup?*

**RQ-2:** *What are the key interoperability challenges in a hybrid VIL environment?*

The outline of the paper is as follows: background and key enablers for the hybrid VIL are mentioned in section 2. The research process and context followed are presented in section 3. Section 4 illustrates the key findings of the research work. Then, the results are discussed in section 5. Lastly, the conclusion and future scopes of the work are presented in section 6.

## 2. BACKGROUND

In this section, a summary of the recent works related to the vehicle-in-the-loop (VIL) testing is presented.

### 2.1 VIL Testbench Setup: SOTA and Concepts

The vehicle-in-the-loop technique is the combination of real-world vehicles with virtual environments to perform experiments in a cost-efficient way Park et al. (2020). The experimental setups for the VIL testing vary with infrastructures based on specific research objectives, testing precision, accuracy of experiments, available technologies, and resources. However, there is a core set of basic requirements to establish a functional VIL testing environment, including physical vehicles (hardware), simulation software, data acquisition systems, computing infrastructure, and safety measures Cheng et al. (2024). To validate the results, there are 3 basic testing approaches: simulation, hybrid (closed-field), and real (on-road) testing Solmaz et al. (2020). Simulation tests are the most economical and low in accuracy as they only involve digital twins and software. On the other hand, real tests include the physical vehicles in real-world scenarios, which are highly expensive but most accurate. Therefore, hybrid tests are often considered as optimal solution as they take both hardware and software in the loop to perform experiments.

### 2.2 Challenges in VIL setup

VIL testing setups have proven their worth for automotive research and development, but they face several challenges. One of the complexities in setting up a VIL laboratory involves a huge investment in infrastructure. Integration of various hardware and software platforms, maintenance and updates of the technologies, and scalability are among the top challenges. However, in this study, we only investigate and focus on the technical challenges associated with the implementation of VIL. The list of challenges is as follows:

- System Integration and Compatibility
- Connectivity and V2X Communication
- Real-time processing
- Data Acquisition and Handling
- Scalability and Flexibility
- High-Fidelity Sensor Simulations
- Mixed Reality and Enhanced Immersion
- Standardization and Bench-marking

Integration of advanced technologies such as radar target simulation and environment perception simulation in VIL testing poses technical challenges Maier et al. (2018). The integration phase is particularly complex due to the numerous components in modern vehicles Rossi et al. (2017). Additionally, the growing complexity of automotive cyber-physical systems and the verification challenges posed by

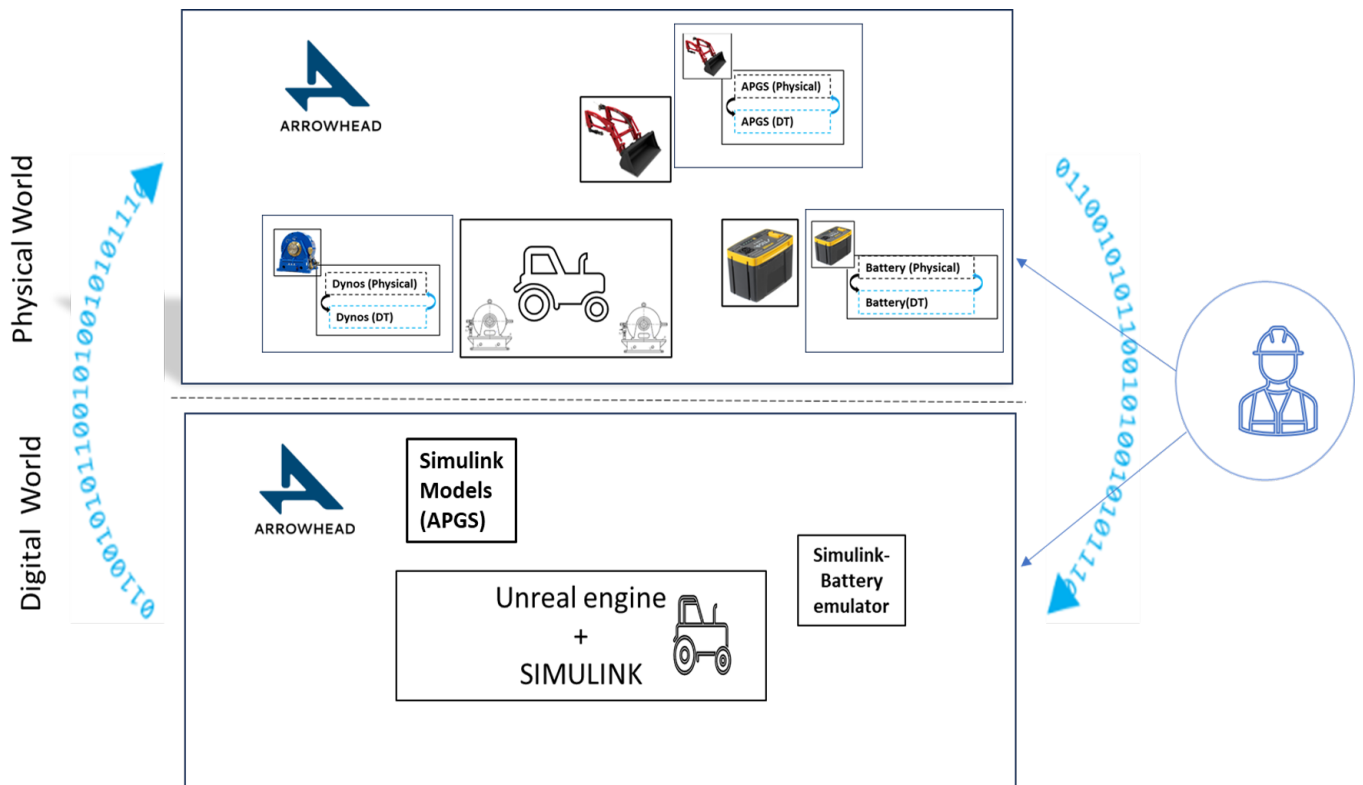


Fig. 1. Overview of the Vehicle-in-the-loop Testing in NUVE-Lab.

distributed software in vehicles create difficulties Raghupatruni et al. (2019). There is also a pressing need for simulations that accurately reflect the complexity of real-world testing environments Babić et al. (2020). Moreover, thorough testing for autonomous vehicles remains a major challenge Chen et al. (2020).

### 2.3 Role of Simulation and Modelling Platforms

Modeling, in general, requires a lot of modeling expertise and is labor intensive, although in this project, we are modifying the existing Valtra tractor model introduced by Jaiswal et al. (2019). The real-time demand of the simulation models also restricts the choices made in modeling and reduces the details that can be applied to the model.

In this research, a setup will be made where the hydraulic system of the tractor is modeled in Simscape, and the front loader is modeled in Mevea. The simulation is performed as a co-simulation where the Simscape model is exported as a functional mock-up unit (FMU) to the Mevea environment (FMI (2024)). The use of the FMU in co-simulation requires a lot of adaptation of the models and can emerge with various issues due to the varying modeling principles in different software.

### 2.4 Role of CPS, DTs and IoT Platforms

IoT platforms and frameworks play a significant role in enabling Vehicle-in-the-Loop (VIL) testing by offering the necessary infrastructure for connecting and monitoring all the components used in the testing procedure. These platforms allow the seamless integration of IoT devices, sensors, and communication technologies within the VIL

setup, enabling the exchange of real-time data and control functions. There are numerous IoT frameworks and platforms currently being used by industries, including the Eclipse Arrowhead Framework, AUTOSAR, BaSys, FIWARE, OCF, IoTivity, and more (Paniagua and Delsing (2020)). The Eclipse Arrowhead framework has been selected for the GORI project to develop networked connectivity between different VIL setups using local cloud automation.

### 2.5 Industry Complaint Open-source IoT Framework

The Eclipse Arrowhead Framework (EAF) is an open-source industrial IoT framework that provides interoperability solutions in Industry 4.0. Delsing (2017). This framework is built on the principles of SOA (Service Oriented Architecture) and leverages the concept of local clouds. Here, a local cloud is a network of interconnected systems and services that function within a limited environment, usually within a particular organization, stakeholder, or region. This architecture emphasizes Standardized communication, late binding, loose coupling, cyber-security, scalability, Dynamic service Discovery, and multi-stakeholder integration. The framework enables real-time communication between different systems irrespective of the technology being used within a local cloud or between systems registered in different local clouds. The framework is compatible with several communication protocols, including HTTP, COAP, MQTT, and OPC UA, as well as transport protocols such as TCP, UDP, and DTLS/TLS, Acharya et al. (2023). Three mandatory core systems are provided to facilitate interaction between systems. The mandatory core systems are:



- (1) the *Service Registry* system, which records the services currently being offered,
- (2) the *Authorization* system that controls system-to-system authorization at a detailed level for secure service exchange,
- (3) the *Orchestrator* system that enables the consumer application to discover the required service endpoint at run time.

In addition to the core systems, there are client systems that essentially function as application systems, either as providers or consumers, that seek to establish communication with one another. Each core system and provider system offers a set of services that are registered with the *Service-Registry* and includes a specific set of interfaces, metadata, and service paths. Any system that wishes to utilize a service must request the service address from the *Orchestrator* system during runtime. The Orchestrator verifies with the Service Registry if the service is now accessible and then consults the Authorization system to determine if the exchange of the specific service is approved between the two systems. After the verification process is completed successfully, the *Orchestrator* provides the relevant service endpoint in response. The consumer system directly contacts the service-provider system. The framework also employs the Gatekeeper and Gateway core systems to provide inter-cloud communication.

2.6 NUVE-Lab Vision

In NUVE-Lab, the ongoing research is focused on building a hybrid VIL testing platform that can accommodate cutting-edge automotive research and development. The current laboratory setup is shown in Fig. 1. Further optimization of digital twins by accurately replicating the physical entities will enhance applications such as predictive maintenance, real-time monitoring, fault diagnosis, etc. Different test scenarios will be developed further by integrating sensors, actuators, programmable logic controllers (PLCs), and Lidars. The longer-term goal is to enable testing across a wide range of vehicles, both with physical and digital prototypes. This will leverage facilities to test all types of prototypes before the realization of products.

3. RESEARCH PROCESS AND CONTEXT

This section explains the research process followed to analyze the integration of physical and software components of the current VIL setup at the NUVE lab. To carry out a systematical investigation of the interconnectivity and interoperability concerns, three test scenarios are developed (as shown in Fig. 2). In the first scenario (T1), dynamometer is attached to the vehicle to perform tests, testing setup is shown in Fig. 2. The second case (T2) is for testing hydraulic systems by attaching an actuator’s Power Need Generation System (APGS) to the vehicle. In the third setup (T3), a battery emulator is connected to the vehicle. All the test scenarios are explained in section 4.

3.1 Challenge Identification

To investigate the interoperability challenges, all the connections are labeled (A, B, C, ..., J) in each test scenario.

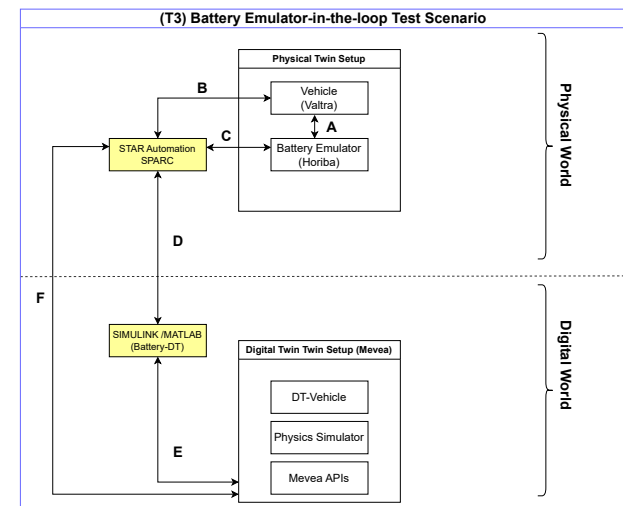
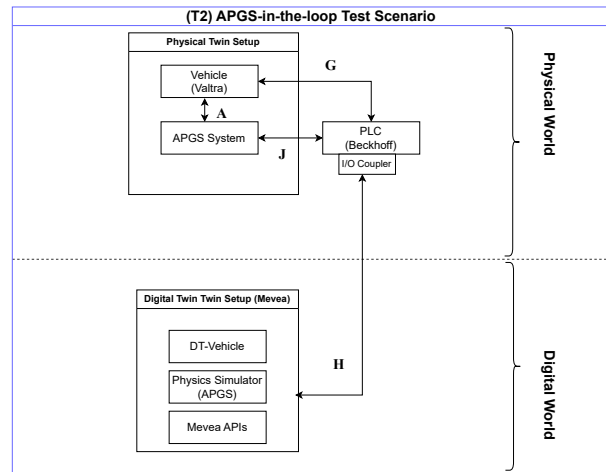
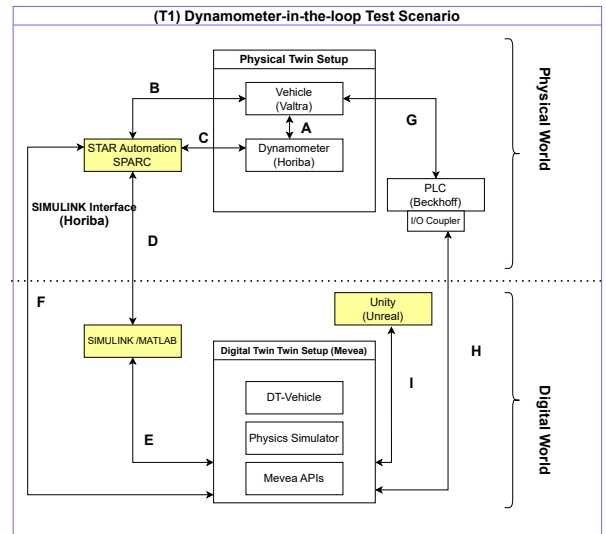


Fig. 2. Test Scenarios in NUVE’s VIL Setup: (T1) Dynamometer Testing with the Vehicle; (T2) Hydraulics System testing with the Vehicle; (T3) Battery Emulator testing with the Vehicle.

Table 1. Challenges Identified for all the 3 Test Scenarios

Connection	Type	Identified Challenge(s)		
		Dynos-in-the-loop (T1)	APGS-in-the-loop (T2)	BE-in-the-loop (T3)
A	P-P	- Mechanical setup and configurations - Version Maintenance and Upgradations - Software Integration	- Electrical setups - Control system integration - Version Maintenance	- Electrical Compatibility - Interface Management
B	P-D	- Protocol Mismatch - API Compatibility - Maintenance and Upgradations	Not Available	- Protocol Mismatch - API Compatibility - Maintenance and Upgradations
C	P-D	- Communication Protocols - Latency	Not Available	- Protocol consistency - Control system-software integration - Latency
D	D-D	- Data Exchange - Protocol Mismatch - Version compatibility - Latency (better than connection 'F')	Not Available	- Data Exchange - Protocol Mismatch - Version compatibility - Latency
E	D-D	- API compatibility - Model Accuracy - Version compatibility - Data Format and representation	Not Available	- Creation of DTs - API compatibility - Model Accuracy - Version compatibility - Data Format and representation
F	D-D	- Protocol Compatibility - API compatibility - Latency (UDP loop) - Model Accuracy	Not Available	- Protocol Compatibility - API compatibility - Data format Compatibility - Latency - Model Accuracy
G	P-P	- Protocol Mismatches (CAN to Mod-bus/ Ethernet/IP) - Latency - I/O Compatibility - Integration with other systems	- Protocol Mismatches (CAN to Mod-bus/ Ethernet/IP) - Latency - I/O Compatibility - Integration with other systems	Not Available
H	P-D	- Data Exchange and Latency - Digital Twin Accuracy - Scalability	- Creation of Digital Twins (DTs) - Data communication - Integration with other systems - Refinement of DTs	Not Available
I	D-D	- Model Accuracy - Data Format Compatibility - Integration of Physics Simulators	Not Available	Not Available
J	P-P	Not Available	- Protocol consistency - Control system-software integration - Latency	Not Available

Three main categories of the connections are shown in Figure 2: physical to physical (P-P), physical to digital (P-D), and digital to digital (D-D). Identified challenges are mentioned in section 4.

## 4. RESULTS

### 4.1 Dynamometer-in-the-Loop Testing (T1)

In this test setup, the dynamometers are tested with the vehicle, as shown in Fig. 2. All the connections and identified challenges are mentioned in table 1. Dynamometers connected with the vehicle (Valtra tractor) are controlled by STARS automation software and exchange data with the digital twin setup. Digital twins are created in the Mevea tool and updated in real-time using SIMULINK models. Some of the significant interoperability challenges faced in this VIL testing are protocol compatibility, latency, API compatibility, and software integration.

### 4.2 APGS-in-the-loop Testing (T2)

In the second test setup, an actuator power need generation (APGS) system was tested in the VIL. The test setup is shown in Fig. 2, and the identified challenges are mentioned in table 1. This test setup is optimized in terms of the number of connections (compared to T1). One of the crucial parts is to create digital twins with granular

information from their physical counterparts. Other challenges include protocol mismatching, input-output (I/O) compatibility, latency and etc.

### 4.3 Battery Emulator-in-the-Loop Testing (T3)

The third setup is to test the battery in the hybrid VIL setup. The experimental setup is presented in Fig. 2. Most of the connections in this setup have familiarity with the first test-case (T1). The identified challenges are listed in table 1. Some essential interoperability barriers in this testing are achieving real-time data and feeding into the twin of the battery. Software integration with the control systems is very crucial for this test setup and can possibly need middleware solutions in the future.

## 5. DISCUSSIONS

### 5.1 Dynamometer-in-the-Loop Testing (T1)

In the dynamometer-in-the-loop test case, the Valtra tractor mounted on dynamometers interacts with Mevea software, where the multi-body dynamics model of the tractor and the virtual environment are modeled (as shown in Fig. 2). The purpose is to evaluate the performance and the behavior of the tractor under realistic driving conditions without the need for a physical test track. For example, in one case, the Mevea software (DT setup) receives the wheel

Table 2. Analysis of each connection and Possible Arrowhead Solutions

Connection	Type (Test Case)	EAF Mapping	Feasibility
A, J	P-P (T1, T2, T3)	Creating individual Arrowhead application systems for vehicles and dynamometers that can connect to sensors and actuators of the hardware. These systems can then provide services that collect, store, and manage these sensors and actuator's data.	The sensors and actuators' information from the hardware can be easily extracted at the next connection levels. Hence, using Arrowhead at this point is not recommended.
B	P-D	Creating individual Arrowhead application systems for the vehicle and STAR automation that can exchange information with each other and control the actuators in the vehicle through Arrowhead services.	Using Arrowhead at this point is not recommended to avoid latency.
C	P-D	Creating individual Arrowhead application systems for the dynamometers and STAR automation that can exchange information with each other and control the actuators in the vehicle through Arrowhead services.	This will create unnecessary latency. Hence, using Arrowhead at this point is not recommended.
D	D-D	Creating individual Arrowhead application systems for STAR Automation and SIMULINK/MATLAB, where information exchange between the two systems takes place via the Arrowhead service exchange process.	Arrowhead can be useful if it can avoid additional latency.
E	D-D	Creating individual Arrowhead application systems for SIMULINK/MATLAB and Mevea. The SIMULINK system can send the input(torque) values to the Mevea through the Arrowhead service exchange process. After testing the DT against the input value at Mevea, the Mevea system can send the outputs (speed of the tires) to the SIMULINK system.	Arrowhead can be used in this scenario as it can be useful to store the test data in a standardized way (senML) into the Data Manager (DM), provided it avoids latency.
F	D-D	Creating individual Arrowhead application systems for STAR automation and Mevea that can exchange information like torque and speed securely via the Arrowhead service exchange process.	Arrowhead can be used to establish direct communication between the two entities if there are no SIMULINK models in the loop for testing.
G	P-P	Creating individual Arrowhead application system for the vehicle and PLC that can connect to sensors and actuators of the Vehicle to the PLC I/O ports. These systems can then provide services that collect, store, and manage these sensors and actuators' data.	The sensor and actuator information from the vehicle can be easily extracted directly from the PLC. Hence, using Arrowhead at this point is not recommended.
H	P-D	Creating individual Arrowhead application systems for the PLC and Mevea. The PLC system can provide services to read the input and output signals from the PLC and send that information to the Mevea system for testing. The Mevea system can send the output values back to the PLC via the Arrowhead service exchange process.	Arrowhead can be utilized here in order to exchange data in a secure way and also to store the data in the Data Manager (DM). Changes to the hardware (like sensors and actuators) can be handled at run time through the automatic discovery of new services, resulting in reduced engineering effort.
I	D-D	Creating individual Arrowhead application systems for Unity and Mevea and exchange information between the two systems using Arrowhead service exchange process and TCP protocol.	Arrowhead can be utilized. Latency needs to be tested. (Best suited when Mevea and Unity operate in different networks or locations.)

speed data from the dynamometers and uses it to calculate the motion of the tractor's simulation model in the virtual

environment. The Mevea software also calculates the driving resistance that the tractor model encounters in the

virtual environment based on the position and velocity of the simulation model and the characteristics of the environment. Therefore, interoperability of data between the physical and digital parts is a must.

The communication between Mevea software and the Horiba system (STARS automation) has to be routed through Matlab/SIMULINK because, at the moment, Horiba and Mevea have no ability to communicate directly with each other. This might cause extra latency, but the current setup appears to be a working solution based on the preliminary communication tests done in NUVE-Lab.

### 5.2 APGS-in-the-loop Testing (T2)

In the APGS-in-the-loop testing scene, the hydraulic system of the Valtra tractor is connected to the APGS system, which can control the hydraulic load of the main hydraulic line and the pressure of the load-sensing line. The performance of the tractor's hydraulic system can be evaluated without the need for any auxiliary equipment, such as a front loader. Alternatively, we can also test how the tractor reacts to the auxiliary equipment without actually mounting the equipment to the tractor. The Mevea software is used to build a multi-body dynamics model of the front loader and the model of the hydraulic system of the Valtra tractor. The modeling can also be done by using Simscape (Mathworks (2024a)), which is an extension for the SIMULINK. The simulation model of the front loader experiences different kinds of loads generated in the virtual environment, such as lifting. The front loader movement is controlled by giving control commands in the virtual environment.

The simulation model provides the pressure and flow information of the hydraulic fluid to the APGS system, which in turn generates similar pressure and flow in the real world in order to load the tractor's hydraulic system accordingly. The speed of the tractor's hydraulic pump can be measured and sent to the simulation model in order to create a feedback loop. This way, a realistic system that mimics the behavior of the tractor and the front loader in different operating conditions can be created.

The FMU adaptation requires resolving several compatibility issues, such as data types, input/output variables, initialization methods, and solver settings. The collaboration with the software supplier is working exemplary, and the raised questions have been resolved without delays, although the fully working test model is still under construction.

The APGS system communicates with Mevea through Bechhoff/EtherCAT. This has been tested, and the tentative results show that the solution works. This raises the question of whether it would be possible for Mevea to communicate with the Horiba system via Ether-CAT. This would eliminate the need for the SIMULINK between Mevea and Horiba and would possibly increase the communication speed.

### 5.3 Battery Emulator-in-the-Loop Testing (T3)

In this testing case, the battery emulator is yet to be fully integrated with the hybrid VIL setup. To ensure

seamless communication between the physical setup and digital twin setup, compatibility across protocols and data consistency is required. Another important aspect at this moment is to create an accurate digital twin that can simulate unforeseen conditions. To work on the mentioned interoperability challenges, the adaptation of IoT middleware solutions to the current VIL setup is needed. One such framework is analyzed below.

### 5.4 Logical Reflections on AH Potential to NUVE-VIL context

Utilizing the industry 4.0 IoT frameworks in VIL testing offers several benefits. The benefits encompass ensuring data quality management, efficient data processing, and enhanced data security practices Javed et al. (2019)Mishra et al. (2015). The EAF follows the principles of Industry 4.0, offering the same advantages outlined earlier. The basic steps in utilizing EAF in any connection are to create an Arrowhead application system for each end of the connection and then establish communication through the service exchange process of EAF explained earlier in section 2.5. In order to gain a deeper understanding of how EAF might be employed, Table 2 presents a comprehensive list of the connections between the various components and systems in NUVE Lab, as well as how EAF can optimize the data exchange process.

One of the primary benefits of Arrowhead is its ability to automatically discover new services during execution. This feature is particularly valuable in the continuously evolving manufacturing landscape, characterized by frequent hardware and software changes. Automatic service discovery techniques minimize the effort needed to synchronize changes in the physical world with the digital realm. For instance, changes in the physical world, like switching between sensors or PLCs, can be easily handled during execution Tripathy et al. (2022). This will significantly reduce the effort involved in re-engineering.

### 5.5 Threads to Validity

The hybrid VIL setup discussed in the paper is an ongoing research work at the NUVE lab at OAMK, Oulu.

*Internal validity* This study is based on the practical work at the laboratory. Therefore, a possible threat to the study findings includes missing the relevant data. We employed a thorough literature survey on the existing literature to analyze our findings. In *external Validity*, concerns the generalizability of the findings and diversification of technologies across different setups. To mitigate this all the available industrial reports and gray literature were refereed to validate the findings.

## 6. CONCLUSION AND FUTURE WORK

In this paper, a hybrid VIL setup of the NUVE lab is demonstrated. This study is focused on building a hybrid VIL environment to assess advanced automotive testing capabilities. Interoperability is observed as a key factor in building such setups while integrating CPS, DTs, software components, and other industry-compliant technologies. The study highlighted the requirements and challenges for

transitioning from traditional VIL to a hybrid testing platform. The physical infrastructure is integrated with cyber-physical systems to facilitate the advanced-level testing. Three test scenarios, including dynamometer-in-the-loop, APGS-in-the-loop, and Battery emulator-in-the-loop, have been designed to investigate and explore the interoperability challenges. The role of modeling and simulation (M/S) platforms in developing and refining digital twins is reviewed. In addition, the potential of the Arrowhead framework to facilitate interoperability in a hybrid VIL setup is examined. These insights from the NUVE lab also lay a foundation for creating more innovative hybrid VIL setups in the future.

This is an ongoing research project, and future work will focus on finding engineering solutions to existing interoperability concerns. The Arrowhead framework will further be tested to enhance interoperability among digital twins, data management, scalability, and collaboration. Furthermore, the creation and refinement of digital twins to achieve better accuracy in mimicking physical components and processes will be focused. The ultimate vision of this hybrid VIL laboratory is to facilitate state-of-the-art automotive testing and innovation.

#### ACKNOWLEDGEMENTS

The research work is carried out under the GO-RI project sponsored by the Academy of Finland. This is a collaborative work between the University of Oulu and Oulu Applied Science University (OAMK), Finland. The authors would like to thank NUVE-Lab and the team for setting up the testing infrastructure. The authors would like to acknowledge the financial support received from the European Union – NextGenerationEU instrument which is funded by the Research Council of Finland under grant number 352726.

#### REFERENCES

- Acharya, S., Wintercorn, O., Tripathy, A., Hanif, M., Van Deventer, J., and Päivärinta, T. (2023). Twins interoperability through service oriented architecture: A use-case of industry 4.0. In *Proceedings of the Annual Symposium of Computer Science 2023 co-located with The International Conference on Evaluation and Assessment in Software Engineering (EASE 2023)*. R. Piskac c/o Redaktion Sun SITE, Informatik V, RWTH Aachen.
- Babić, A., Vasiljević, G., and Mišković, N. (2020). Vehicle-in-the-loop framework for testing long-term autonomy in a heterogeneous marine robot swarm. *IEEE Robotics and Automation Letters*, 5(3), 4439–4446.
- Chaudhry, N., Yousaf, M.M., and Khan, M.T. (2019). Security assessment of data management systems for cyber physical system applications. *Journal of Software Evolution and Process*. doi:10.1002/smr.2241.
- Chen, Y., Chen, S., Xiao, T., Zhang, S., Hou, Q., and Zheng, N. (2020). Mixed test environment-based vehicle-in-the-loop validation—a new testing approach for autonomous vehicles. In *2020 IEEE intelligent vehicles symposium (IV)*, 1283–1289. IEEE.
- Cheng, J., Wang, Z., Zhao, X., Xu, Z., Ding, M., and Takeda, K. (2024). A survey on testbench-based vehicle-in-the-loop simulation testing for autonomous vehicles: Architecture, principle, and equipment. *Advanced Intelligent Systems*, 2300778.
- Delsing, J. (2017). *IoT Automation: Arrowhead Framework*. CRC Press.
- FMI (2024). Functional mock-up interface specification, version 3.0.1. <https://fmi-standard.org/docs/3.0.1/>. [Online; accessed 17-May-2024].
- Jaiswal, S., Korkealaakso, P., Åman, R., Sopenan, J., and Mikkola, A. (2019). Deformable terrain model for the real-time multibody simulation of a tractor with a hydraulically driven front-loader. *IEEE Access*, 7, 172694–172708. doi:10.1109/ACCESS.2019.2956164.
- Javed, A., Yousefnezhad, N., Robert, J., Heljanko, K., and Främling, K. (2019). Access time improvement framework for standardized iot gateways. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 220–226. IEEE.
- Lv, S., Qin, Y., Gan, W., Xu, Z., and Shi, L. (2024). A systematic literature review of vehicle-to-everything in communication, computation and service scenarios. *International Journal of General Systems*, 1–31.
- Maier, F.M., Makkapati, V.P., and Horn, M. (2018). Environment perception simulation for radar stimulation in automated driving function testing. *Elektrotech. Informationstechnik*, 135(4-5), 309–315.
- Mathworks (2024a). Simscape. <https://se.mathworks.com/products/simscape.html>. [Online; accessed 16-May-2024].
- Mathworks (2024b). Simulink. <https://se.mathworks.com/products/simulink.html>. [Online; accessed 16-May-2024].
- Mevea (2024). Mevea simulation software. <https://mevea.com/solutions/software/>. [Online; accessed 16-May-2024].
- Mishra, N., Lin, C.C., and Chang, H.T. (2015). A cognitive adopted framework for iot big-data management and knowledge discovery prospective. *International Journal of Distributed Sensor Networks*, 11(10), 718390.
- Oulu University of Applied Science (OAMK) (2023). Nuvelab. <https://oamk.fi/nuve-lab/>. [Online; accessed 04-April-2024].
- Paniagua, C. and Delsing, J. (2020). Industrial frameworks for internet of things: A survey. *IEEE Systems Journal*, 15(1), 1149–1159.
- Park, C., Chung, S., and Lee, H. (2020). Vehicle-in-the-loop in global coordinates for advanced driver assistance system. *Applied Sciences*, 10(8), 2645.
- Raghupatruni, I., Goeppel, T., Atak, M., Bou, J., and Huber, T. (2019). Empirical testing of automotive cyber-physical systems with credible software-in-the-loop environments. In *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*, 1–6. IEEE.
- Rossi, R., Galko, C., Narasimman, H., and Savatier, X. (2017). Vehicle hardware-in-the-loop system for adas virtual testing. *Toward a Common Software/Hardware Methodol. Futur. Adv. Driv. Assist. Syst. DESERVE Approach*, 251–267.
- Solmaz, S., Rudigier, M., and Mischinger, M. (2020). A vehicle-in-the-loop methodology for evaluating automated driving functions in virtual traffic. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, 1465–1471. IEEE.



- Su, H., Xiao, B., Zhou, M., Qi, W., Sandoval, J., and Kim, S.T. (2022). Theory, applications, and challenges of cyber-physical systems 2021. *Complexity*. doi: 10.1155/2022/9861298.
- Tripathy, A., van Deventer, J., Paniagua, C., and Delsing, J. (2022). Opc ua service discovery and binding in a service-oriented architecture. In *IEEE 5th International Conference on Industrial Cyber-Physical Systems*.
- Wang, X., Hu, X., Ren, Z., and Tian, T. (2023). Knowledge-graph based multi-domain model integration method for digital-twin workshops. doi:10.21203/rs.3.rs-2630784/v1.
- Zhao, X., Gao, Y., Jin, S., Xu, Z., Liu, Z., Fan, W., and Liu, P. (2023). Development of a cyber-physical-system perspective based simulation platform for optimizing connected automated vehicles dedicated lanes. *Expert Systems with Applications*, 213, 118972.

## New Chemical Kinetics Mechanism for Simulation of Natural Gas/Hydrogen/Diesel Multi-fuel Combustion in Engines

Mohammad Mahdi Salahi<sup>\*</sup>, Amin Mahmoudzadeh Andwari<sup>\*</sup>, Alireza Kakoe<sup>\*\*</sup>, Kian Golbaghi<sup>\*\*</sup>,  
Jari Hyvönen<sup>\*\*\*</sup>, Ayat Gharehghani<sup>\*\*\*\*</sup>, Maciej Mikulski<sup>\*\*</sup>, Eric Lendormy<sup>\*\*\*</sup>

<sup>\*</sup> Machine and Vehicle Design (MVD), Materials and Mechanical Engineering, University of Oulu, FI-90014 Oulu, Finland ([mahdi.salahi@oulu.fi](mailto:mahdi.salahi@oulu.fi), [amin.m.andwari@oulu.fi](mailto:amin.m.andwari@oulu.fi))

<sup>\*\*</sup> University of Vaasa, School of Technology and Innovation, Wolffintie 34, FI-65200, Vaasa, Finland  
<sup>\*\*\*</sup> Wärtsilä Finland OY, Jarvikatu 2-4, 65101 Vaasa, Finland

<sup>\*\*\*\*</sup> School of Mechanical Engineering, Iran University of Science and Technology, 16846, Tehran, Iran

**Abstract:** Reactivity Controlled Compression Ignition (RCCI) stands out as a promising combustion method for the next wave of internal combustion engines, offering cleaner and more efficient operation, particularly in heavy-duty engines. A key approach within this strategy involves pairing diesel as the high reactivity fuel with natural gas (NG) as the low reactivity counterpart. Further optimization can be achieved by introducing hydrogen to replace portions of NG, thereby enhancing combustion quality while reducing greenhouse gas emissions. For accurate numerical simulation of engines employing this strategy, specialized chemical kinetics reaction mechanism tailored for internal combustion engines becomes essential. To facilitate computationally efficient 3-D Computational Fluid Dynamics (CFD) simulations, the mechanism has been reduced to include 60 species and 372 reactions, with N-heptane acting as a diesel fuel surrogate. This compact mechanism is optimized to align with experimental ignition delay time (IDT) data for N-heptane. The accuracy of the mechanism's predictions for IDT and laminar burning velocity (LBV) is validated using available experimental data. Furthermore, 3-D CFD and quasi-dimensional multi-zone engine simulations are performed with the new mechanism to validate engine operating parameters against experimental data.

**Keywords:** Diesel, Natural Gas, Simulation, Combustion, Engines, Mechanism

### 1. INTRODUCTION

Low-temperature combustion (LTC) is an advanced combustion concept for internal combustion engines that has attracted global attention in recent years. Among the LTC modes, Reactivity-Controlled Compression Ignition (RCCI), in particular, has demonstrated promising results in terms of reducing NO<sub>x</sub> emissions, brake thermal efficiency, and combustion phasing control when operated under certain conditions (Salahi et al., 2017). The unique operation of RCCI engines involves the intake of a low-reactivity fuel and the direct cylinder injection of a high-reactivity fuel (Hossein Fakhari et al., 2024). This process creates reactivity stratification, providing more control over the heat release rate. Recent advancements in RCCI engines have explored the use of alternative fuels and investigated combustion strategies to extend RCCI mode to higher loads (Fakhari et al., 2023). However, there is a recognized need for further enhancements in simulation models for these engines (Dwarshala et al., 2023).

Simulating combustion in internal combustion engines is crucial due to its impact on engine performance, emissions, and environmental regulations (Gössnitzer et al., 2022). Engine simulation methods have evolved significantly, with studies highlighting the use of simulation software to predict engine performance accurately.

Reliable and practical chemical kinetics are essential for the accuracy of modeling internal combustion engines (Kakoe et al., 2022). Vasudev et al. (Vasudev, Mikulski, et al., 2022) provide a detailed examination of chemical kinetics within the context of multi-zone models for low-temperature combustion engines. The review emphasizes the difficulties in striking a balance between computational efficiency and model complexity, and it suggests that developments in methods such as tabulated chemistry will lead to significant gains in simulation speed. Aziz et al. (Aziz et al., 2022) provided a comprehensive review of the impact of various fuels on the performance and emissions of RCCI engines, highlighting the benefits of using low-reactivity fuels like gasoline, natural gas, and alcohol-based fuels, as well as high-reactivity fuels like diesel and biodiesel. They explored fuel management strategies and the use of additives to enhance Ammonia and hydrogen has also attracted a lot of attention in marine and power plant applications (Salahi et al., 2022). Natural gas, with its great fuel economy, clean combustion characteristics, abundance, and comparatively low cost, is one of the most potential alternative fuels to meet ever-tougher pollution regulations (Gharehghani et al., 2023). The analysis of dual-fuel engines reveals that researchers need a more appropriate dual-fuel chemical kinetic mechanism involving diesel and natural gas (W. Zhang et al., 2019).

Optimizing the chemical kinetic mechanism for natural gas and diesel dual-fuel engines presents challenges, including the need for accurate prediction of combustion characteristics and emissions, understanding the impact of methane content on ignition delay, and ensuring mechanism accuracy across different fuel ratios (Aminian et al., 2023). Many researchers studied heavy substances as diesel fuel representatives. Guo et al. (Guo et al., 2021) generated a simplified mechanism (155 species, 645 reactions) for a diesel and natural gas dual-fuel RCCI engine, assuming n-hexadecane as diesel fuel. Using cross-reaction analysis, Liu et al. (Liu et al., 2021) developed a reduced combustion mechanism (150 species, 847 reactions) for a multi-component fuel mixture consisting of toluene, n-dodecane, methane, and methylcyclohexane. They validated it through testing against experimental ignition delay and flame speed data, ensuring its accuracy over a wide range of pressures, temperatures, and equivalence ratios. Zhou et al. (Zhou et al., 2023) introduced a new reduced mechanism (141 species, 739 reactions) for diesel (represented by n-dodecane) and natural gas combustion in HCCI applications. To create a compact model, they used simplification techniques such as directed relation graphs (DRG), DRG with error propagation (DRGEP), and sensitivity analysis.

It is generally acknowledged that the n-heptane oxidation processes accurately depict the properties of diesel fuel combustion in a simplified way (H. Wang et al., 2013). Recently, researchers have shown interest in utilizing reduced mechanisms in which n-heptane is the diesel fuel surrogate. Rahimi et al. (Rahimi et al., 2010) introduced a semi-detailed chemical kinetic mechanism (76 species, 464 reactions) for n-heptane and methane, optimized by a genetic algorithm. The mechanism has been continuously employed to simulate multi-fuel engines (Yin et al., 2024). Hockett et al. (Hockett et al., 2016) developed a reduced chemical mechanism (141 species, 709 reactions) known as CSU141 by combining two detailed mechanisms of n-heptane and natural gas (represented by a combination of methane, ethane, and propane). De Bellis et al. (De Bellis et al., 2022) modified this mechanism by adding a NO<sub>x</sub> sub-mechanism to develop a multi-zone phenomenological combustion model for RCCI dual-fuel engines.

Among the recent articles focusing on n-heptane as the chemical surrogate for the diesel fuel, Zhang et al. (W. Zhang et al., 2019) presented a diesel/natural gas dual fuel mechanism (79 species, 224 reactions), employing peak concentration analysis (PCA) and rate of production (ROP) methods to simplify the GRI-mech 3.0 natural gas mechanism, and then combined it with the 95/5 v/v mechanism of diesel combustion (represented by n-heptane and aromatic combinations). They validated the reduced mechanism through chemical kinetics calculations and CFD simulations. Although it accurately predicts combustion characteristics and pollutant emissions, the scope is limited to medium-load conditions. Huang et al. (Huang et al., 2019) developed a reduced mechanism (143 species, 746 reactions) based on detailed mechanisms for n-heptane, n-butylbenzene, NG, and polycyclic aromatic hydrocarbons (PAH) using the methods of DRGEP, ROP, and sensitivity analysis. They validated their mechanism against

various parameters such as ignition delay and laminar flame speed in the context of HCCI applications, demonstrating its accuracy at varying rates of natural gas substitution. However, the complexity of the mechanism might pose challenges for practical implementation in simulations. Schuh et al. (Schuh and Winter, 2020) proposed a dual fuel combustion mechanism (75 species, 344 reactions), specifically focusing on the ignition and combustion of methane/propane/n-heptane fuel blends in high-pressure environments. They analyzed and updated the Arrhenius parameters to extend the mechanism's application range and included the San Diego nitrogen sub-mechanism to allow for consideration of NO<sub>x</sub> formation.

As discussed, NG/diesel is one of the most promising fueling strategies for RCCI combustion. Although some mechanisms exist to simulate combustion with these fuels, the present work aims to introduce a compact mechanism designed for high-speed simulations, tailored specifically to the operating conditions of these engines. The mechanism is finally validated through simulations of some RCCI engine cases with both 3-D CFD and multi-zone methods and also fundamental ignition delay time (IDT) and laminar burning velocity (LBV) test data.

## 2. MATERIALS AND METHODS

The current work targets to develop a chemical kinetics mechanism capable of simulating combustion processes in engines employing multi-fuel low temperature combustion (LTC) strategies, such as RCCI. The targeted fuels include diesel fuel, chosen for its high reactivity, and low reactivity fuels can be a blend of natural gas and hydrogen (H<sub>2</sub>). It is assumed that n-heptane (C<sub>7</sub>H<sub>16</sub>) serves as the surrogate for diesel fuel, as demonstrated in numerous scientific investigations on light diesel fuel combustion in engines. Additionally, methane (CH<sub>4</sub>) and ethane (C<sub>2</sub>H<sub>6</sub>) hydrocarbons are incorporated into the mechanism to simulate natural gas reactions.

The mechanism is intended for use in 3-D Computational Fluid Dynamics (CFD) simulations and fast multi-zone models, necessitating a limited number of species and reactions to maintain computational affordability. Hence, skeletal and reduced base mechanisms are employed to construct a multi-fuel mechanism.

In the simulation of combustion processes like RCCI, the auto-ignition of the high reactivity fuel (e.g., diesel) is pivotal in determining the onset of combustion. Unlike applications such as furnace burners, this process begins at relatively high pressures (above 20 bar) and low temperatures (below 900 K) in engine environments. Therefore, selecting a proper set of reactions that accurately predict the ignition of n-heptane under such conditions is crucial. To address this, the skeletal n-heptane mechanism developed by Chang et al. (Chang et al., 2022) is chosen to provide the sub-mechanisms pertinent to n-heptane combustion. This skeletal mechanism is based on the detailed n-heptane mechanism developed by Zhang et al. (K. Zhang et al., 2016), which has been validated against experimental IDT and LBV, establishing its efficacy for n-heptane combustion simulation under engine-like conditions.

To consider reactions related to natural gas (NG) combustion, the GRI 3.0 mechanism (Smith et al., 2000), widely acknowledged as a comprehensive model for NG combustion, is merged with the n-heptane mechanism. To enable the mechanism to predict nitrogen oxide emissions, some corresponding species and reactions are added from the mechanism of Wang et al. (B. Wang et al., 2023).

After merging base mechanisms and adding needed species and reactions, to result a compact mechanism, the DRGEP method with sensitivity analysis is employed to reduce the mechanism. In this, operation the IDT for n-heptane and methane/n-heptane blend mixtures and LBV for methane, methane/H<sub>2</sub> and H<sub>2</sub> cases are considered as the base cases.

The resulting mechanism comprises 60 species and 372 reactions, which makes it a really compact one comparing to other existing NG/diesel mechanisms.

As previously mentioned, it is crucial for a mechanism intended for RCCI engine simulations to accurately capture the low-temperature auto-ignition of the highly reactive fuel. In this work, the ignition delay resulting from the n-heptane fuel mechanism is compared to the data presented in the study by Zhang et al. (K. Zhang et al., 2016), and adjustments are made to improve agreement with experimental data, particularly for elevated pressures (20 bar) and low temperatures. In RCCI engines, the auto-ignition initiates in mixtures with lower temperatures than 1000 K. Therefore, the combustion IDT should be predicted well by mechanism in this low temperature conditions.

This tuning process involves modifying certain reactions to enhance the mechanism's predictive capability for experimental data.

All merging, reduction, simulations, and optimization processes are conducted using Converge CFD 3.1 software. In the tuning process, the most sensitive reactions under the specified conditions (n-heptane fuel, high pressure, low temperature) are first identified. The sensitivity to these reactions is illustrated in Fig. 1.

In the subsequent stage, the 20 most sensitive reactions are selected, and their pre-exponential factors are optimized using the NLOPT optimization algorithm, which is integrated into the chemistry module of the Converge software. This optimization process aims to align the simulated IDT with the experimental data reported in (K. Zhang et al., 2016).

To assess the effectiveness of the tuning process, a comparison between the IDT resulting from the initial merged mechanism (prior to tuning) and the tuned mechanism against experimental data is presented in Fig. 2. Accurate modeling of the start of combustion in RCCI engines hinges on the mechanism's ability to predict first-stage ignition delay times, which significantly impacts the low-temperature heat release during engine combustion.

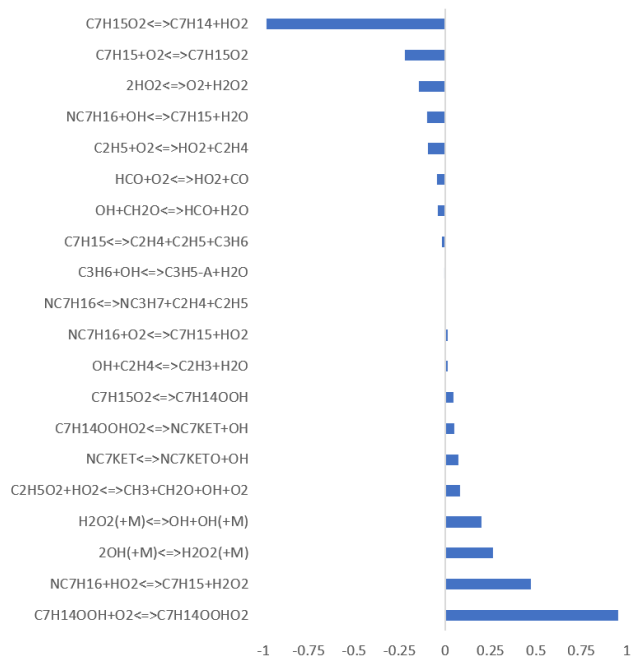


Fig. 1. Sensitivity analyzes for IDTs to identify the most sensitive reactions

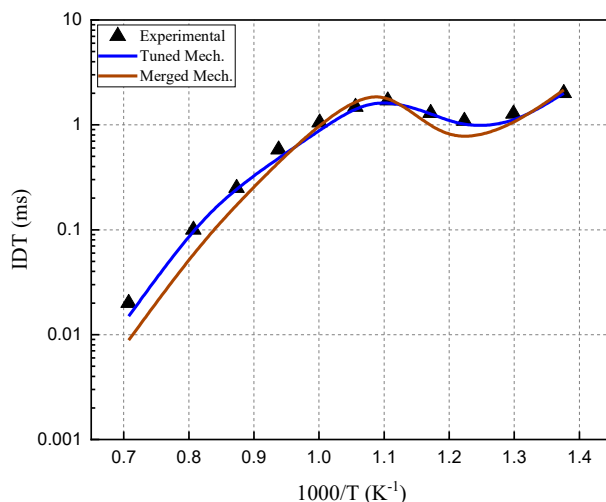


Fig. 2. Comparison of IDT for tuned and merged mechanism with experimental data from Zhang et al. (K. Zhang et al., 2016)

Figure 2 illustrates that following tuning, the mechanism demonstrates improved precision in predicting IDT for n-heptane fuel, particularly within the low-temperature zone. This enhancement renders the mechanism more suitable for simulating engines employing low temperature combustion (LTC) strategies.

### 3. VALIDATION OF THE NEW MECHANISM

In this section, the performance of the new mechanism is evaluated through simulations compared against available experimental test data. Initially, the IDT and LBV produced by the mechanism are scrutinized against experimental data. Subsequently, the applicability of the mechanism is explored in simulating RCCI and dual-fuel engines with a 3-D CFD software and also a multi-zone code.

### 3.1 IDT and LBV simulations

In the context of applying a mechanism for simulation in CI engines, precise prediction of ignition onset time is paramount. Therefore, investigating the IDT resulting from the mechanism across various scenarios is crucial. As the high reactivity fuel in this work is n-heptane and its auto-ignition starts the combustion process, the ability of the mechanism to predict the IDT for n-heptane mixtures is studied in this part.

Figure 3 shows the comparison of IDT values predicted from the new mechanism and mechanisms from Hockett (Hockett et al., 2016) and Rahimi (Rahimi et al., 2010), which are two mechanisms that are widely used for simulation of dual fuel and RCCI combustion in NG/Diesel fueled engines, with experiments of Zhang for n-heptane fuel (K. Zhang et al., 2016). The initial pressure in the tests and simulations is 20 bar and 38 bar, for Fig. 3 (a) and (b) respectively. Hockett mechanism has 150 species and 872 reactions and Rahimi mechanism has 76 species and 483 reactions. The proposed mechanism is seen to give better predictions than the two other mechanisms, although it has less number of species and reactions.

As stated earlier, in the target combustion strategies (RCCI and dual fuel), the auto-ignition process starts from regions with high concentration of the fuel with high reactivity and then the flame propagates into the mixture containing the low reactivity fuel (or fuel blend). Therefore, in the following, the LBV in the mixtures containing target low reactivity fuels, i.e. NG and hydrogen is investigated.

As the first evaluation of flame propagation properties, a 1-D simulation of laminar flame speed in natural gas (NG) was conducted using Converge CFD software. The test case utilized experimental data as reported in (Rozenchan et al., 2002). In this scenario, the initial temperature (pre-burning) was set at 298 K, and initial pressures of 1 atm, 2 atm, and 5 atm were investigated. Figure 4 illustrates the comparison between simulated LBV values and the experimental data reported for different pressures in fuel equivalence ratio ( $\phi$ ) ranged from 0.6 to 1.4. The results indicate that the new mechanism reliably predicts flame propagation within NG mixtures across the investigated range.

The mechanism is supposed to have the potential to be used for the cases of hydrogen addition to the NG fuel. Graphs of Figure 5 shows the result of LBV simulation for H<sub>2</sub>/NG fuel blends in both air and O<sub>2</sub>/He oxidizer environment. Comparison to experimental cases reported by Donohoe et al. (Donohoe et al., 2014) shows that the new mechanism has a good ability to predict flame propagation in H<sub>2</sub>/NG blend mixtures, as well as mixtures of air and NG. Specially, it is seen that the simulations have a very good accuracy to predict the results close to experimental data, for lower equivalence ratios than 1 in the low reactivity mixture, which is the case that usually happens in the target dual fuel and RCCI engines.

### 3.2 Simulation of RCCI NG/diesel fueled engines

To verify the mechanism ability in simulation of combustion in NG/diesel fueled engines, two marine engines with RCCI combustion mode are selected to be both experimentally and numerically investigated. First one is the 6L20CRDF research engine, a variant of the Wärtsilä 20 platform, featuring a 6-cylinder inline dual fuel (DF) configuration (Kakooe et al., 2023). The second investigated marine engine is a Wärtsilä single-cylinder research engine (SCRE), which adheres to the DF variant of Wärtsilä's W31 engine platform (Åstrand, 2016). Key specifications of both test engines are outlined in Table 1.

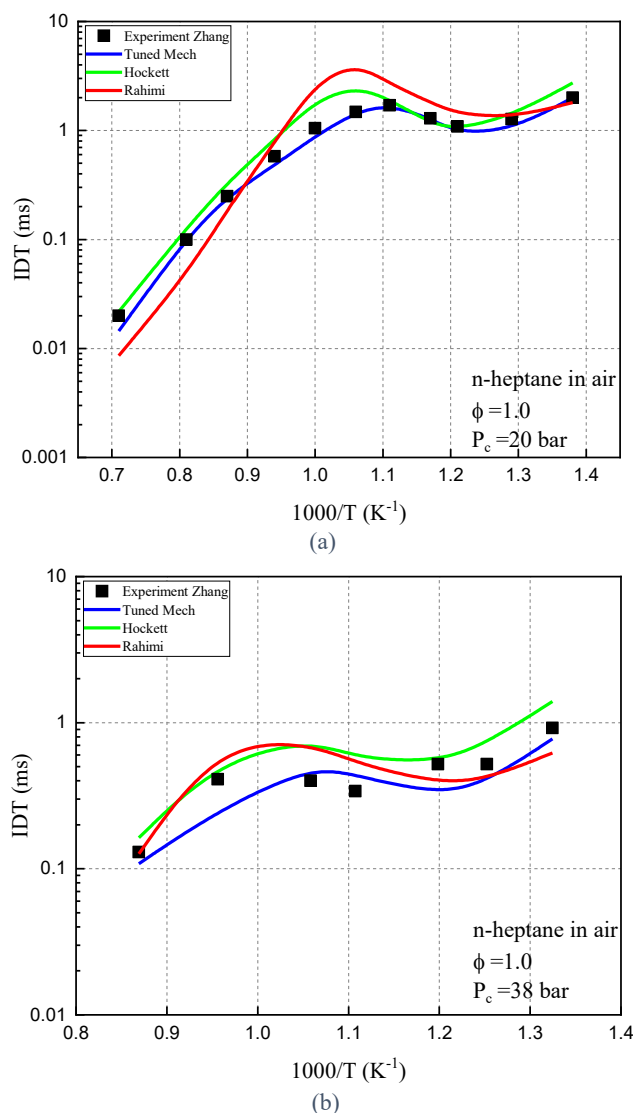


Fig. 3. Comparison of IDT for the new mechanism with two other mechanisms (Hockett (Hockett et al., 2016) and Rahimi (Rahimi et al., 2010)) and also experimental data from Zhang et al. (K. Zhang et al., 2016), for (a) pressure of 20 bar and (b) pressure of 38 bar



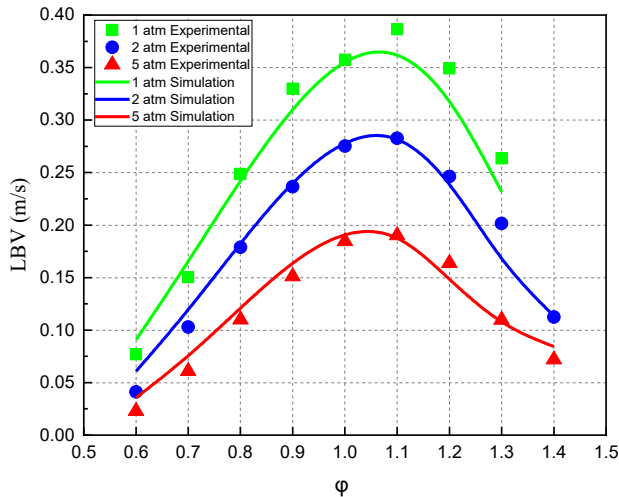


Fig. 4. LBV for NG mixtures in different pressures, comparison of simulations (lines with small marks) with experimental data (marks) from Rozenchan et al. (Rozenchan et al., 2002)

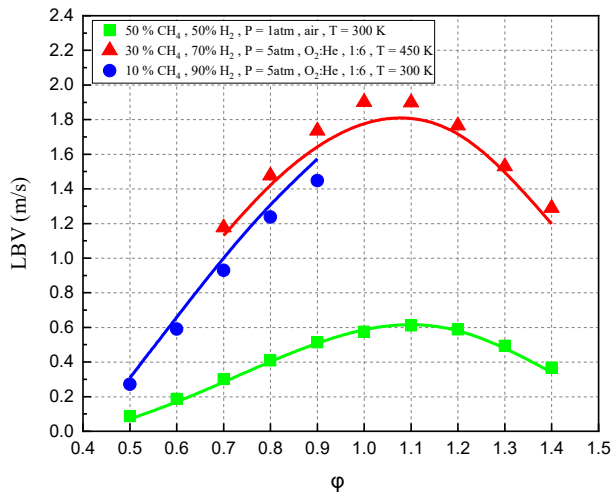


Fig. 5. LBV for NG/H<sub>2</sub> mixtures in different pressures and blend ratios comparison of simulations (lines with small marks) with experimental data (marks) (Donohoe et al., 2014)

### 3.4 3-D CFD simulation

In this part, 3-D CFD simulation of a closed cycle (IVC to EVO) for Wärtsilä W20 engine with RCCI combustion is performed using Converge CFD v3.1 software. The SAGE solver, specializing in transient chemical kinetics, was used in transient mode to compute species mass fractions at each time step before solving the transport equation with the PISO technique. Diesel fuel properties were represented by the "DIESEL2" surrogate from the CONVERGE library for spray and mixing modeling. Liquid injection was modeled using the Blob injection sub-model, with parcel diameter set to the nozzle hole size. Diesel spray was simulated with the KH-RT model considering high injection pressure. Heat transfer was modeled using the Standard Wall Function, while turbulence was captured via the RNG  $k-\epsilon$  model. Adaptive mesh refinement was used for a base grid size of 0.002 m.

Computational runs are conducted to evaluate RCCI operation of the engine using diesel and NG fuels under mid-load conditions. For the simulated case, the energy share of diesel and NG fuels were 7% and 93% respectively. In this scenario, diesel fuel injection has commenced about 70° before piston TDC, allowing for partial mixing of diesel fuel with the in-cylinder mixture, facilitating RCCI combustion mode.

Figure 6 illustrates a comparison between simulated and measured in-cylinder pressure and heat release rate (HRR) to assess the performance of the mechanism in simulating diesel/NG combustion. Normalized values are presented in the figure, relative to their maximum values. The depicted results demonstrate a high level of agreement between simulation and measurement. Both the pressure trace and heat release rate are simulated with acceptable accuracy. Important operating parameters such as SOC, maximum pressure and maximum HRR are seen to be predicted with a good accuracy. The results prove the effectiveness of the new mechanism for simulating combustion processes in mixtures containing diesel fuel and NG under internal combustion engine conditions.

Table 1: Specification of the Wärtsilä test engines

Engine model	Wärtsilä W20	Wärtsilä W31
Displacement	8.8 liter	32.45 liter
Stroke/Bore	1.4	1.39
Test Speed	1000 rpm	External compressor
Air System	Two-stage turbocharged (in series)	with air temperature and pressure control (up to 10 bar)
High reactivity fuel system	Common rail	Common rail 2.0 with twin needle injector; and multi-injection capability
Low reactivity fuel system	Low-pressure; multi-point, upstream of the intake valve	Low-pressure, multi-point, upstream of the intake valves
Valve train	4 valves per cylinder, fully variable hydraulic valve train	four valves with swirl + tumble ports; variable intake valve closure (VIC); fixed exhaust valve opening (EVO)

### 3.5 Quasi-dimensional simulation

In this section, simulations are conducted using the UVATZ code, a multizone model developed at Vaasa University specifically designed to model combustion in RCCI engines (Vasudev, Cafari, et al., 2022). UVATZ simulates low-temperature combustion concepts driven by chemical kinetics and in the present study, is parameterized for natural gas and diesel RCCI combustion. The model accounts for fuel and thermal stratification, in-cylinder turbulence, IVC temperature, and residual gas composition. It consists of 12 zones, two disc-shaped for the cylinder head and piston boundary layers, and 10 annular zones, with the outermost in contact with the liner, capturing bulk inhomogeneity. Heat and mass flows are modeled through gradient-based transport.

HRF stratification is simplified, with mass linearly distributed across the zones, and evaporation enthalpy proportional to HRF mass.

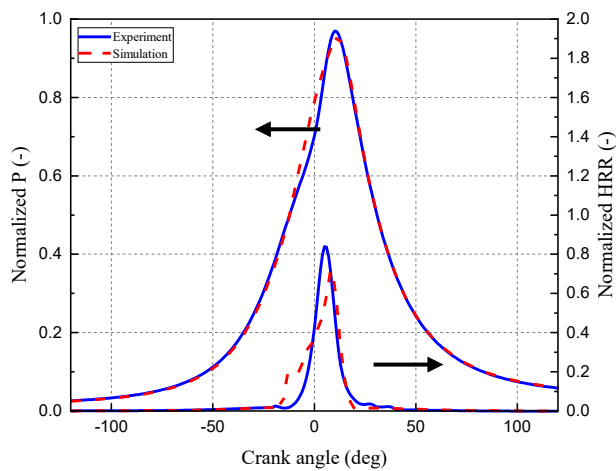


Fig. 6. Comparison of in-cylinder pressure and HRR (normalized values) between experiments and simulation for an NG/diesel RCCI engine in CFD simulation

The simulated engine under study is a Wärtsilä single-cylinder research engine (SCRE), which adheres to the dual fuel (DF) variant of Wärtsilä's W31 engine platform (Åstrand, 2016).

In the tests conducted for this segment, the load is about 25% of the full load conditions. Liquefied natural gas (LNG) with a methane number of 80 serves as the low reactivity fuel, while light fuel oil (LFO) acts as the high reactivity fuel. The start of injection of LFO occurs  $65^\circ$  before top dead center (TDC) to ensure the onset of RCCI combustion. Notably, 70% of the total fuel energy is attributed to NG, while the remaining 30% comes from LFO.

The 12-zone configuration of the UVATZ model, which is employed in the present work, has the ability to represent the subtleties of fuel and thermal stratification, in-cylinder turbulence, intake valve closing (IVC) temperature, and the composition of residual burned gas. These zones interact with one another through the transfer of heat, mass, and work, making them dynamic entities in the simulation world. The model also accounts for heat loss to the walls, which is modeled using a specific correlation (Kakooe et al., 2023).

For simulations, n-heptane serves as a surrogate for LFO, whereas a mixture of methane and ethane is employed to represent NG. A 13-zone model is utilized for quasi-dimensional modeling of the RCCI combustion process.

Figure 7 presents a comparison between measured and simulated in-cylinder pressure curves and heat release rate. The simulation process, executed on a PC with an i7-13700H processor, required 219 seconds for a full-cycle simulation to converge. For the presented engine test case, it took 4 cycles for the model to converge. It is evident that certain parameters, like SOC and peak pressure, have been accurately predicted. This comparison highlights the mechanism's suitability for integration into multi-zone

model codes to simulate the RCCI combustion process.

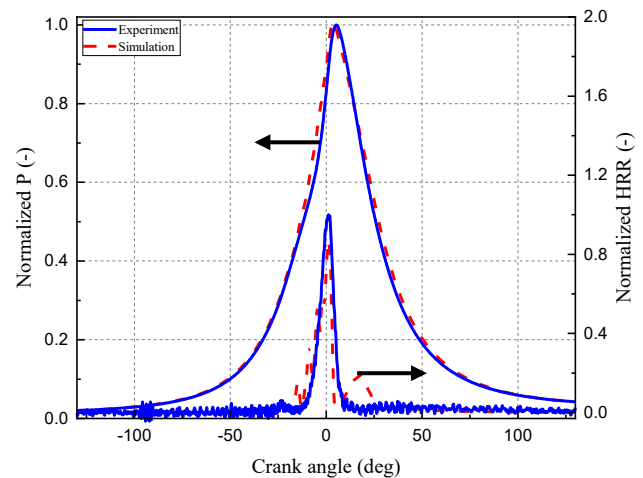


Fig. 7. Comparison of in-cylinder pressure for experiments and simulation for an NG/diesel RCCI engine with MZM code

#### 4. CONCLUSION

A reduced chemical kinetics mechanism for NG, H<sub>2</sub> / n-heptane blends was proposed in this study. This mechanism contained 60 species and 372 reactions, making it cost-effective for simulation of combustion engine cases. The mechanism was formed based on merging existing mechanisms for n-heptane, and NG. The mechanism was tuned for better resolving n-heptane auto-ignition in low temperatures and elevated pressures. The mechanism's accuracy was initially assessed using fundamental Induction Delay Time (IDT) and Laminar Burning Velocity (LBV) experimental data from the literature for the blended target fuels. It was demonstrated that 0-D simulations employing the new mechanism could accurately predict the IDT for n-heptane fuel mixtures with acceptable precision, better than some of other widely used mechanisms for NG/diesel fueled engines. Furthermore, conducted 1-D simulations substantiated the mechanism's capability to predict LBV in mixtures of target low reactivity fuels, considering mixtures with NG and NG/H<sub>2</sub> blends. After verification with data from fundamental tests, 3-D CFD simulation utilizing this new mechanism were validated with real engine test data, for NG/diesel cases. Good agreement between test results and simulations demonstrates the mechanism's suitability for simulating engines operating in multi-fuel modes such as RCCI and dual fuel, with diesel as the high reactivity fuel and a blend of NG and H<sub>2</sub> as the low reactivity fuel.

#### ACKNOWLEDGEMENTS

Authors would like to acknowledge the financial supports provided by Business Finland through CASEMATE project (Ref. 2911/31/2022) under Wärtsilä ZEM (zero emission marine) ecosystem.

## REFERENCES

- Aminian, E., Gharehghani, A., and Mirsalim, M. (2023). Proposing a chemical kinetic mechanism for biodiesel/NG blend in RCCI engine condition. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 09544070231186278. doi:10.1177/09544070231186278
- Åstrand, U. (2016). The Wärtsilä 31 - The World's Most Efficient 4-Stroke Engine. *Journal of The Japan Institute of Marine Engineering*, 51(2). doi:10.5988/jime.51.203
- Aziz, M. A., Yusop, A., Nafiz, D., and Kumarasamy, S. (2022). Effect of Alcohol, Natural Gas and Renewable Diesel on Reactivity Controlled Compression Ignition Engine: A Critical Review. *SSRN Electronic Journal*. doi:10.2139/ssrn.4126186
- Chang, Y., Jia, M., Wang, P., Niu, B., and Liu, J. (2022). Construction and derivation of a series of skeletal chemical mechanisms for n-alkanes with uniform and decoupling structure based on reaction rate rules. *Combustion and Flame*, 236. doi:10.1016/j.combustflame.2021.111785
- De Bellis, V., Malfi, E., Lanotte, A., Fasulo, G., Bozza, F., Cafari, A., Caputo, G., and Hyvönen, J. (2022). Development of a phenomenological model for the description of RCCI combustion in a dual-fuel marine internal combustion engine. *Applied Energy*, 325, 119919. doi:10.1016/j.apenergy.2022.119919
- Donohoe, N., Heufer, A., Metcalfe, W. K., Curran, H. J., Davis, M. L., Mathieu, O., Plichta, D., Morones, A., Petersen, E. L., and Güthe, F. (2014). Ignition delay times, laminar flame speeds, and mechanism validation for natural gas/hydrogen blends at elevated pressures. *Combustion and Flame*, 161(6). doi:10.1016/j.combustflame.2013.12.005
- Dwarshala, S. K., Rajakumar, S. S., Kummitha, O. R., Venkatesan, E. P., Veza, I., and Samuel, O. D. (2023). A Review on Recent Developments of RCCI Engines Operated with Alternative Fuels. *Energies*, 16(7). doi:10.3390/en16073192
- Fakhari, A. H., Gharehghani, A., Salahi, M. M., Mahmoudzadeh Andwari, A., Mikulski, M., Hunicz, J., and Könnö, J. (2023, August 28). *Numerical Investigation of Ammonia-Diesel Fuelled Engine Operated in RCCI Mode*. doi:10.4271/2023-24-0057
- Gharehghani, A., Salahi, M. M., Andwari, A. M., Mikulski, M., and Könnö, J. (2023). Reactivity enhancement of natural gas/diesel RCCI engine by adding ozone species. *Energy*, 274. doi:10.1016/j.energy.2023.127341
- Gössnitzer, C., Posch, S., and Pirker, G. (2022). Development of a framework for internal combustion engine simulations in openfoam. *World Congress in Computational Mechanics and ECCOMAS Congress*. doi:10.23967/eccomas.2022.089
- Guo, X., Chen, Y., Huang, H., Chen, Y., Liu, M., Lei, H., Deng, B., and Chen, C. (2021). Development of a Diesel/Natural Gas Mechanism Model for the CFD Simulation of Dual-Fuel Engine. *ACS Omega*, 6(33), 21543–21555. doi:10.1021/acsomega.1c02514
- Hockett, A., Hampson, G., and Marchese, A. J. (2016). Development and Validation of a Reduced Chemical Kinetic Mechanism for Computational Fluid Dynamics Simulations of Natural Gas/Diesel Dual-Fuel Engines. *Energy and Fuels*, 30(3), 2414–2427. doi:10.1021/acs.energyfuels.5b02655
- Hossein Fakhari, A., Gharehghani, A., Mahdi Salahi, M., and Mahmoudzadeh Andwari, A. (2024). RCCI combustion of ammonia in dual fuel engine with early injection of diesel fuel. *Fuel*, 365. doi:10.1016/j.fuel.2024.131182
- Huang, H., Lv, D., Zhu, J., Zhu, Z., Chen, Y., Pan, Y., and Pan, M. (2019). Development of a new reduced diesel/natural gas mechanism for dual-fuel engine combustion and emission prediction. *Fuel*, 236, 30–42. doi:10.1016/j.fuel.2018.08.161
- Kakooe, A., Gharehghani, A., and Mostafaei, M. (2022). Development of a reduced chemical kinetic mechanism for biodiesel/natural gas mixture. *Fuel*, 312, 122920. doi:10.1016/j.fuel.2021.122920
- Kakooe, A., Vasudev, A., Smulter, B., Hyvonen, J., and Mikulski, M. (2023). A Predictive 1D Modeling Framework for Reactivity-Controlled Compression Ignition Engines, via a Chemistry-Based, Multizone Combustion Object. *SAE Technical Papers*. doi:10.4271/2023-24-0001
- Liu, Z., Yang, L., Song, E., Wang, J., Zare, A., Bodisco, T. A., and Brown, R. J. (2021). Development of a reduced multi-component combustion mechanism for a diesel/natural gas dual fuel engine by cross-reaction analysis. *Fuel*, 293, 120388. doi:10.1016/j.fuel.2021.120388
- Rahimi, A., Fatehifar, E., and Saray, R. K. (2010). Development of an optimized chemical kinetic mechanism for homogeneous charge compression ignition combustion of a fuel blend of n-heptane and natural gas using a genetic algorithm. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 224(9), 1141–1159. doi:10.1243/09544070JAUTO1343
- Rozenchan, G., Zhu, D. L., Law, C. K., and Tse, S. D. (2002). Outward propagation, burning velocities, and chemical effects of methane flames up to 60 ATM. *Proceedings of the Combustion Institute*, 29(2). doi:10.1016/s1540-7489(02)80179-1
- Salahi, M. M., Esfahanian, V., Gharehghani, A., and Mirsalim, M. (2017). Investigating the reactivity controlled compression ignition (RCCI) combustion strategy in a natural gas/diesel fueled engine with a pre-chamber. *Energy Conversion and Management*, 132. doi:10.1016/j.enconman.2016.11.019
- Salahi, M. M., Mahmoudzadeh Andwari, A., Könnö, J., and Gharehghani, A. (2022). *Hydrogen and ammonia fuelled internal combustion engines, a pathway to carbon-neutral fuels future*.
- Schuh, S., and Winter, F. (2020). Dual Fuel Reaction Mechanism 2.0 including NOx Formation and Laminar Flame Speed Calculations Using Methane/Propane/n-Heptane Fuel Blends. *Energies*, 13(4). doi:10.3390/en13040778

- Smith, G. P., Golden, D. M., Frenklach, M., Moriarty, N. W., Eiteneer, B., Goldenberg, M., Bowman, C. T., Hanson, R. K., Song, S., Gardiner Jr., W. C., Lissianski, V. V., and Qin, Z. (2000). GRI-Mech 3.0. In [http://www.Me.Berkeley.Edu/Gri\\_Mech/](http://www.Me.Berkeley.Edu/Gri_Mech/).
- Vasudev, A., Cafari, A., Axelsson, M., Mikulski, M., and Hyvonen, J. (2022). Towards Next Generation Control-Oriented Thermo-Kinetic Model for Reactivity Controlled Compression Ignition Marine Engines. *SAE Technical Papers*. doi:10.4271/2022-01-1033
- Vasudev, A., Mikulski, M., Balakrishnan, P. R., Storm, X., and Hunicz, J. (2022). Thermo-kinetic multi-zone modelling of low temperature combustion engines. *Progress in Energy and Combustion Science*, 91. doi:10.1016/j.peccs.2022.100998
- Wang, B., Dong, S., Jiang, Z., Gao, W., Wang, Z., Li, J., Yang, C., Wang, Z., and Cheng, X. (2023). Development of a reduced chemical mechanism for ammonia/n-heptane blends. *Fuel*, 338. doi:10.1016/j.fuel.2022.127358
- Wang, H., Jiao, Q., Yao, M., Yang, B., Qiu, L., and Reitz, R. D. (2013). Development of an n-heptane/toluene/polyaromatic hydrocarbon mechanism and its application for combustion and soot prediction. *International Journal of Engine Research*, 14(5), 434–451. doi:10.1177/1468087412471056
- Yin, C., Wang, C., Shen, X., and Zhang, Z. (2024). Experimental and simulation study of diesel/methane/hydrogen triple-fuel combustion progression in a heavy-duty optical engine. *International Journal of Hydrogen Energy*, 62, 562–578. doi:10.1016/j.ijhydene.2023.10.198
- Zhang, K., Banyon, C., Bugler, J., Curran, H. J., Rodriguez, A., Herbinet, O., Battin-Leclerc, F., B'Chir, C., and Heufer, K. A. (2016). An updated experimental and kinetic modeling study of n-heptane oxidation. *Combustion and Flame*, 172. doi:10.1016/j.combustflame.2016.06.028
- Zhang, W., Chang, S., Wu, W., Dong, L., Chen, Z., and Chen, G. (2019). A diesel/natural gas dual fuel mechanism constructed to reveal combustion and emission characteristics. *Energy*, 179, 59–75. doi:10.1016/j.energy.2019.04.106
- Zhou, S., Zhou, W., Xi, H., Shreka, M., and Zhang, Z. (2023). Development and validation of a new reduced diesel/natural gas mechanism under engine-relevant conditions. *Energy*, 24(7), 3119–3129. doi:10.1177/14680874221143900

## Driving force model for a real-time control concept of a hybrid heavy duty vehicle

Jari Ruuska\* Aki Sorsa\* Isa Banagar\*\* Perttu Niskanen\*\*  
Amin Mahmoudzadeh Andwari\*\* Mika Ruusunen\* Emil Kurvinen\*\* Juho Könnö\*\*

\*Environmental and Chemical Engineering, Control Engineering, University of Oulu, Finland,

\*\*Materials and Mechanical Engineering, Machine and Vehicle Design, University of Oulu, Finland,  
{forename.surname}@oulu.fi

**Abstract:** The electrification of heavy vehicles and work machinery is developing rapidly. The main motivators are green transition and requirements from the customers. In Finland, there are many high-tech market-leading companies in this segment. Mass-produced equipment and machines are suitable for general applications and thus tailoring design for specific conditions and/or needs results in better productivity and efficiency. In heavy electric vehicle applications, the challenge is to make new products economically viable and configure them to meet customer needs. In these applications, the number of solutions is an order of magnitude higher than in traditional mechanical solutions. However, electronic solutions enable new features and energy efficiency improvements to have measurable benefits in the application. The research investigates the effects of electric axle solutions for hybrid heavy duty vehicles. Modelling and simulations consider both the effects of engine and usage of battery charge and surroundings of vehicle, for example road profile, traffic, outdoor temperature, and friction. A system level model of a vehicle has been utilized to simulate its longitudinal dynamics interacting with estimated surroundings followed by model-based control. The planned route can be made further favorable by utilizing real-time model predictive control (MPC) receiving online data from changing conditions. MPC gives new suggestions for optimal battery usage based on deviations from the best matching model from a database. Control strategy is important when considering economic benefits for a hybrid heavy duty vehicle with a high degree of freedom in system design.

*Keywords:* Electrification, Green transition, Model Predictive Control, Model Based System Engineering, Systematic Machine Design

### 1. INTRODUCTION

An example of promising solutions for pollution reduction are electric and hybrid electric vehicles (EV/HEV), which can be exploited for a safe environment and sustainable transportation. Designing these vehicles requires different optimization procedures, for example components, systems, and controls (Ehsani et al. 2021). A review article of path tracking strategies used in autonomous vehicle control design discusses different elements of modelling process including the criteria for evaluating the controller's performance (Ruslan et al. (2023)). Extremely important part for enabling the optimization of the battery usage during the route is to have a competitive battery management system. Advanced battery management, which consists of three progressive layers. A comprehensive overview of each layer is presented, and future trends of next-generation battery management are discussed (Dai et al. 2021). A broad review to optimize the power flow in EV powertrains using multispeed discrete transmission, continuously variable transmission and multi-motor configurations. The potentials and challenges

regarding for example environmental issues are discussed. They can be applied to hybrid vehicles as well. As the overall development is proceeding rapidly, it is getting more and more challenging to be able to answer all demands. A key issue is to develop optimum vehicle fulfilling, for example tighter emission regulations. This leads to reduced emissions of new vehicles, more research of advanced materials for energy storage, better vehicle connectivity and more investments in autonomous technologies. However, it causes, for example sustainability issues in production and mining, higher electricity demand requires new electricity production and socio-economic issues on technology migration (Mazali et al. 2022). The EV vehicle inverse dynamics model was developed. Then vehicle states and kinematic constraints were used to formulate the servo constraints. Finally, a procedure for optimizing trajectories was developed. The results show that the optimal trajectory uses the least amount of energy (Min et al. 2023). Hybrid powertrains having two or more different energy sources, questions arise in terms of HEV structure selection, components sizing, and energy management control. Control variables



optimization is vital to find the set of optimal control rules for the minimum fuel consumption. The dynamic programming approach is a common method because of its unique ability to find the global optimum solution with a certain degree of precision. This computationally demanding optimization method combined with a gradient-based optimization algorithm was used in a systematic way to reduce execution time and to increase the precision of the result (Cipek et al. 2020). The optimization of battery usage in a hybrid vehicle to minimize fuel consumption is a very complex problem. Basically, this means increasing the efficiency of the combustion engine efficiency and recovering electrical energy by charging the battery when driving or braking (Anselma 2022). In (Lei et al. 2018), four operating modes were used, and they were electric driving mode, driving and charging mode, combustion engine driving mode and hybrid driving mode.

A review of architectures and control strategies for the dual-motor coupling propulsion system used in battery electric vehicles is presented in this article. The article describes different architectures, reviews the means of mechanical coupling and transmission, electromechanical configurations, and summarizes approaches to the control of this emerging class of battery electric vehicles. Discussion comparing the advantages and disadvantages of dual-motor coupling propulsion system technology for battery electric vehicles is presented, as well as research challenges and prospects being discussed. (Wang et al. 2022)

The integrated energy management and engine control system for HEV is introduced. The synergy of artificial intelligent control and prior information, for example about route, can be exploited to boost the control performance together, with the engine being optimally controlled. (Zhang et al. 2022)

Having as an objective to decrease fuel consumption, the implementation of an adaptive, optimal neuro-fuzzy inference control was introduced (Saju et al. 2022). When evaluating fuel and electrical energy capabilities, it is usually assumed that the route and velocity profiles are known (Anselma 2022). By applying models, an optimal battery usage plan can be developed. For example, nonlinear programming, genetic algorithms and dynamic programming have been used in such optimization tasks. The problem, however, is that the overall driving power demand must be approximated accurately, which is very difficult in practical cases. For real-time control, the equivalent consumption minimization strategy and model predictive control (MPC) have been used. For MPC to be efficient, the model used must be accurate for future driving information estimation, which is not necessarily the case (Peng et al. 2017). Being capable of managing multi-variable problems and to consider constraints on states and control actions, with capability to predict future behavior of

the process, MPC is widely used for trajectory tracking. This literature review discusses the research conducted from 2015 until 2021 on model predictive path tracking control (Stano et al. 2023).

The design of a path tracking controller for autonomous vehicles is addressed in this paper. The Reference Aware MPC is reformulated to guarantee closed-loop stability, while maintaining a safe and comfortable ride, and minimizing wear and tear of vehicle components. For usability in online operation, a novel model for the nonlinear curvature response of the vehicle is proposed by means of Kalman filtering. (Pereira et al. 2023)

Different technologies' potential for fuel consumption improvement of heavy-duty vehicle has been investigated in literature (Dünnebeil and Keller Heidelberg 2015); Schade (unpublished). According to high interaction between different technology packages system-level simulation should be implemented to overcome the complexity of powertrain design (Delgado et al. 2017). Developing a hybrid powertrain, system-level simulations enable the possibility to calculate vehicle fuel consumption and battery state of the charge for which are the main control strategy objectives (Enang and Bannister 2017).

The aim of this research is to develop a tool which could give a good platform to scheme suitable structures for electric axle. The goal of this paper is to first discuss shortly the concept of the planning process and then more in detail those parts, which concern the control strategy of vehicle environment and main variables effecting on it. This paper considers an approach where an approximation of the overall driving power is made with the simple driving force equation.

## 2. SYSTEM MODELLING FOR OPTIMIZATION

The optimization of battery usage of a hybrid vehicle requires a couple of models. First, it needs a model for the resistive forces acting on a vehicle that must be overcome with the force provided by either the combustion engine or battery. This overall force is called the driving force in (Lei et al. 2017) and (Koch et al. 2021). The optimization scheme also needs a model of the vehicle including the battery model describing the usage and charging of the battery. This paper concentrates on the model of the driving force while the vehicle model is described in (Banagar et al. 2024).

### 2.1 The driving force model

The driving force model is described in literature for example in (Lei et al. 2017), (Koch et al. 2021), (Anselma 2022) and (Chu et al. 2022). The references use different notations and thus the notations in (Koch et al. 2021) are adopted here. The driving force denoted by  $F_W$  is the sum of four forces which are functions of vehicle velocity ( $v$ ) or acceleration ( $a$ ). These forces are the resistance due to road slope with angle  $\alpha$ , rolling friction, aerodynamic drag and the force to overcome vehicle

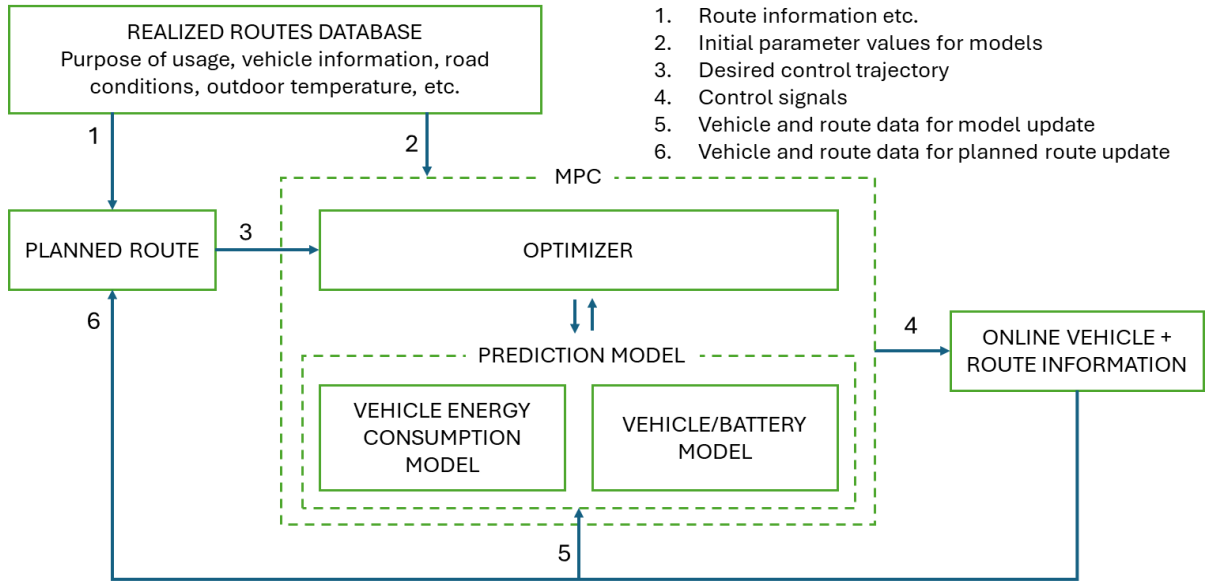


Fig. 2. The overall MPC-based control concept.

inertia. The driving force is calculated in equation (1) from (Chu et al. 2022):

$$F_W = mgsin(\alpha) + mgf_r \cos(\alpha) + \frac{c_w A_f v^2}{2.115} + me_i a. \quad (1)$$

Above in equation (1),  $m$  is the vehicle mass,  $g$  is the gravitational constant,  $f_r$  is the friction coefficient,  $c_w$  is the drag coefficient,  $A_f$  is the vehicle frontal area and  $e_i$  is an equivalence factor. The torque required at wheels is given by equation (2):

$$T_W = F_W r_W. \quad (2)$$

Above,  $r_W$  is the radius of the wheels.

### 2.2 Vehicle model

The target vehicle is a rigid chassis heavy-duty truck with an 8x4 axle configuration and a gross weight of 33 tons. The plan is to replace one axle with an e-axle. The e-axle system includes for example a 15-kwh battery and an electric motor with a rated torque capacity of 880 N.m. The longitudinal dynamic model of the vehicle has been developed by implementing the AVL Cruise M vehicle module. AVL Cruise M /AVL Cruise M/ is a dedicated tool for vehicle and powertrain components simulation. Different powertrain components of the targeted vehicle have been modeled using an available dataset from the vehicle. A map-based model has been selected for the internal combustion engine and electric motor. The battery pack has been simulated implementing an Equivalent Circuit Model (ECM) component of AVL Cruise M. More detail about the parameters and simulation set up has been provided in (Banagar et al. 2024). The schematic of the vehicle powertrain architecture and the simulation set-up have been depicted in Fig 1

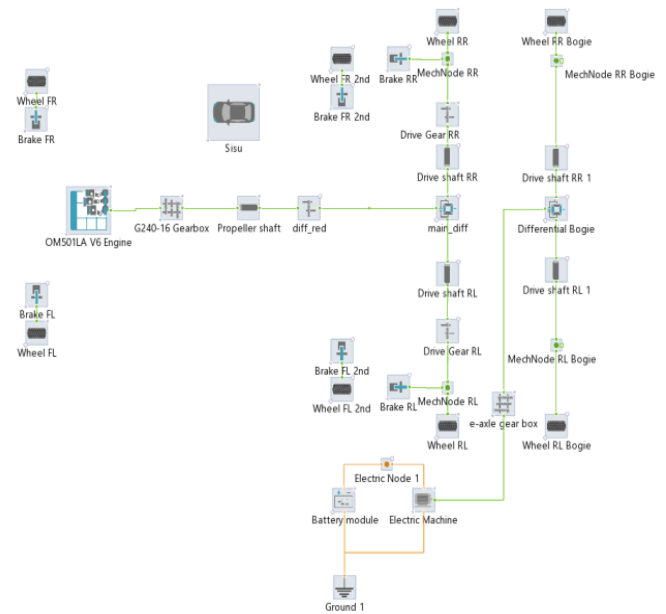


Fig. 1. The schematic of the vehicle's powertrain architecture.

### 2.3 Definition of control concept

The aim of this research is to develop a general approach to optimally use HEV's electric axle and thus reduce fuel consumption. The approach is based on two models which describe the energy consumption and battery usage of HEV. This paper introduces the model for energy consumption and compares its results to real measured data. The overall control concept is presented in Fig. 2. The main idea is to plan the route beforehand. However, no model can describe that perfectly. Thus, a simple model that can capture the main trends is used here but the battery usage plan is updated online as the route proceeds. Also, the models can

be updated online because there are a lot of variables that are not considered with models.

### 3. RESULTS AND DISCUSSION

#### 3.1 Data description

The data used was from a real truck that drove a route from Tornio to Rovaniemi in Finland. The most important variables used for calculating the driving force according to (1) were speed and altitude. The data included also information about actual engine torque that was used for comparison to see if (1) and (2) can describe the torque needed with adequate accuracy. At this point, it is already worth mentioning that (1) and (2) give purely theoretical power requirement and thus a perfect correlation is neither expected nor needed. Instead, a rough estimate is enough because control actions are aimed to be updated continuously.

Data included about 116000 data points sampled at 0.05 second intervals. A moving average filtering without overlap was applied to lower the sampling time to 1 s. The computations are carried out with this data while for model validation the data is further filtered with a 2 min moving median filter. The altitude and speed measured are shown in Figs. 3 and 4, respectively.

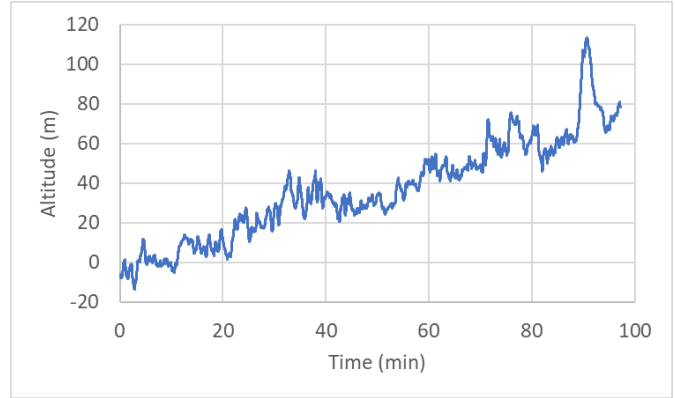


Fig. 3. The measured altitude.

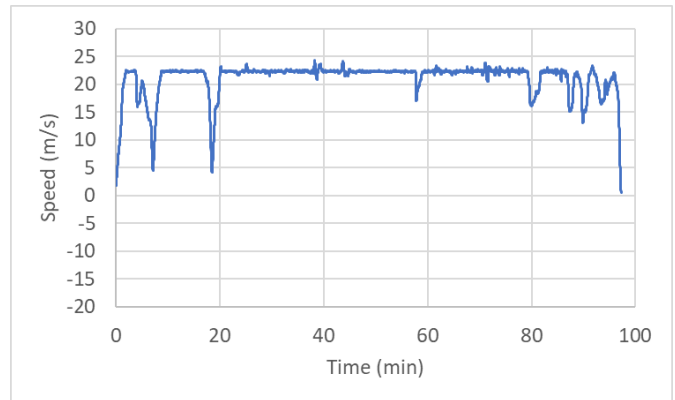


Fig. 4. The measured speed.

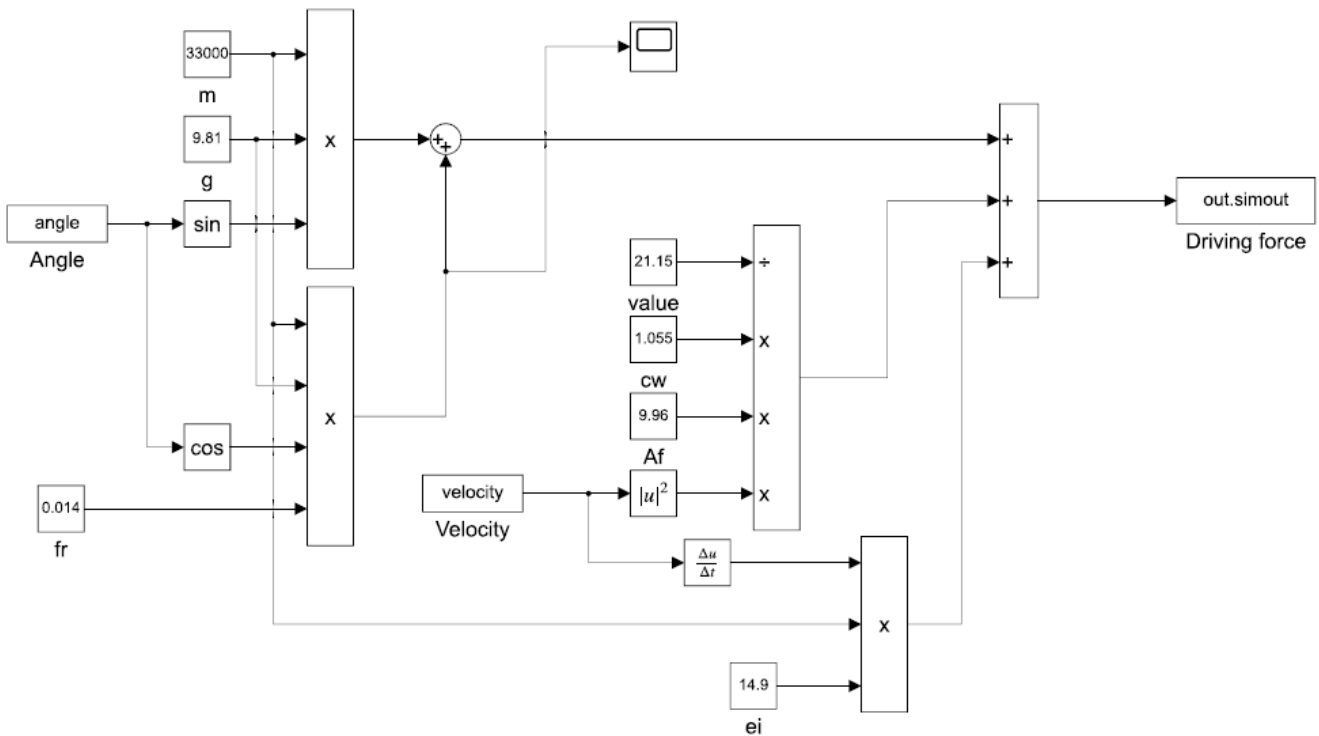


Fig. 5. Simulink® implementation of the driving force model.

### 3.2 The Simulink model of the driving force

The model described in 2.1 is implemented in Simulink® environment. The parameter values are taken from /Kinnunen (2023)/ and they are given in Table 1. The model is shown in Fig. 2. Also, the driving force is a function of distance. Thus, the overall driving force required is obtained as the sum of the model output in Fig. 5.

Table 1. The parameter values used.

$m$	$f_r$	$c_w$	$A_f$	$e_i$
33000 kg	0.014	1.055	9.96 m <sup>2</sup>	14.9

### 3.3 Driving force model results

Driving force was computed according to (1) with the simulator in Fig. 5. The 2 min median filter was applied to obtain the driving force shown in Fig. 6. The high peaks of the driving force are associated with accelerations while the low peaks are associated with decelerations observed in Fig. 4.

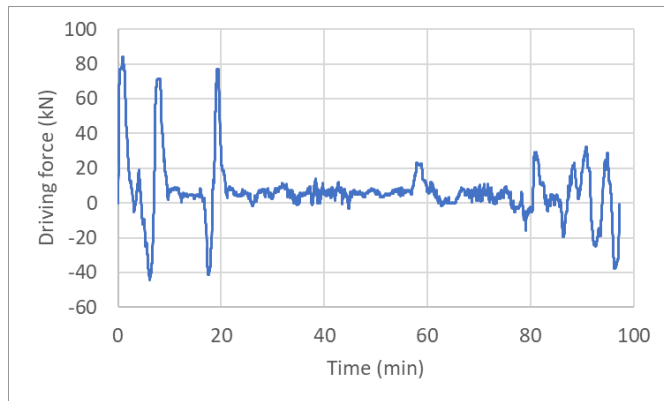


Fig. 6. The simulated driving force.

Figure 7 shows the driving force computed as a function of the actual engine percentage torque measured. The correlation between these is calculated to be 0.67. As expected, the theoretical model cannot explain the actual measurement data perfectly but only the trend is captured. This is, however, expected to be enough because the control actions will be continuously updated with MPC in the overall concept presented in Fig. 2.

Figure 8 shows the fuel consumption as a function of actual engine percentage torque. The figure shows that they are highly correlated. This means that high fuel consumption is associated with high torque. This observation combined with the relationship in Fig. 7 tells us that driving force peaks observed in Fig. 6 should be avoided. This knowledge can be used when defining the optimal battery consumption trajectory for MPC.

### 3.4 Control concept

The first steps towards the control concept introduced in Fig. 2 are presented in this paper. The driving force model is based on physics, and it is expected to give adequate information for identifying the most promising instances to

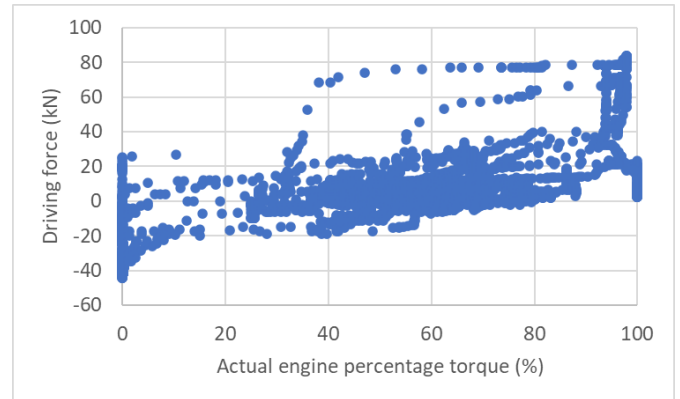


Fig. 7. The driving force computed as a function of actual engine percentage torque.

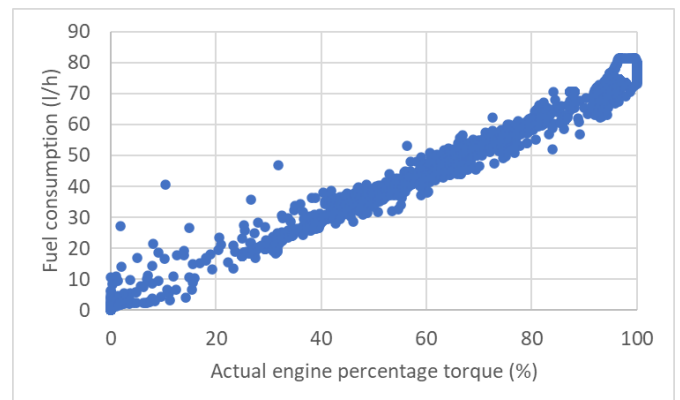


Fig. 8. The fuel consumption as a function of actual engine percentage torque.

use the e-axle for reducing the fuel consumption of the vehicle. In the future, the driving force model is further studied and further validated with more data. Of especial interest are the parameters of the model which are aimed to be defined automatically from the actual data. The driving force model is also complemented with the vehicle model to gain the information needed for making the battery usage plan.

When the models mentioned above are linked to each other, MPC will be implemented into the control approach. It will first be tested in a simulator environment with the collected data. First implementations use constant model parameters, but the alternative to continuously update them and thus adapt the models to current situation will be studied. This is because many variables have an influence on the vehicle's energy consumption and battery operation, and these are not readily included in the models. Such variables are for example the purpose of transport, constraints of battery usage, environmental temperature, tires, road conditions, speed limits, traffic and so on.

## 4. CONCLUSIONS

In this work the concept of planning process for e-axle used in heavy-duty vehicle was presented. The control approach using driving force and vehicle models was introduced. By choosing to use a widely accepted driving force model for

computing the needed energy for driving certain routes, the acceptance of the method is easier to achieve. The driving force model was presented in this study and its results were compared with real measured data. It was noticed that the theoretical driving force model was able to capture the main trend of the engine power demand. This is expected to be enough for the MPC based control strategy.

The data used in this study was from a real truck. Future includes the usage of data collected from a truck in which a prototype of an e-axle is implemented to validate the model. A more distant goal for the future is the implementation of MPC to the prototype truck. As the model will be utilizing online data, the approach is going towards digital twin. This kind of tool can also be seen as a useful asset for several parties of industry branch in question to render the digitalization and green transition.

#### ACKNOWLEDGEMENTS

This work was funded by Business Finland via project ‘NO DAMAGE Truck - Net-effect Optimized Data And Modeling Assisted Green and Efficient Truck’ (1179/31/2022). The data was obtained from Korsu Oy.

#### REFERENCES

- Anselma, P.G. (2022). Computationally efficient evaluation of fuel and electrical energy economy of plug-in hybrid electric vehicles with smooth driving constraints. *Applied Energy* 307, 118247
- AVL Cruise M <https://www.avl.com> › avl-cruise-m
- Banagar, I., Huhtala, T., Könnö, J., and Andwari, A.M.(2024) Configurable Design Approach for Heavy-duty Vehicle Powertrain Design. doi:10.3217/xh0h5-xg862
- Cipek, M. (2020). A novel cascade approach to control variables optimisation for advanced series-parallel hybrid electric vehicle power-train. *Applied Energy* 276 (2020), 115488
- Chu, Q., Sun W., and Zhang Y. (2022). A Data-Driven Method for the Estimation of Truck-State Parameters and Braking Force Distribution. *Sensors* 22(21), 8358.
- Dai, H. (2021). Advanced battery management strategies for a sustainable energy future: Multilayer design concepts and research trends. *Renewable and Sustainable Energy Reviews* 138 (2021), 110480
- Delgado, O., Rodríguez, F., Muncrief, R., Berlin, B. ], Brussels, ], San, Washington, F., Rexeis, M., Williams, P., Dorobantu, M., Laferriere, M., and Boenning, M. (2017). Fuel Efficiency Technology in European Heavy-Duty vehicles: Baseline and Potential For The 2020-2030 Time Frame. [www.theicct.org](http://www.theicct.org)
- Dünnebeil, F., and Keller Heidelberg, H. (2015). Monitoring emission savings from low roll-ing resistance tire labelling and phase-out schemes MRV Blueprint based on an example from the European Union. [www.ifeu.de](http://www.ifeu.de)
- Ehsani, M. (2021). State of the Art and Trends in Electric and Hybrid Electric Vehicles. Vol. 109, No. 6, June 2021| *PROCEEDINGS OF THE IEEE*, 967-984
- Enang, W., and Bannister, C. (2017). Modelling and control of hybrid electric vehicles (A comprehensive review). In *Renewable and Sustainable Energy Reviews* (Vol. 74, pp. 1210–1239). Elsevier Ltd.
- Kinnunen, J. (2023). Concept modeling of energy efficiency for heavy-duty trucks with e-axle equipped trailer. Master thesis. University of Oulu, 97 pp.
- Koch L., Buse, D.S., Wegener, M., Schoenberg, S., Badalian, K., Dressler, F., and Andert, J. (2021). Accurate physics-based modeling of electric vehicle energy consumption in the SUMO traffic microsimulator. 2021 IEEE Intelligent Transportation Systems Conference (ITSC), Indianapolis, USA.
- Lei, Z., Qin, D., Liu, Y., Peng, Z., and Lu, L. (2017). Dynamic energy management for a novel hybrid electric system based on driving pattern recognition. *Applied mathematical Modelling* 45, 940– 954.
- Mazali, I.I. (2022). Review of the Methods to Optimize Power Flow in Electric Vehicle Powertrains for Efficiency and Driving Performance. *Applied sciences*, MDPI, Appl. Sci. 2022, 12, 1735.
- Min, C. (2023). Trajectory optimization of an electric vehicle with minimum energy consumption using inverse dynamics model and servo constraints. *Mechanism and Machine Theory* 181 (2023), 105185.
- Peng, J., He, H., and Xiong, R. (2017). Rule based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming. *Applied Energy* 185, 1633–1643.
- Pereira, G.C. (2023). Adaptive reference aware MPC for lateral control of autonomous vehicles. *Control Engineering Practice* 132 (2023), 105403
- Ruslan, N.A.I. (2023). Modelling and control strategies in path tracking control for autonomous tracked vehicles: A review of state of the art and challenges. *Journal of Terramechanics* 105 (2023), 67–79
- Saju, C. (2022). The implementation of optimal control based on an adaptive neuro-fuzzy inference system that decreases internal combustion engine fuel consumption is the paper’s main contribution. *Sustainable Energy Technologies and Assessments* 52 (2022), 102087
- Schade, W. (unpublished) GHG-TransPoRD Reducing greenhouse-gas emissions of transport beyond 2020: linking R&D, transport policies and reduction targets GHG-TransPoRD Reducing greenhouse-gas emissions of transport beyond 2020: linking R&D, transport policies and reduction targets Title: Aligned R&D and transport policy to meet EU GHG reduction targets Authors: Wolfgang Schade, Michael Krail (with contributions from partners). <http://www.ghg-transpord.eu/>
- Stano P. (2023). Model predictive path tracking control for automated road vehicles: A review. *Annual Reviews in Control* 55 (2023), 194–236
- Wang, Z. (2022) A review of architectures and control strategies of dual-motor coupling powertrain systems for battery electric vehicles. *Renewable and Sustainable Energy Reviews* 162 (2022), 112455
- Zhang, W. (2022). Learning-based supervisory control of dual mode engine-based hybrid electric vehicle with reliance on multivariate trip information. *Energy Conversion and Management* 257 (2022), 115450



# Model Predictive Control for Integrated Photovoltaic (PV) and Electrolyser System<sup>\*</sup>

Ali Reza Pirouzfard,<sup>\*</sup> Sambeet Mishra,<sup>\*</sup> Gaurav Mirlekar,<sup>\*</sup>  
Koteswara Rao Putta<sup>\*\*</sup>

<sup>\*</sup> *Department of Electrical Engineering, Information Technology and Cybernetics, University of South-eastern Norway (USN), Porsgrunn 3918, Norway (e-mail: gaurav.mirlekar@usn.no).*

<sup>\*\*</sup> *GCH2SOL AS, Trondheim 7028, Norway.*

---

**Abstract:** This paper investigates the integration of photovoltaic (PV) systems and proton exchange membrane (PEM) electrolyzers to advance clean energy production and mitigate carbon emissions. The integration of PEM electrolyzers with PV systems presents a promising solution for sustainable hydrogen production. This study utilizes Model Predictive Control (MPC) algorithms to manage the temperature of PEM electrolyzers, crucial for enhancing performance and longevity. Temperature management is vital as lower temperatures increase overpotential, reducing efficiency, while higher temperatures improve performance but can accelerate membrane degradation. The paper simulates the PV-PEM electrolyser system using existing models to identify key parameters affecting system performance, employing MPC for efficient temperature regulation. The methodology involves modeling the integrated PV system, which includes Maximum Power Point Tracking (MPPT) algorithms, a DC-DC converter, and a PEM electrolyser. The MPPT algorithm ensures maximum power output from PV panels under varying irradiance levels, while buck-boost converters regulate voltage to meet electrolyser requirements. The electrolyser model considers mass and energy balance equations to understand the dynamics of hydrogen production and temperature control. Results from simulations indicate that the PV system's power generation is directly influenced by solar radiation and temperature. The study confirms that higher irradiance leads to greater power output, emphasizing the need for feasibility studies. The implementation of MPC algorithms demonstrates effective temperature control, ensuring stable operation and reduced membrane degradation. The integration of PV systems with PEM electrolyzers, coupled with advanced control strategies like MPC, offers a viable pathway for enhancing renewable hydrogen production. This approach addresses the intermittency challenges of renewable energy sources and optimizes system performance.

*Keywords:* Advance process control, Process simulation, Renewable energy systems

---

## 1. INTRODUCTION

The transition to clean energy is critical for reducing fossil fuel dependence and minimizing carbon emissions. The International Energy Agency's annual outlook highlights various scenarios to address these challenges, notably the NetZero Emissions by 2050 pathway, which aims to stabilize global temperatures at 1.5°C and provide universal modern energy access by 2030 World Energy Outlook 2023. Central to this effort are photovoltaic (PV) systems and electrolyzers, with solar PV projected to account for over half of new renewable power capacity by 2030 World Energy Outlook 2023. In 2022 alone, solar PV generation surged by 26% to 1293 TWh, underscoring its pivotal role in decarbonization World Energy Outlook 2023. In addition

to solar energy, hydrogen-based fuels are considered a clean energy source for decarbonization. Hydrogen production methods include both fossil fuels and renewable sources. Conventional methods like steam reforming and coal gasification dominate current production, accounting for about 96% of the total Arsal et al. (2023) Hydrogen forecast to 2050. Renewable-based methods, particularly electrolysis, are gaining traction due to their clean energy potential Arsal et al. (2023). Proton Exchange Membrane (PEM) and Alkaline electrolyzers are the most efficient and commercially available technologies Hydrogen forecast to 2050. Integrating PEM electrolyzers with photovoltaic (PV) systems presents a promising solution for reducing emissions and achieving sustainability. However, electrolyzers currently contribute only 4% to hydrogen production from renewable energies, primarily due to their higher average costs. Forecasts from DNV Hydrogen forecast to 2050 suggest that the average costs of electrolyzers are expected to decrease by 25% by 2030 and by 50% by 2050,

---

<sup>\*</sup> We gratefully acknowledge the financial support from the Department of Electrical Engineering, Information Technology and Cybernetics at the University of South-eastern Norway (USN). corresponding author's e-mail: gaurav.mirlekar@usn.no

based on current market insights. Water electrolysis stands as a leading industrial method for producing nearly pure hydrogen, highlighting its future significance. Moreover, electrolyzers play a pivotal role in converting energy into gas within Power-to-Gas (P2G) systems, which transform renewable energy sources such as wind, solar, geothermal, and hydro into gas. Although currently underutilized, this approach is projected to grow significantly, with hydrogen production via electrolysis expected to reach 22% by 2050 Arsad et al. (2023).

Renewable energy sources, however, face intermittency challenges due to varying climatic conditions. This instability necessitates auxiliary energy systems to ensure consistent hydrogen production, as electrolyzers require a minimum current density for safe operation Afshari et al. (2021). A grid-connected setup, leveraging Maximum Power Point Tracker (MPPT) controllers, can optimize power from PV panels to electrolyzers Gutiérrez-Martín et al. (2024). Enhancing system performance further, buck-boost converters can regulate voltage from PV arrays to meet electrolyser requirements Ruuskanen et al. (2020). In this study, the simulation of PEM electrolyzers is conducted using existing models Abdin et al. (2015), Marangio et al. (2009), Cavaliere (2023), and Falcão and Pinto (2020) to identify and analyze the key parameters affecting the system.

Furthermore, Model Predictive Control (MPC), offer efficient temperature management, critical for enhancing system performance and longevity Scheepers et al. (2021) Majumdar et al. (2023). Simplified control models, like those using piece-wise affine and multi-parametric approaches, have shown promise in minimizing hydrogen production costs and managing operational constraints Flamm et al. (2021) Ogumerem and Pistikopoulos (2020). This paper utilizes an energy balance equation and a state-space model for temperature control in PEM water electrolyzers, adapting MPC techniques for improved efficiency. It has been done by regulating water flow rate in to the system as the manipulating variable to stable the cell temperature and leads lower membrane degradation of the PEM electrolyser.

## 2. PROCESS DESCRIPTION

The objective of this study is to model a PV system, incorporating MPPT algorithms, a DC-DC converter, and a PEM electrolyser, and to apply model predictive control (MPC) algorithms to manage the electrolyser's temperature. As illustrated in Fig.1, electricity generated by photovoltaic panels is regulated by DC-DC converters, primarily buck-boost types, to stabilize the output current, which is then supplied to the PEM electrolyser. This process splits water into hydrogen and oxygen, which are collected in separate containers for drying and further use or storage. MPC are crucial due to the operational conditions of the electrolyzers. PEM electrolyzers typically operate at temperatures between 60-90°C and pressures around 30 bar Arsad et al. (2023). In terms of temperature of the electrolyser, lower temperatures increase overpotential, reducing efficiency, while higher temperatures enhance performance by improving membrane ionic conductivity and reaction kinetics, thus reducing overpotentials Cavaliere (2023).

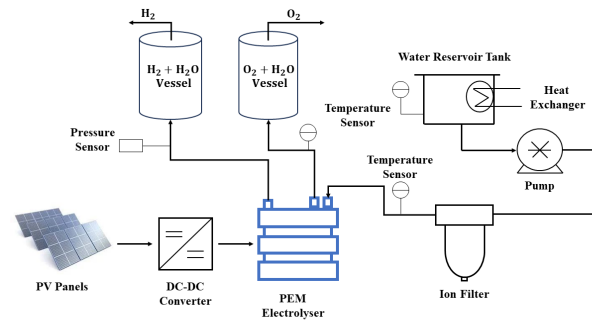


Fig. 1. Process schematic of integrated PV and PEM Electrolyser System.

However, higher temperatures can accelerate the degradation of polymer membranes, primarily due to thermal, chemical, and mechanical stresses, leading to thinning, unzipping, loss of functional groups, or membrane rupture Babic et al. (2017). Monitoring and controlling the temperature is therefore essential to prevent degradation and maintain efficiency.

The increase in electrolyser temperature results from endothermic reactions during water electrolysis and Joule heating, where electric current generates heat as it passes through the conductor Ogumerem and Pistikopoulos (2020). Sudden increases in hydrogen production can raise temperatures, affecting membrane stability and lifespan. Typical methods to maintain the desired temperature range include cooling airflow or adjusting the cooling water flow rate. In the system depicted in Fig.1, a water reservoir, temperature sensor, and heat exchanger keeps the water temperature constant. Water is pumped through an ion filter to reduce resistance and manage flow rates before being directed to the electrolyser. Temperature sensors at the inlet and outlet monitor changes. MPC algorithms, based on Ogumerem and Pistikopoulos (2020), control temperature fluctuations by adjusting the water flow rate, acting as a coolant. The manipulated variable is the water flow rate, and the state variable is the temperature derived from the energy balance equation. The MPC algorithms are based on a linear state-space model derived from solving the energy balance and mass balance equations for each part of the electrolyser, simplifying system modeling.

## 3. METHODOLOGY

Mathematical modeling and governing equation of the integrated system including photovoltaic (PV) system with MPPT, DC-DC converter, PEM electrolyser and MPC algorithms has been presented in this section.

### 3.1 Photovoltaic system

In PV systems solar radiation, containing photons, excites electrons upon contact with PV panels, creating a P-N junction in semiconductors and generating current. The more photons that hit the PV panels, the greater the current produced. Therefore, irradiance information based on the system's location is critical for calculations. Fig. 2 demonstrates, the irradiance data for the University of South-East Norway (USN) in Porsgrunn, Norway (latitude 59.138, longitude 9.672) for June 2020.

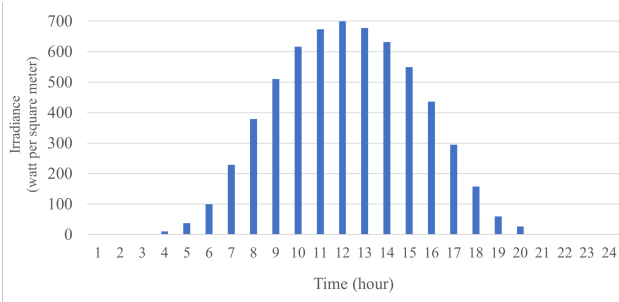


Fig. 2. Hourly solar irradiance  $\frac{W}{m^2}$ .

### 3.2 Modeling of photovoltaic Modules

The current voltage equation for modeling the PV system can be represented as follows:

$$I_{pv} = I_{ph} - I_0 \left[ \exp \left( q \frac{V_{pv} + I_{pv} \cdot R_s}{AKT_j} \right) - 1 \right] \quad (1)$$

$$I_{ph} = I_{ph-G} \cdot (1 + \alpha_{sc} \cdot \Delta T) \quad (2)$$

$$I_{ph-G} = I_{sc} \cdot \left( \frac{G}{G_{ref}} \right) \quad (3)$$

$$\Delta T = T_j - T_{j-ref} \quad (4)$$

In which,  $I_{pv}$  (A) is the photovoltaic current,  $I_{ph}$  (A) is the photo-current,  $I_0$  (A) is the reverse saturation current of the diode,  $q$  is the electron charge which is equal to  $1.602 \times 10^{-19}$  (C),  $K$  is the Boltzmann's constant which is equal to  $1.381 \times 10^{-23}$  ( $\frac{J}{K}$ ),  $A$  is diode ideality factor,  $T_j$  is junction temperature of the panels ( $^{\circ}K$ ),  $R_s$  is series of resistance and  $V_{pv}$  is the voltage across PV cell (V). Also in the Eq.2 to Eq.4,  $I_{ph-G}$  represent current depending on short circuit current, reference irradiance of  $G_{ref} = 1000 \frac{W}{m^2}$  and reference temperature of  $T_{j-ref} = 25^{\circ}C$ . Also  $\alpha_{sc}$  is the temperature coefficient of short-current ( $^{\circ}K$ ) found on the datasheet. For finding reverse saturation current at any temperature ( $I_0$ ), the Eq.5 can be obtained as follow:

$$I_0 = \frac{I_{sc}}{\left[ \exp \left( \frac{V_{oc}}{V_{th} \cdot T_{j-ref}} \right) - 1 \right]} \cdot \left[ \exp \left( \frac{-q \cdot \frac{E_g}{AK}}{\frac{1}{T_j} - \frac{1}{T_{j-ref}}} \right) \right] \cdot \left( \frac{T_j}{T_{j-ref}} \right)^{\frac{3}{A}} \quad (5)$$

In which  $E_g$  represents the band gap energy considered as 1.12 electron volt (eV) Chander et al. (2015). So by substituting Eqs. 2 and 5 in the Eq.1, the final photovoltaic current is obtained. Also in order to find  $R_s$  in the Eq.1, term "-1" added to the exponential equation:

$$R_s = - \frac{dV_{pv}}{dI_{pv}} \Big|_{V_{pv}=V_{oc}} - \frac{1}{W} \quad (6)$$

and:

$$W = q \cdot \frac{I_{sc}}{AKT_j} \quad (7)$$

The term of  $-\frac{dV_{pv}}{dI_{pv}}$ , obtained by experiment or by  $I_{pv} - V_{pv}$  characteristic of the manufacturer mentioned on datasheets. Also, for solving Eq.1, Newton's method has been used.

### 3.3 Maximum Power Point Tracker algorithm

According to Afshari et al. (2021), maintaining minimum current density levels is crucial for safety at varying pressures, as hydrogen concentration decreases relative to oxygen at the anode at lower current densities. To address this,

the Maximum Power Point Tracking (MPPT) algorithm ensures maximum power output from PV panels under different irradiance levels. As defined by Zhou et al. (2010), MPPT, combined with a DC-DC converter, allows a photovoltaic generator to produce optimal power consistently, regardless of changes in irradiance and temperature, by operating at the optimal voltage and current ( $V_{opt}, I_{opt}$ ). Various MPPT algorithms exist, with the Perturb and Observe (P&O) method being used in this study. The P&O method iteratively seeks the maximum power point by continuously evaluating the PV module's current and voltage. The MPPT algorithm based on the P&O method illustrates how the algorithm identifies the maximum power point and corresponding voltage under different hourly irradiance conditions. This algorithm is repeated for each hourly irradiance to maximize the daily power output.

### 3.4 DC-DC converters

The governing equation of the DC-DC converter is presented. Three main types of converters—boost, buck, and buck-boost—are commonly employed in integrated photovoltaic and electrolyser systems to adjust the final voltage. Average modeling methods are typically utilized for simulating these converters. Ruuskanen et al. (2020) highlights that current ripple reduces the efficiency of alkaline electrolyzers, necessitating better power electronics control, potentially applicable to PEM electrolyzers as well. Energy suppliers using photovoltaic systems and batteries require DC-DC converters to adjust voltage and current levels, as these converters are essential for modifying electrical voltage levels between generators and loads Mohan (2012). Simulations show that solar power output and voltage fluctuate throughout the day due to irradiance changes, while electrolyzers require a steady 320V. Buck-Boost converters adjust output voltage above or below the input based on the switch duty ratio  $D$ . The output voltage of a Buck-Boost converter, determined by the duty cycle  $D$ , is given by  $\frac{V_{out}}{V_{in}} = \frac{D}{1-D}$  Hart (2011). With  $D > 0.5$ , it operates as a boost converter; with  $D < 0.5$ , as a buck converter. This hybrid model combines boost behavior when the IGBT is on and buck behavior when off. By implementing average model methodologies, the final model representing the converter is given by the following equations:

$$\frac{di_L}{dt} = \frac{1}{L} (DV_{in} - (1-D)v_C) \quad (8)$$

$$\frac{dv_C}{dt} = \frac{1}{C} \left( (1-D)i_L - \frac{v_C}{R} \right) \quad (9)$$

### 3.5 Modeling of electrolyser system

The electrolyser cell consists of a pair of conductive electrodes immersed in an electrolyte that facilitates ionic conduction. Upon the application of voltage across these electrodes, oxidative processes occur at the anode, while reduction reactions occur at the cathode. The anode and cathode are connected through the flow of current. In this study, PEM electrolyser model is employed for simulation purposes. A detailed explanation of operating characteristics of PEM electrolyzers presented in Table.1. To accurately determine the total voltage of an electrolyzer, it is essential to understand the governing equations derived

from the mass balance in each part of the electrolyzer (Rabascall and Mirlekar (2023)). Accordingly, three control volumes have been defined within the electrolyzer, including the anode, cathode, and membrane sections. The overall water flow across the membrane can be described through three processes: diffusion, electro-osmotic drag, and hydraulic pressure effects representing membrane mass transfer dynamics. Additionally, energy balance equation is chosen from the literature to represent electrolyser temperature (Ogumerem and Pistikopoulos (2020)).

Table 1. Standard Characteristics of Polymer Electrolyte Membrane (PEM).

PEM electrolyser characteristics			
Technology maturity	Commercial	Anode	IrO <sub>2</sub> , RuO <sub>2</sub>
Electrolyte	Polymer (Solid)	Cathode	Pt, Pt-Pd
Cell temperature, °C	60 - 90	System energy consumption, kWh/Nm <sup>3</sup>	4.5 - 7.0
Operating Pressure (bar)	15 - 30	H <sub>2</sub> Capacity (Nm <sup>3</sup> /h)	<40
Cell Voltage (V)	1.8 - 2.2	H <sub>2</sub> purity	99.9
Current Density (A/cm <sup>2</sup> )	0.6 - 2	Stack lifetime, hr	<20,000
Power density (W/cm <sup>2</sup> )	Up to 4.4	System lifetime, yr	10 - 20
Voltage Efficiency (%)	67 - 82	Charge carrier	H <sup>+</sup>

The anode mass transfer module computes the flows of oxygen and water, as well as their respective partial pressures. At the anode, water undergoes oxidation to yield oxygen, electrons, and protons. The cathode mass transfer dynamics module computes the partial pressures and flow rates of hydrogen and water at the cathode, where protons undergo reduction. The objective of these calculations is to determine the partial pressures of each species at both the cathode and anode sides. It is essential to note that the calculated partial pressures are utilized in subsequent calculations to determine the total potential of the electrolyzer and the energy balance equation.

### 3.6 Model predictive control

We utilize a derived model based on the energy balance of the electrolyser, as discussed in previous section. This model captures the temperature dynamics of the electrolyser and can be manipulated by adjusting the water flow rate into the system. In MPC, the process of forecasting future states and outputs, formulating, and solving an optimization problem is iterated at each time step. This iterative approach, known as a sliding horizon strategy, ensures continuous feedback and control adaptation. A linear state space model has been employed for MPC algorithms. A brief description of the linear state space model presented as follow. The general form of linear state space model can be written as,

$$x_{k+1} = Ax_k + Bu_k + v_k \rightarrow \text{State equation} \quad (10)$$

$$y_k = Cx_k + Du_k + w_k \rightarrow \text{Measurement equation} \quad (11)$$

Here,  $A$ ,  $B$ ,  $C$ , and  $D$  represent system matrices.  $A \in R^{n_x \times n_x}$ ,  $B \in R^{n_x \times n_u}$ ,  $C \in R^{n_y \times n_x}$ , and  $D \in R^{n_y \times n_u}$ . Additionally,  $v_k \in R^{n_x}$  and  $w_k \in R^{n_y}$  are zero-mean random variables with specific variances.  $v_k$  represents

process noise, while  $w_k$  denotes measurement noise. It is assumed that  $v_k$  and  $w_k$  are uncorrelated (stochastically independent), i.e.,  $\text{corr}[v_k, w_j] = 0$  for all  $k$  and  $j$ . This implies that random disturbances affecting measurements are unrelated to the randomness in the system states or processes themselves.

For designing linear MPC, the nonlinear derived model is linearized to obtain a continuous time linear state space model, that is, the derivation of energy balance equation on PEM electrolyser has been implemented. It is important to note that the MPC algorithm has been implemented on electrolyser to control operating temperature based on the energy balance equation as follows:

$$c_p M \frac{dT}{dt} = n_{cell} I (V_{oc} - V_0) + \dot{M}_{H_2O, in}^{an} c_{pH_2O} (T - T_{in}^{an}) - H_{rad} \quad (12)$$

Where  $c_p$  is the specific heat capacity (J/kg.K),  $V_0$  is the thermoneutral voltage expressed as a function of temperature anode pressure and  $H_{rad}$  is the the heat loss from radiation.

Finally, MPC is designed using the discrete time linear state space model. The initial step in linearizing a non-linear model involves establishing an equilibrium point to derive a linear model around it. It is assumed that the actual system dynamics approximate the nominal trajectories, that is, they are near the defined operating points. Also the cost function can be used in order to implement the optimization. So:

$$\min_u J = \sum_{k=1}^{N-1} ((y_k - y_k^R)^T Q R (y_k - y_k^R)) + \sum_{k=0}^{M-1} (\Delta u_k^T R \Delta u_k) \quad (13)$$

By understanding the MPC concepts and fundamentals, the implementation of this algorithms on integrated system has been done in Simulink by use of MPC block. While the `mpcDesigner` command used to set the mentioned values for the model. Finally, the MPC block used with the designed plant of electrolyser to set the temperature at the set point value. It is also of the essence to mention that the reference value for the electrolyser operating point considered as 63°C while weighting matrices in the MPC controller considered on the error between the actual altitude and the target to be in the range of 5°C and positive water flow rate based on the pump operation.

## 4. SIMULATION RESULTS

In this section simulation results of an integrated dynamic system, encompassing photovoltaic panels with an MPPT algorithm, a DC-DC converter, a PEM electrolyser, and MPC implementation are presented.

### 4.1 Photovoltaic system simulation results

Power generation is directly influenced by solar radiation as discussed in Sections 2 and 3. The solar irradiance is illustrated in Fig. 2. Additionally, temperature affects the current-voltage produced by the PV system, with a standardized temperature of 25°C assumed for all simulations. Fig. 3 illustrates the current-voltage and power-voltage diagrams at a different radiation levels and a temperature of 25°C, based on MATLAB mathematical simulations. As depicted in Fig. 3, the maximum power output of the PV system is approximately 200 watts, aligning with

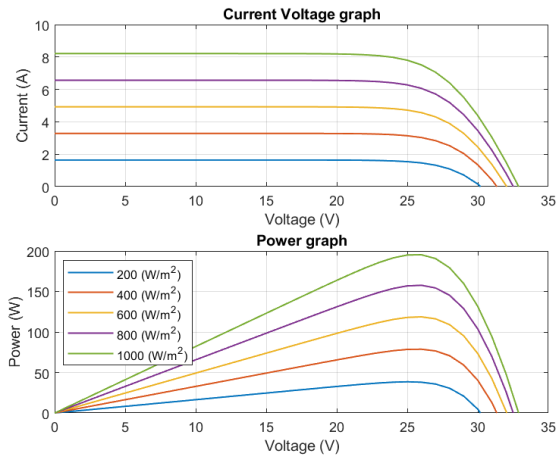


Fig. 3. Current-voltage and power-voltage simulation of photovoltaic panel.

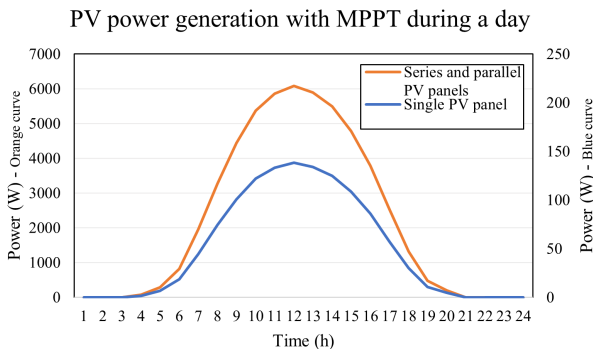


Fig. 4. Simulation of PV system's power during a day.

the findings of Ma et al. (2013) under similar conditions of  $1000 \left(\frac{W}{m^2}\right)$  and a temperature of  $25(^{\circ}C)$ . The behavior of these curves closely mirrors those presented in the cited study. Due to the inherent characteristics of photovoltaic systems, the current-voltage and power-voltage profiles vary with different levels of irradiance. It is typical for manufacturers to display these curves at various standard irradiances such as 200, 400, 600, 800, and  $1000 \left(\frac{W}{m^2}\right)$ . The KC200GT PV module, utilized in the study by Ma et al. (2013), demonstrates how PV characteristics are impacted under various irradiance levels, maintaining a constant temperature of  $25^{\circ}C$ . Fig.3 shows that the behavior of these curves is consistent with those found in PV panel datasheet<sup>1</sup>, further validating the model's accuracy.

As shown in Fig. 3, higher solar irradiance results in greater power output from photovoltaic panels due to increased photon absorption. As this study primarily focuses on the electrolyser and the implementation of MPC algorithms for temperature control, the photovoltaic system is treated solely as a renewable energy source, with simulation and modeling efforts emphasizing accurate representation and validation against existing literature.

From Fig. 3, it is evident that the system yields maximum power output around 25 volts per radiation level. To achieve this maximum power, the Maximum Power Point

<sup>1</sup> <https://www.energymatters.com.au/images/kyocera/KC200GT.pdf>

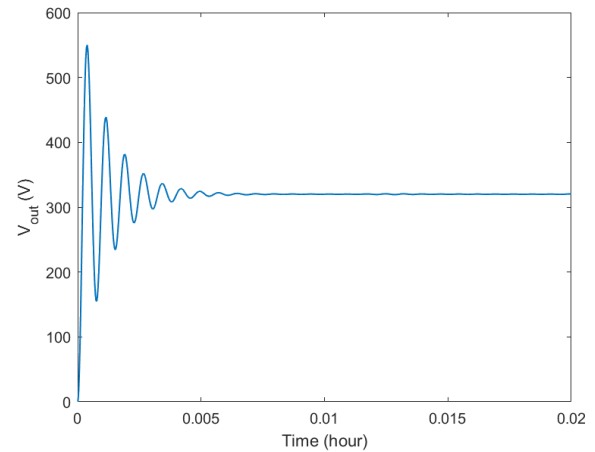


Fig. 5. Simulation of buck-boost converter with output voltage of 320(V).

algorithm is integrated into the PV system, ensuring that the output is optimized for peak power generation. Simulation of the system over 24 hours, based on hourly irradiance data (Fig. 2), reveals a correlation between power production and irradiance fluctuations, as shown in Fig. 4. However, since the location lacks a radiation level of  $1000 \frac{W}{m^2}$ , the system operates below its maximum potential, peaking at approximately 148 W at noon.

The power generated by the PV system needs to be transmitted to the electrolyser to meet the energy demand. This necessitates the use of a DC-DC converter to regulate the fluctuating power output from the PV panels, ensuring stable energy supply to the system.

#### 4.2 DC-DC converter simulation results

The simulations of buck-boost converters are conducted to validate the methodology employed in this study against existing literature. A buck-boost converter, ideal for renewable energy systems with inherent power fluctuations, ensures a stable 320-volt supply to the electrolyser, as demonstrated in simulations based on Eqs. 8 and 9, is depicted in Fig. 5. The design parameters, including inductor, resistor, and capacitor values are calculated accordingly to achieve desired output voltage, operating at the lowest irradiance. This was accomplished with a frequency of 20kHz and a voltage deviation of  $\Delta V_{out} = 0.05 \cdot V_{in}$  to reach the final voltage of 320V. The decision to design the buck-boost converter to achieve an output voltage of 320 volts was driven primarily by the need to align with grid specifications (for the scenarios where can be combined PV system with the grid), particularly the root mean square voltage typical in many grid systems. This voltage level ensures that the power produced by the PV system is compatible with the grid infrastructure, facilitating a seamless integration of the renewable energy source with the existing power network. Ensuring compatibility with the grid voltage is crucial for efficient energy transfer, minimizing losses, and ensuring stability in the power system, which ultimately enhances the reliability and effectiveness of both the renewable integration and the grid operation.

The simulation of a single input voltage from the PV panels is shown in Fig. 5. We now proceed to depict the



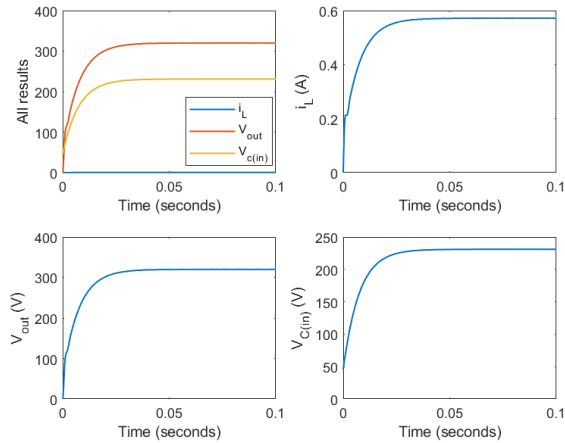


Fig. 6. Simulation of buck-boost converter with an inlet capacitor.

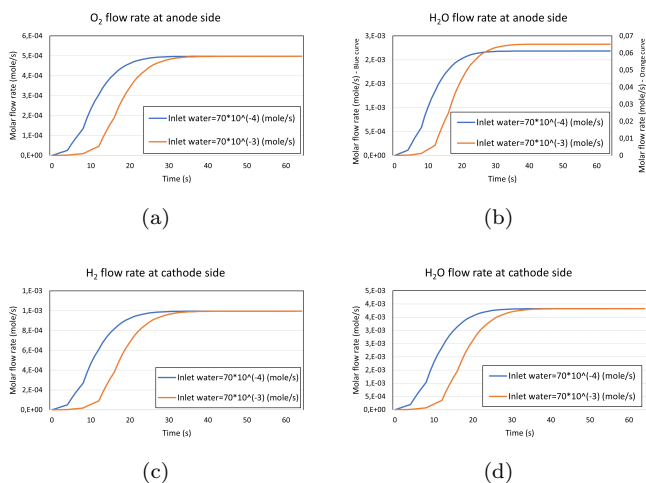


Fig. 7. Hydrogen, oxygen and water produced flow rate from cathode and anode part of electrolyser with different inlet water flow rate.

the comprehensive simulation and modeling of a DC-DC converter, which adjusts various hourly changing voltage arrays from the PV panels to a desired voltage level. It is crucial to highlight that in this mathematical simulation, the duty cycle is adjusted on an hourly basis in response to changes in solar radiation. This adjustment ensures that the output voltage from the PV system aligns with the required 320V for the DC bus. Adjustments in the duty cycle are managed by an appropriate controller within the circuit, as discussed in Safari and Mekhilef (2010). Over-shooting fluctuations can be mitigated by integrating a capacitor at the input of the buck-boost converter circuit. This capacitor functions by storing excess voltage and charging, thus smoothing out the overshoots in the circuit. Xiao (2017) have provided models that demonstrate this effect. Implementation of these models shows that initial fluctuations are effectively eliminated from the system, as illustrated in Fig.6 under the same conditions mentioned in Fig.5.

### 4.3 Electrolyser simulation results

The simulation results of the PEM electrolyser are presented in the following section. Figure 7 presents the simulation modeling results of the different water flow rate into the cathode side of the electrolyser. Notably, increasing the water flow rate does not enhance hydrogen production, which is primarily dependent on the amount of electric current supplied to the electrolyser. According to Ogumerem and Pistikopoulos (2020), increasing the current flow directly boosts hydrogen output. While the water flow rate may not influence hydrogen production, as illustrated in Figure 7, it plays a critical role in cooling the electrolyser. Figure 7 uses water flow rates of  $70 \times 10^{-4}$  and  $70 \times 10^{-3}$  moles per second for comparison. The analysis shows that hydrogen production remains unaffected by changes in the water flow rate but is expected to increase with higher current. Additionally, the study indicates that a water flow rate below  $50 \times 10^{-4}$  mole per second can lead to abnormal outcomes, such as negative water production at the anode with simulated specifications. The system also can be faced with upper limits on the water flow rate due to the capacities of the pump and ion filter components.

Temperature and pressure significantly influence electrolyser performance, making their accurate measurement and control crucial for effective system modeling. To accurately model these systems, it is essential to understand the pressures at the anode and cathode, which influence several critical parameters. Additionally, the differential pressure ( $\Delta P$ ) across the membrane (used for pressure effect based on the Darcy's law) is fundamental for accurately predicting water transport through it. The dependency of the electro-osmotic drag coefficient ( $n_d$ ) on pressure underscores the complexity of electrolyser dynamics under varying operational pressures. The thermal sensitivity of the electrolyser also poses significant challenges. Fluctuations in temperature can compromise membrane integrity and, in extreme cases, could lead to hazardous conditions if hydrogen and oxygen mix explosively. This study, therefore, adheres to operating temperatures for PEM electrolyzers typically between  $60^\circ\text{C}$  and  $90^\circ\text{C}$  with the pressure of 30 bar, aligning with industry standards to minimize risks and optimize performance. The membrane, assumed fully hydrated, exhibits conductivity solely dependent on temperature. The polarisation curve comprises various potentials, including  $V_{oc}$ ,  $V_{act}$ ,  $V_{con}$ , and  $V_{ohm}$ . The contribution of each potential to the polarisation curve is depicted in Fig. 8. Simulations conducted at a symmetric pressure of 1 bar for the cell demonstrate that higher temperatures correlate with reduced overpotential in the electrolyser (Fig. 9(a)). Conversely, as depicted in Fig. 9(b), maintaining a constant temperature of  $60^\circ\text{C}$  while increasing pressure leads to heightened overpotential in the electrolyser. Understanding these dynamics is essential for optimizing electrolyser performance and guiding design decisions.

### 4.4 Model predictive control implementation results

The system description of the PEM electrolyser has been elucidated thus far. Modeling is a crucial precursor to implementing advanced control systems like MPC algorithms, as they rely heavily on accurate simulation models.

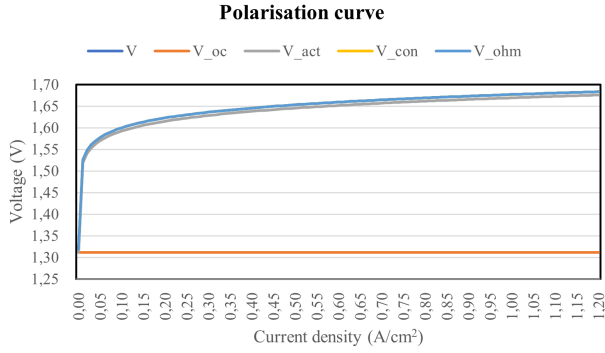


Fig. 8. Share of each potential to the final polarisation curve of the PEM electrolyser.

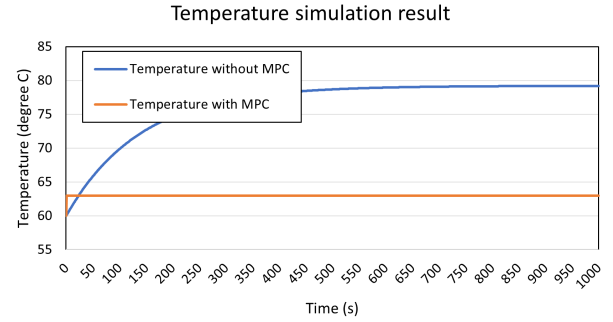
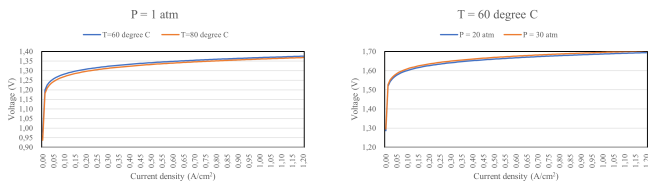


Fig. 10. Simulink result of simulated energy balance model of the PEM electrolyser.



(a) Influence of cell temperature. (b) Influence of cell pressure.

Fig. 9. Influence of cell temperature and pressure on the electrolyser.

Commencing with MPC algorithms, as discussed in Section 3.6, it is crucial to comprehend the model governing the system, derived from the energy balance equation elucidated by Ogumerem and Pistikopoulos (2020). To maintain consistency with the referenced study, a linear state-space form was chosen. To determine matrices  $A$  and  $B$ , it is essential to derive the derivation form of the energy balance formula based on the states and control variables. The state variable considered as the temperature in the Eq.12 while the control variable is the water flow rate into the system by which the cell temperature can be controlled. This involves selecting state, control, and output parameters, with the electrolyser temperature chosen as the state variable to be controlled by the water flow rate. Additionally, the temperature difference from the set point temperature could serve as the output variable. Choosing an operating point around 60°C for the electrolyser facilitates the derivation of  $A$  and  $B$  matrices as follows which will be gain as the derivation form of Eq.12 based on state variable wch is temperature and control variable that is water flow rate effecting in reducing cell temperature with following equation form:

$$A = \left[ \frac{\left( \frac{d(n_{cell}I(V_{oc}-V_0) + \dot{M}_{H_2O,in}^{an} c_p H_2O (T - T_{in}^{an}) - H_{rad})}{dT} \right)}{C_p M_{H_2O}} \right] \quad (14)$$

$$B = \left[ \frac{\left( \frac{d(n_{cell}I(V_{oc}-V_0) + \dot{M}_{H_2O,in}^{an} c_p H_2O (T - T_{in}^{an}) - H_{rad})}{d\dot{M}_{H_2O,in}^{an}} \right)}{C_p M_{H_2O}} \right] \quad (15)$$

To determine the  $A$  matrix, it is essential to define the control operating point. This parameter can be determined by setting the energy balance equation to zero and substituting the set temperature of the operating point.

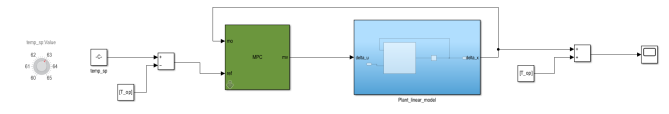


Fig. 11. MPC block implementation in Simulink.

$$\dot{M}_{H_2O,in}^{an} = \frac{1}{c_p H_2O (T_{op} - T_{in}^{an})} (-n_{cell}I(V_{oc} - V_0) + H_{rad}) \quad (16)$$

To simulate the model in state space form, a sampling time of 0.1 seconds was specified, and the system was named 'plant' using the 'ss' function in MATLAB. Additionally, the model was discretized using the 'c2d' function. Following the implementation of the state space model, simulation was conducted using the Simulink environment in MATLAB. Fig. 10 highlights the need for implementing MPC algorithms to regulate the setpoint temperature within the range of 60-65°C. To achieve this, the MPC block was integrated into the Simulink diagram alongside the designed system labeled 'plant.' Fig. 11 illustrates the schematic of the interconnected blocks within the Simulink environment, including the MPC block.

In Fig. 11, the output of the simulation generated by the linear state space model is fed back to the MPC block for comparison with the predetermined reference setpoint. This feedback loop enables the system to dynamically adjust the flow rate to regulate the temperature at the desired value. Expanding upon this concept, the MPC block continuously receives feedback from the simulation results and compares it with the target setpoint temperature of 63°C. By analyzing this feedback, the MPC algorithm calculates the necessary adjustments to the control inputs, ensuring that the system maintains the temperature within the specified range. Moreover, the simulation output, depicted in Fig.10, provides a visual representation of how effectively the MPC algorithm controls the system. It showcases the temperature response over time, demonstrating the system's ability to track and stabilize the temperature around the desired setpoint. Utilizing the 'mpcDesigner' tool in the MATLAB workspace enables users to configure and fine-tune the controller settings effectively. In summary, the simulation results encapsulated in the series of figures, indicate successful temperature regulation by the MPC algorithm. This was achieved through system modeling, controller design, and feedback mecha-

nisms that adjust the flow rate to stabilize temperature. The results validate the robustness and precision of the designed MPC system in achieving and maintaining the desired operational conditions within the electrolyser.

## 5. CONCLUSIONS

The study reviewed extensive literature on hydrogen production, particularly focusing on solar-powered proton exchange membrane (PEM) electrolysers, which were identified as effective but requiring stable operation. Model Predictive Control (MPC) was chosen for its capability to regulate electrolyser temperature, enhancing operational efficiency. Simulations were performed for integrated system, including photovoltaic (PV) panels and PEM electrolysers. MPC was tested, proving successful in maintaining desired electrolyser temperatures. The system incorporated Maximum Power Point Tracking (MPPT) to optimize power output from PV panels, and DC-DC converters were evaluated, with the buck-boost converter providing stable power adjustments suitable for renewable energies. The simulations, which also considered various operational parameters, demonstrated the potential for high efficiency and detailed system behavior under varying conditions. The study achieved successful system modeling and simulation, demonstrating a stable adjustment time that was approximately 10 minutes faster to the selected set point with MPC algorithms compared to without. Additionally, the findings suggest that more detailed models are necessary for practical application and scaling up the project.

## REFERENCES

- Abdin, Z., Webb, C., and Gray, E.M. (2015). Modelling and simulation of a proton exchange membrane (pem) electrolyser cell. *International journal of hydrogen energy*, 40(39), 13243–13257.
- Afshari, E., Khodabakhsh, S., Jahantigh, N., and Toghyani, S. (2021). Performance assessment of gas crossover phenomenon and water transport mechanism in high pressure pem electrolyzer. *International Journal of Hydrogen Energy*, 46(19), 11029–11040.
- Arsad, A., Hannan, M., Al-Shetwi, A.Q., Begum, R., Hosain, M., Ker, P.J., and Mahlia, T.I. (2023). Hydrogen electrolyser technologies and their modelling for sustainable energy production: a comprehensive review and suggestions. *International Journal of Hydrogen Energy*.
- Babic, U., Suermann, M., Büchi, F.N., Gubler, L., and Schmidt, T.J. (2017). Critical review—identifying critical gaps for polymer electrolyte water electrolysis development. *Journal of The Electrochemical Society*, 164(4), F387.
- Cavaliere, P. (2023). *Water Electrolysis for Hydrogen Production*. Springer Nature Switzerland AG, Gewerbestrasse 11, 6330 Cham, Switzerland.
- Chander, S., Purohit, A., Nehra, A., Nehra, S., and Dhaka, M. (2015). A study on spectral response and external quantum efficiency of mono-crystalline silicon solar cell. *International journal of renewable energy research*, 5(1), 41–44.
- Falcão, D. and Pinto, A. (2020). A review on pem electrolyzer modelling: Guidelines for beginners. *Journal of cleaner production*, 261, 121184.
- uchi, F.N., and Lygeros, J. (2021). Electrolyzer modeling and real-time control for optimized production of hydrogen gas. *Applied Energy*, 281, 116031.
- Gutiérrez-Martín, F., Díaz-López, J., Caravaca, A., and Dos Santos-García, A. (2024). Modeling and simulation of integrated solar pv-hydrogen systems. *International Journal of Hydrogen Energy*, 52, 995–1006.
- Hart, D.W. (2011). Power electronics–dc-dc converter. *Valparaiso University, Valparaiso, Indiana*, 2, 198–205.
- Hydrogen forecast to 2050 (2022). Hydrogen forecast to 2050. <https://www.dnv.com/focus-areas/hydrogen/forecast-to-2050.html>. Accessed on 15 Jan 2024.
- Ma, J., Man, K.L., Ting, T., Zhang, N., Guan, S.U., Wong, P.W., et al. (2013). Approximate single-diode photovoltaic model for efficient i-v characteristics estimation. *The Scientific World Journal*, 2013.
- Majumdar, A., Haas, M., Elliot, I., and Nazari, S. (2023). Control and control-oriented modeling of pem water electrolyzers: A review. *International Journal of Hydrogen Energy*.
- Marangio, F., Santarelli, M., and Calì, M. (2009). Theoretical model and experimental analysis of a high pressure pem water electrolyzer for hydrogen production. *International journal of hydrogen energy*, 34(3), 1143–1158.
- Mohan, N. (2012). *Power electronics: a first course*. John Wiley & Sons.
- Ogumerem, G.S. and Pistikopoulos, E.N. (2020). Parametric optimization and control for a smart proton exchange membrane water electrolysis (pemwe) system. *Journal of Process Control*, 91, 37–49.
- Rabascall, J.B. and Mirlekar, G. (2023). Sustainability analysis and simulation of a polymer electrolyte membrane (pem) electrolyser for green hydrogen production. In *Proceedings of the 64th International Conference of Scandinavian Simulation Society, SIMS 2023 Västerås, Sweden, September 25-28, 2023*, 110–117.
- Ruuskanen, V., Koponen, J., Kosonen, A., Niemelä, M., Ahola, J., and Hämmö, M. (2021). and reactive power of water electrolyzers supplied with thyristor converters. *Journal of Power Sources*, 459, 228075.
- Safari, A. and Mekhilef, S. (2010). Simulation and hardware implementation of incremental conductance mppt with direct control method using cuk converter. *IEEE transactions on industrial electronics*, 58(4), 1154–1161.
- uller, M., Carmo, M., and Lehnert, W. (2021). Temperature optimization for improving polymer electrolyte membrane-water electrolysis system efficiency. *Applied Energy*, 283, 116270.
- World Energy Outlook 2023 (2023). World Energy Outlook 2023. <https://www.iea.org/reports/world-energy-outlook-2023>. Accessed on 12 Jan 2024.
- Xiao, W. (2017). *Photovoltaic power system: modeling, design, and control*. John Wiley & Sons.
- Zhou, S., Kang, L., Sun, J., Guo, G., Cheng, B., Cao, B., and Tang, Y. (2010). A novel maximum power point tracking algorithms for stand-alone photovoltaic system. *International journal of control, automation and systems*, 8, 1364–1371.

# Dynamic Reactor Modelling and Operability Analysis of Xylose Dehydration to Furfural Using an Extractive-reaction Process in an Agitated Cell Reactor

Markku Ohenoja\*, Pekka Uusitalo\*, Fernando Russo Abegão\*\*, Abdullahi Adamu\*\*, Kamelia Boodhoo\*\*, Mika Ruusunen\*

\**Environmental and Chemical Engineering Research Unit, Control Engineering Group, Faculty of Technology, P.O.Box 4300, University of Oulu, Oulu, 90014, Finland (e-mail: [markku.ohenoja@oulu.fi](mailto:markku.ohenoja@oulu.fi)).*

\*\* *School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom*

---

**Abstract:** Valorisation hemicellulose into furans chemicals is of great interest to create sustainable furan alternatives to fossil-derived products. A route of particular interest is acid-catalysed dehydration of the hemicellulose pentoses in aqueous medium, with simultaneous extraction of furfural using organic solvent. Agitated Cell Reactor (ACR) could be effectively used to intensify this process and decouple mixing from the long reaction time. This study presents a mathematical model for dehydration of C5 sugars to produce furfural in an ACR. The model can be used to study the effect of feed concentration to the product properties, the concentration profiles along the reactor length, and the dynamic behaviour of the system under feed disturbances or flow rate adjustments. The model was successfully fitted to the experimental data of a laboratory scale ACR for the target product. A simulation study was conducted to analyse the controllability of the process. Operability analysis with the nominal input space and the design space was used for mapping the most feasible region for the process design to meet the flexibility or controllability already at the design phase of the reactor system.

**Keywords:** Simulation, Controllability, Process intensification, Agitated cell reactor, Furfural.

---

## 1. INTRODUCTION

In order to minimise impact on climate change and reduce dependency of non-renewable fossil resources, it is necessary to convert fuels and chemical production to more sustainable synthesis routes using renewable and circular feedstocks, such as biomass. Lignocellulose biomass is the most abundant type of biomass worldwide, being a valuable resource with efficient conversion routes for production of biorenewable chemicals being developed. The biorefining processes being engineered, fractionate lignocellulose into cellulose, hemicellulose and lignin using a range of technologies. The majority of the processes create relatively pure streams of cellulose fibres and/or lignin. However, hemicellulose (HMC), a branched heteropolymer composed mostly of pentoses and some hexose sugars, often ends partially hydrolysed in very diluted aqueous streams, often contaminated with other soluble species such as inorganics, acids and extractives (Wan Azelee et al., 2023). To increase the usage of lignocellulose as a renewable raw material, it is therefore required to be able to convert HMC sugars effectively and overcome the issues associated with diluted streams, and consequent difficulty and cost of separation.

A promising strategy is to dehydrate the HMC sugars into furans using an acid catalyst, where all C5 sugars are converted into furfural (FUR) and all C6 sugars are converted into 5-hydroxymethylfurfural (5-HMF) (Esteban et al., 2020; Ricciardi et al., 2021). This eliminates the need to purify the individual sugars, and creates product convergence, reducing downstream separation requirements. Simultaneously, it produces two high-value platform compounds for fuels and

chemicals, with furfural being recently investigated for production of sustainable aviation fuels, and 5-HMF being used in production of renewable monomers (Mariscal et al., 2016).

Nevertheless, due to the diluted nature of the HMC hydrolysates, it is necessary to process these efficiently in aqueous medium, as water cannot be removed without paying a large energy penalty. This creates challenges since furans are reactive intermediates and, in aqueous medium, can easily undergo self-condensation into humins. Additionally, in the case of 5-HMF, it can also decompose by rehydration into levulinic and formic acid (Esteban et al., 2020; Tong et al., 2010).

To improve product yields, process intensification techniques can be employed. In particular, extractive-reaction technology (Trambouze and Piret, 1960) can significantly enhance the yields as it prevents furan decomposition by carrying out *in situ* extraction of the furan products from the aqueous reaction medium using an immiscible organic solvent (Esteban et al., 2020; Ricciardi et al., 2021).

To ensure the mass transfer of furans between the aqueous and organic solvent media is efficient, good mixing is required. Nevertheless, the required reaction times are in the order of several minutes to a couple of hours, and tubular reactors are not able to offer good mixing for such slow flow conditions. Continuous stirred tank reactors (CSTRs) can offer good mixing at longer residence times, but typically operate at low concentrations of reactants and high concentrations of products, with such conditions promoting furan degradation. It

is therefore important to select appropriate reactor technology capable of decoupling mixing from reaction time. A good compromise between good mixing and long residence times is achieved by using a cascade of CSTRs, where the reaction progresses fractionally in each tank, and where there is good mixing in each stage of the cascade. For small scale production and process development, setting a cascade of CSTRs can be costly due to the number of reactors and agitators required, and difficult to control. Alternatively, mixing intensification and decoupling between mixing and residence time can be achieved using an Agitated Cell Reactor (ACR).

The ACR is suitable for both homogeneous and multiphase reactions. In (Toftgaard Pedersen et al., 2017), several applications are listed, where the ACR has been used including continuous processing of slurries (Browne et al., 2011), hydrodechlorination of organic waste (Gómez-Quero et al., 2011), functionalization of carbon nanotubes (Salice et al., 2012), and biocatalytic oxidation reactions (Gasparini et al., 2012; Jones et al., 2012). Specific to the furan production in this work, the ACR is expected to intensify mixing and increase the liquid-liquid interfacial area. Consequently, this will enhance mass transfer rates, allowing the reduction of the amount of extraction solvent used, and resulting in small reaction vessels and lower separation costs downstream. This will also help to control the reaction system selectivity.

To achieve a holistic view on process intensification and a fully integrated process with optimal performance, it is also necessary to account for the process operability, controllability and safety during the conceptual process design (Tian et al., 2018). The operability analysis was casted into process intensification in (Carrasco and Lima, 2017). An operability-based algorithm was presented and used for input–output analysis of nonlinear systems comprising the process intensification goals in the calculation. Furthermore, an optimisation-based approach was applied to determine the feasible operating envelope (desired input set satisfying the design constraints and the desired output set).

This study aims to develop the necessary mathematical tools for optimizing the design of a novel, intensified reactor while taking into account the process operability. Therefore, a dynamic model was developed to simulate the response of the ACR during dehydration of xylose, the most representative sugar in HMC hydrolysate streams produced in biorefining processes. A validation of the model was performed for experimental data representing different feed properties and reactor conditions. The model was then used to analyse the operability of the ACR system for this biorefining application to gain insight on the effect of foreseen process disturbances to the process performance, and thus to support the process design already at the conceptual design phase.

The rest of the article is structured as follows: Section 2 presents the laboratory-scale ACR reactor and the developed dynamic model, as well as the applied methodology for the operability analysis. In Section 3, the model identification and validation are performed. Section 4 presents the case for the operability study, followed by conclusions in Section 5.

## 2. MATERIALS AND METHODS

### 2.1 System description

ACRs make use of a reactor block with multiple cylindrical horizontal cells in series, with each cell containing a cylindrical agitator positioned loosely inside. The entire reactor block is then oscillated horizontally, causing the agitators to move freely inside the cells, rolling back and forth and disrupting the flow through each cell. Each cell in the reactor then replicates a CSTR, with the entire reactor block constituting the cascade. Different agitator sizes with different material densities and cylindrical geometries (e.g. solid, hollow tube, coil) can be chosen depending on what is more beneficial for each application.

An ACR with a total volume of 100 mL (10 mL per cell  $\times$  10 cells) is modelled. The reactor is fed with an aqueous mixture of xylose and sulphuric acid catalyst and pure MIBK (methyl isobutyl ketone) as the organic solvent for extractive phase.

Five components were considered in the models: xylose ( $Xyl$ ), furfural in aqueous phase ( $Faq$ ), furfural in organic phase ( $Forg$ ), degradation products ( $DG$ ), and intermediate products ( $Int$ ). It is assumed that only furfural would transfer between the aqueous and organic phase, while other components remained in the aqueous phase. The degradation products describe a variety of organic compounds, particularly humins, formed during self-condensation reactions of the xylose, intermediates and furan products. Finally, the intermediate products can be various carbohydrate-based compounds resulting from partial reaction of xylose, and which react further to produce furfural or degradation products.

At the initial state, it is assumed that the entire reactor volume is filled with reactants and MIBK solvent at a volumetric ratio of 1:1, and the operation pressure is achieved. The temperature at which reactions starts is assumed to be 90 °C at  $t = 0$  min, from which point the temperature increases until it reaches the final operation temperature.

The agitation frequency of the ACR is assumed to be fixed at 5 Hz. Thus, the mixing intensity and its effects to the mass transfer are omitted in modelling.

### 2.2 Model structure

The model structure here consists of a cascade of continuous stirred tank reactors (CSTRs) connected in series with forward flow and an optional backflow term (Roemer and Durbin, 1967). A similar approach for ACR modelling has been proposed in (Toftgaard Pedersen et al., 2017).

The component balances were defined for the mentioned components  $C_{i,Xyl}$ ,  $C_{i,Faq}$ ,  $C_{i,Forg}$ ,  $C_{i,Deg}$  and  $C_{i,Int}$  in each reactor  $i$  resulting in a set of ordinary differential equations (ODEs), represented by (1).

$$\frac{dC_{i,j}}{dt} = r_{i,j} + \frac{v_{aq}}{V_{aq}}(C_{i-1,j} - C_{i,j}) + \frac{v_B}{V_{aq}}(C_{i+1,j} + C_{i-1,j} - 2C_{i,j}) \quad (1)$$



In (1), the  $v_{aq}$  is the flow rate of aqueous phase,  $V_{aq}$  is the volume of aqueous phase in reactor cell,  $v_B$  is the backflow rate, and  $r_{i,j}$  is the formation or consumption of component  $j$  due to reactions or mass transfer between phases. There is no backflow stream leaving from the first unit, thus the term  $2C_{i,j} = 0$  for  $i = 1$ . Similarly, with the last unit, there is no incoming backflow term from the next unit and the backflow term is  $(C_{i-1,j} - C_{i,j})$ . For furfural in the organic phase, the flow term in the balance equation uses the flow rate and volume of organic phase ( $v_{org}$  and  $V_{org}$ , respectively), instead of  $v_{aq}$  and  $V_{aq}$ . The model was implemented to MATLAB® and a variable-step, variable-order method ‘ode15s’ was used to solve the ODEs.

### 2.3 Kinetics for xylose dehydration to furfural

The reaction mechanism is assumed to follow the pathways presented by (Jakob et al., 2022). The studied reaction pathway can be seen in Fig. 1. The reaction rates are first order (2–6) and follow an Arrhenius temperature dependence (7). The acid catalyst concentration ( $C_{cat}$ ) is accounted by multiplying the intrinsic reaction rates ( $k_{x0}$ ) with the ion concentration of  $H_2SO_4$  catalyst (8). The temperature dependency of the dissociation constant  $K_A$  and the ion concentration  $C_{H^+}$  are calculated as in (9) and (10), respectively (Guo et al., 2021).

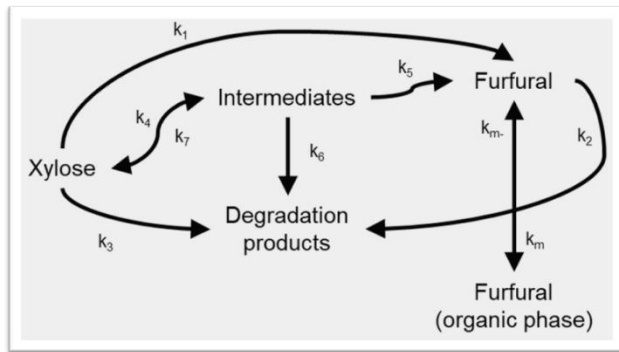


Fig. 1. Reaction mechanism.

In Fig. 1, and Figs. 2 to 6, the  $k_x$  are the apparent rate constants,  $k_m$  is the mass transfer rate constant for furfural from aqueous phase to organic phase, and  $k_{m-}$  is the mass transfer rate constant for furfural from organic phase to aqueous phase.  $T$  is the operation temperature and  $T_{ref}$  the reference temperature,  $E_{a,x0}$  the activation energy and  $R$  is the ideal gas constant.

$$r_{i.Xyl} = -k_1 C_{i.Xyl} - k_3 C_{i.Xyl} - k_4 C_{i.Xyl} + k_7 C_{i.Int} \quad (2)$$

$$r_{i.Faq} = k_1 C_{i.Xyl} - k_2 C_{i.Faq} + k_5 C_{i.Int} - k_m C_{i.Faq} + k_{m-} C_{i.Forg} \quad (3)$$

$$r_{i.Forg} = k_m C_{i.Faq} - k_{m-} C_{i.Forg} \quad (4)$$

$$r_{i.Deg} = k_3 C_{i.Xyl} + k_2 C_{i.Faq} + k_6 C_{i.Int} \quad (5)$$

$$r_{i.Int} = k_4 C_{i.Xyl} - k_7 C_{i.Int} - k_5 C_{i.Int} - k_6 C_{i.Int} \quad (6)$$

$$k_{x0}(T) = k_{x0}(T_{ref}) \exp \frac{E_{a,x0}}{R} \left( \frac{1}{T_{ref}} - \frac{1}{T} \right) \quad (7)$$

$$k_x = C_{H^+} \times k_{x0} \quad (8)$$

$$K_A = \exp(0.0152 \times T + 2.636) \quad (9)$$

$$C_{H^+} = C_{cat} + \frac{1}{2} \left( -K_A - C_{cat} + \sqrt{(K_A + C_{cat})^2 + 4C_{cat}K_A} \right) \quad (10)$$

The rate constants and activation energies are given in (Jakob et al., 2024). In addition, the mass transfer coefficient value and partition coefficient value were based on (Weingarten et al., 2010).

### 2.4 Intensification factor and temperature profile

In order to simulate the intensified reactor with improved mixing, and thus mass transfer, an intensification factor  $F$  was applied to the model. For simplicity, a single lumped value was assumed that multiplies the rate constants and the forward mass transfer constant, as in (11). The value for  $F$  was estimated using the experimental data.

In addition to the intensification factor, the model fitting comprises an unknown, transient temperature profile for the reactor start-up. The temperature increment to the target operation temperature was assumed to follow an exponential function seen in (12). The value for time constant  $h$  was estimated using experimental data.

$$k_{intensified} = k_x \times F \quad (11)$$

$$T(t) = T(0) - (T(t) - T(0)) \exp(-h \times t) \quad (12)$$

### 2.5 Operability analysis

Process operability analysis focuses on evaluating in what extent the desired process outputs (controlled variables, desired output space) can be achieved with the assumed range of input variables (manipulated variables, available input space) and expected disturbances. The analysis does not assume any specific control structure. The same analysis can also be conducted to evaluate whether the process design satisfy the desired intensification targets and constraints (Carrasco and Lima, 2017).

An open-source MATLAB® tool ‘Operability App’ was applied for the operability study in this work. The operability index ( $OI$ ) for a design region is used to evaluate the feasible process design in terms of operability.  $OI$  calculation uses the desired outputs set ( $DOS$ ) and achievable output set ( $AOS$ ) divided into subregions and applies computational geometry tools to determine the ratio of achieved  $DOS$  subregions and the total number of  $DOS$  subregions (13).

$$OI = \frac{\mu(AOS \cap DOS)}{\mu(DOS)} \quad (13)$$

$AOS$  is the set of controlled variables that the process design can achieve for the available input set (operational and design variable space) and disturbance set. Process model is used to

describe the process behaviour and, thus, to define *AOS*. The applicability of the software tool for process achievability analysis, intensification, and modularization has been demonstrated in (Gazzaneo et al., 2020) and (Gazzaneo, 2021), where also further details of the method can be found.

### 3. MODEL IDENTIFICATION AND VALIDATION

#### 3.1 Model fitting

The model performance was qualitatively evaluated against the experimental data acquired from a synthetic feed stream with xylose monosaccharides as a starting compound. The feed concentration was 42.3 g/L, operation temperature at 125 °C, and catalyst concentration 0.2 M. The volumetric ratio of the phases was 1:1. The feed rates of both phases are calculated based on the target residence time of aqueous phase in the reactor. The nominal residence time is 120 min, corresponding a feed rate of 0.42 mL/min of both phases.

The model was simulated with different values of  $F$  and  $h$ . Figure 2 presents the fitted concentration values at the reactor exit with  $F = 12$  and  $h = 0.015 \text{ min}^{-1}$ . It can be seen that the xylose consumption and total furfural formation are in a good agreement between the model and experimental data. The experimental results for furfural in aqueous phase show very small concentrations (<0.2 g/L). However, the simulation shows higher concentration of furfural in aqueous phase comparatively to the experiment. This implies that the mass transfer of furfural from aqueous phase to organic phase is underestimated in the model.

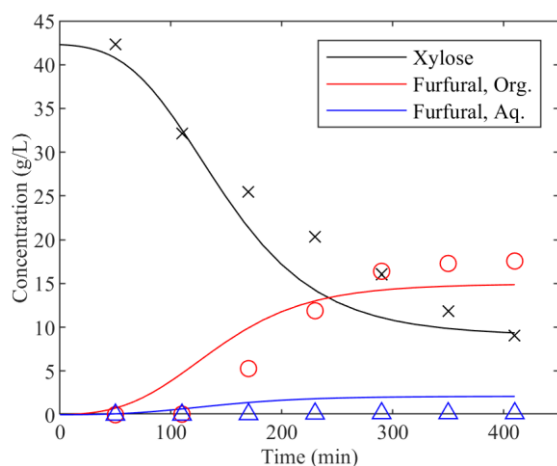


Fig. 2. Evolution of concentration as a function of time. Lines represent the simulation. Tick marks represent the experimental data.

#### 3.2 Transfer to real hemicellulose streams

The model input parameters ( $C_{1,xy1}$ ,  $T$ ,  $v_{aq}$ ,  $v_{org}$ ,  $C_{cat}$ ) were then changed to represent three other experimental cases: one with the same synthetic stream but a different residence time, and two experimental campaigns with real HMC streams.

In Table 1, the error between the predicted and experimental total furfural concentrations (organic + aqueous) at steady-state for the four experimental data sets are shown, when the simulations were run with fixed values of  $F$  and  $h$ . As

expected, the error is higher for the three cases, where the model was extrapolated to. For the synthetic stream with lower residence time ( $RT$ ), the model overestimates the furfural concentration. For the first real HMC stream, the direction of the modelling error is the same. However, for the second real HMC stream, the model slightly underestimated the furfural concentration. Taking into account the complexity of the modelled system and simplicity of the fitting approaches, the proposed modelling approach was effective and provided acceptable results with respect to the target compound furfural.

Table 1. Error in steady-state furfural concentration.

Case	Absolute error	Relative error
Synthetic, $RT=120$ min	-0.77 g/L	-4.3 %
Synthetic, $RT=60$ min	2.75 g/L	36.8 %
Real HMC #1	2.22 g/L	25.2 %
Real HMC #2	-1.46 g/L	-10.6 %

#### 3.3 Implications

Sugar dehydration reactions are complex networks (Dussan et al., 2015), making analysis of hemicellulose streams dehydration challenging due to the wide range of sugars present. Moreover, it is often difficult to find available kinetic data for all sugars using the same catalyst, reactor type (which can mask mixing, mass transfer and other limitations) and process conditions. Nevertheless, the high potential for valorisation of these streams, and the appeal of using dehydration reactions as a route to product convergence and reduction of downstream processing requirements, highlights the importance of developing the understanding of these systems and be able to simulate them in simple and effective ways.

The models developed demonstrate that transferability of kinetic data between batch and intensified flow reactor systems is possible by employing a simple Intensification Factor, whilst still retaining the overall system behaviour, and providing practical tools to model the system dynamic behaviour of intensified ACRs. This is expected to notably speed up modelling, design and preliminary optimisation calculations, contributing to the overall sustainability of the biorefinery.

When looking at the concentration values at the reactor exit for the real HMC stream #1, depicted in Fig. 3, the modelled predictions for xylose concentration deviate remarkably from the experimental observations. The model suggests that the concentration of monosaccharide xylose should evolve to substantially lower values than what was observed in the experiments. There is indeed a natural explanation to this model-data mismatch: the real HMC streams can contain a significant amount of oligosaccharides that undergo hydrolysis to the corresponding monosaccharides at similar reaction conditions. Therefore, while some of the xylose is converted, some more monomeric xylose is formed. Thus, for HMC hydrolysate stream where the amount of oligomers is

significant, oligomers hydrolysis should also be accounted for in future model developments, incorporating literature hydrolysis kinetic models and validation against experimentation which includes analysis of oligomers concentration over time.

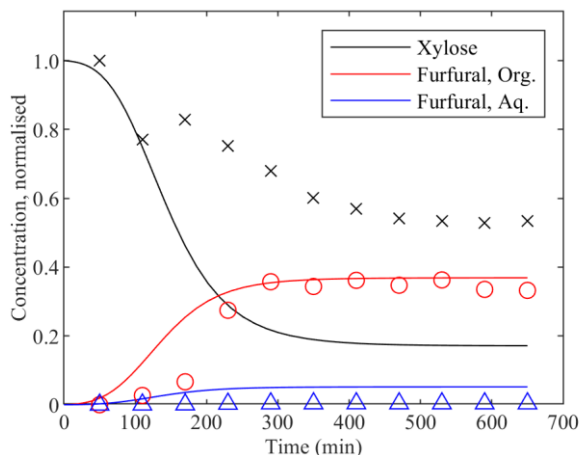


Fig. 3. Concentration (normalised) as a function of time for real HMC stream. Lines represent the simulation. Tick marks represent the experimental data.

## 4. OPERABILITY STUDY

### 4.1 Simulation scenario

Due to the heterogeneity of the biomass properties, and the disturbances in upstream processes, the concentration of xylose in the feed stream has natural variability. In the operability study, a simulation scenario is built, where different transients take place during the continuous processing. Namely, it is assumed that the feed concentration is subject to a sudden +25% change after the first steady-state is reached. After that, a second change takes place with -25% xylose concentration with respect to the nominal value. Again, after the steady-state, a third change occur with the feed concentration returning its nominal value.

The desired output space is determined by two process performance variables dependent on the dynamic performance of the system; The cumulative amount furfural production ( $y_1$ ) and the cumulative amount of degradation and intermediate products ( $y_2$ ) during the simulated period. The rationale for the first is obvious, as furfural is the target product. For the second output variable, production should be minimized in order to avoid excessive carbon (e.g. humins) formation that can possibly lead to reactor fouling. The ranges are given in Table 2.

The available input space is also presented in Table 2. For the design variables, the reactor size ( $u_{des,1}$ ) and the catalyst loading ( $u_{des,2}$ ) are used as available inputs. The volume of ACR reactor can be altered by changing the mixing elements. The operational inputs are assumed to be the aqueous phase feed flow rate ( $u_{op,1}$ ) and the feed ratio of aqueous and organic phases ( $u_{op,2}$ ). Higher value for the feed ratio corresponds to larger volume of reacting phase (aqueous) in the reactor.

Table 2. Variable ranges in the operability study.

	Nominal	Minimum	Maximum
$u_{des,1}$ (mL)	100	60	100
$u_{des,2}$ (mol/L)	0.2	0.1	0.3
$u_{op,1}$ (mL/min)	0.42	0.30	0.60
$u_{op,2}$ (-)	0.5	0.4	0.6
$y_1$ (kmol)		6	10
$y_2$ (kmol)		0	2

### 4.2 Results

For the nominal process design, Figure 4 presents the furfural production rate at the reactor exit as a function of time, when a 25% stepwise increment in xylose feed concentration takes place at  $t=0$ . It can be seen that the due to the long residence time, the new steady-state is achieved after a period of 216 min (2% settling time). Naturally, this prevents establishing an efficient feedback control of ACR system with the studied reactions and a careful design of the system is needed to establish feedforward control for feed disturbances entering the system.

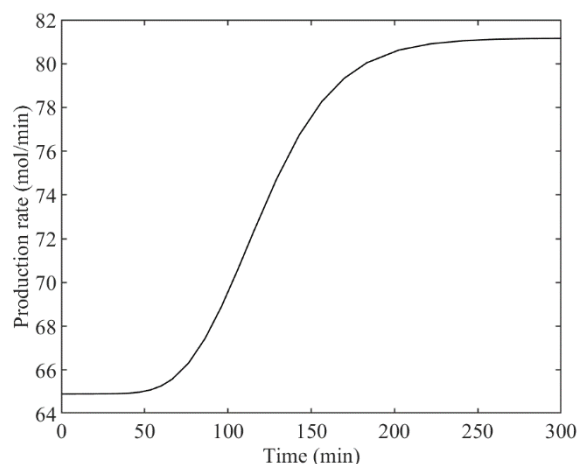


Fig. 4. Simulated dynamic response of the furfural production in nominal conditions to +25% increment of xylose feed at 0 min.

$OI$  for a design region was evaluated for nine different process design ( $u_{des,1}$ ,  $u_{des,2}$ ), as presented in Table 3.  $OI$  is a scalar number describing the proportion of subregions in the desired output space that can be achieved with the available operational inputs in each process design. From Table 3, it can be seen that the process design with reactor size ( $u_{des,1}$ ) of 80 mL and catalyst concentration of 0.2 mol/L shows the best operability ( $OI = 67\%$ ).

Each design can also be interpreted with the achievable output space, as depicted in Fig. 5. There, the desired output space is given as the light grey shaded area divided into nine subregions. Three examples of process designs are given and illustrated as the dark grey polyhedrons. Design 4 shows a

narrow range of obtained outputs, which all lie outside the desired output space. Thus, also  $OI$  is equal to zero (see Table 3). This is due to the combination of limited residence time and low catalyst concentration, contributing to moderate production of both the target furfural and the degradation products.

Table 3. Operability indices for the different designs.

Design	$u_{des,1}$	$u_{des,2}$	$OI$
1	60	0.1	0
2	60	0.2	22
3	60	0.3	56
4	80	0.1	0
5	80	0.2	67
6	80	0.3	33
7	100	0.1	11
8	100	0.2	33
9	100	0.3	33

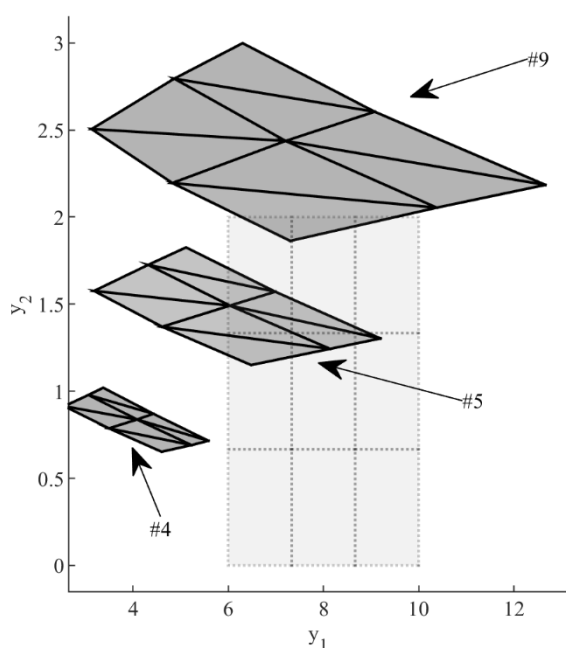


Fig. 5. Achievable output space (represented by the dark grey areas) for designs 4, 5 and 9. The desired output space is given in light grey colour.

Design 9 suggests that the system can achieve cumulative furfural production ( $y_1$ ) under feed disturbances up to 12.7 kmol. However, this design also contributes to substantial degradation product formation, thus most of the operational space lying outside the desired region. According to Table 3, the design 9 can achieve 33% of the design space, but Figure 5 reveals that only small part of those design regions can be

reached with this input space. Finally, design 5 seems to have the highest  $OI$  (67%) as it can reach at least a small portion of six different regions in the desired output space. This process design shows the best performance in terms of operability.

#### 4.3 Implications

In comparison, a different process design optimisation target could be based on the xylose conversion and furfural yield (to the organic phase) in steady-state conditions. By looking at the simulation results from the operability study, the maximum conversion (96.9%) would be attained for the process design with maximal residence time, i.e.  $u_{des,1} = 100$  mL and  $u_{op,1} = 0.3$  mL/min, with  $u_{des,2} = 0.3$  and  $u_{op,2} = 0.6$ . Also, yield is preferred by long residence time, with maximum (60.0%) seen in  $u_{des,1} = 100$  mL,  $u_{op,1} = 0.45$  mL/min,  $u_{des,2} = 0.3$  and  $u_{op,2} = 0.6$ .

However, with respect to  $y_1$  and  $y_2$ , these designs and operation conditions would provide dynamic behaviour ( $y_1$ : 3.13 and 4.87 kmol) and excessive degradation products formation ( $y_2$ : 2.51 and 2.79 kmol) outside the expected output range of the operability study. This example highlights the need of dynamic analysis and incorporation of dynamic performance as a part of process design, especially in processes where considerable feed disturbances cannot be avoided, or the process dynamics are slow mitigating the feedback control possibilities.

From the operability analysis results, it can be observed that the desired output space is somewhat optimistic, and most of the process states cannot be achieved by changing the selected design and operational variables. The degradation product formation and furfural production are affected to the same direction by the different inputs, and thus the degradation product formation cannot be pushed down without negative effect on the furfural production.

The operability could be improved by allowing larger ranges for the operational inputs. However, as both inputs affect the residence time of reaction phase, the extension of ranges can have a competing effect to the residence time. There are also limits with the ratio ( $y_2$ ) as both phases need to flow through the reactor, and the high amount of organic phase would result in higher capacity demand for the downstream separation steps of furfural and MIBK.

As a future work, the model should be developed to comprise also the temperature balance. This would allow to simulate the operation temperature as an operational input and use in control design to compensate the feed variations. Temperature balance would also allow more accurate predictions of concentrations along the reactor length as the temperature gradient affect to the reactions rates, ion concentration, and mass transfer coefficients. Addition of catalyst mass balance would also establish using the catalyst concentration as an operational variable instead of (fixed) design variable.

Finally, the operability analysis could also involve the ACR mixing intensity as an input. This has been modelled for example in (Miller, 2021), where its effect to the mass transfer was studied. The mixing intensity can then be linked to the furfural mass transfer between the phases, and to the

intensification factor coefficient. However, additional experimental work would be needed to identify these interactions.

## 5. CONCLUSIONS

Mathematical model aiming to describe the xylose dehydration to furfural using an extractive-reaction process in an agitated cell reactor was presented. A lumped intensification factor was found useful to transfer the kinetic data between batch and intensified flow reactor systems. The proposed modelling approach was effective and provided acceptable validation to experimental data. Operability of the ACR system was studied with the model. The demonstration highlighted how most of the process states cannot be achieved by changing the selected design and operational variables within their ranges. Instead, the desired output space needs to be reconsidered. This implies that the operability measures should be seen as inputs already at the conceptual process design.

## ACKNOWLEDGEMENTS

This project has received funding from the Bio-based Industries Joint Undertaking (JU) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 887226. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Bio-based Industries Consortium.

## REFERENCES

- Browne, D. L., Deadman, B. J., Ashe, R., Baxendale, I. R., and Ley, S. V. (2011). Continuous Flow Processing of Slurries: Evaluation of an Agitated Cell Reactor. *Organic Process Research & Development*, 15(3), 693–697. doi: 10.1021/op2000223
- Carrasco, J. C., and Lima, F. V. (2017). Novel operability-based approach for process design and intensification: Application to a membrane reactor for direct methane aromatization. *AIChE Journal*, 63(3), 975–983. doi: 10.1002/aic.15439
- Dussan, K., Girisuta, B., Lopes, M., Leahy, J. J., and Hayes, M. H. B. (2015). Conversion of Hemicellulose Sugars Catalyzed by Formic Acid: Kinetics of the Dehydration of D-Xylose, L-Arabinose, and D-Glucose. *ChemSusChem*, 8(8), 1411–1428. doi: 10.1002/cssc.201403328
- Esteban, J., Vorholt, A. J., and Leitner, W. (2020). An overview of the biphasic dehydration of sugars to 5-hydroxymethylfurfural and furfural: A rational selection of solvents using COSMO-RS and selection guides. *Green Chemistry*, 22(7), 2097–2128. doi: 10.1039/C9GC04208C
- Gasparini, G., Archer, I., Jones, E., and Ashe, R. (2012). Scaling Up Biocatalysis Reactions in Flow Reactors. *Organic Process Research & Development*, 16(5), 1013–1016. doi: 10.1021/op2003612
- Gazzaneo, V. (2021). Multimodel Operability Framework for Design of Modular and Intensified Energy Systems. *Graduate Theses, Dissertations, and Problem Reports*, West Virginia University, 8069. doi: 10.33915/etd.8069
- Gazzaneo, V., Carrasco, J. C., Vinson, D. R., and Lima, F. V. (2020). Process Operability Algorithms: Past, Present, and Future Developments. *Industrial & Engineering Chemistry Research*, 59(6), 2457–2470. doi: 10.1021/acs.iecr.9b05181
- Gómez-Quero, S., Cárdenas-Lizana, F., and Keane, M. A. (2011). Liquid phase catalytic hydrodechlorination of 2,4-dichlorophenol over Pd/Al<sub>2</sub>O<sub>3</sub>: Batch vs. continuous operation. *Chemical Engineering Journal*, 166(3), 1044–1051. doi: 10.1016/j.cej.2010.07.032
- Guo, W., Zhang, Z., Hacking, J., Heeres, H. J., and Yue, J. (2021). Selective fructose dehydration to 5-hydroxymethylfurfural from a fructose-glucose mixture over a sulfuric acid catalyst in a biphasic system: Experimental study and kinetic modelling. *Chemical Engineering Journal*, 409, 128182. doi: 10.1016/j.cej.2020.128182
- Jakob, A., Likozar, B., and Grilc, M. (2022). Aqueous conversion of monosaccharides to furans: Were we wrong all along to use catalysts? *Green Chemistry*, 24(21), 8523–8537. doi: 10.1039/D2GC02736D
- Jakob, A., Likozar, B., and Grilc, M. (2024). Model-assisted optimization of xylose, arabinose, glucose, mannose, galactose and real hemicellulose streams dehydration to (hydroxymethyl)furfural and levulinic acid. *ChemSusChem*, Accepted. doi: 10.1002/cssc.202400962
- Jones, E., McClean, K., Housden, S., Gasparini, G., and Archer, I. (2012). Biocatalytic oxidase: Batch to continuous. *Chemical Engineering Research and Design*, 90(6), 726–731. doi: 10.1016/j.cherd.2012.01.018
- Mariscal, R., Maireles-Torres, P., Ojeda, M., Sádaba, I., and Granados, M. L. (2016). Furfural: A renewable and versatile platform molecule for the synthesis of chemicals and fuels. *Energy & Environmental Science*, 9(4), 1144–1189. doi: 10.1039/C5EE02666K
- Miller, J. M. J. (2021). *Solvent Extraction and Mass Transfer Assessment in Novel Extraction Technologies*, Phd thesis, University of Leeds. <https://etheses.whiterose.ac.uk/29223/>
- Ricciardi, L., Verboom, W., Lange, J.-P., and Huskens, J. (2021). Production of furans from C5 and C6 sugars in the presence of polar organic solvents. *Sustainable Energy & Fuels*, 6(1), 11–28. doi: 10.1039/D1SE01572A
- Roemer, M. H., and Durbin, L. D. (1967). Transient Response and Moments Analysis of Backflow Cell Model for Flow Systems with Longitudinal Mixing. *Industrial &*



- Engineering Chemistry Fundamentals*, 6(1), 120–129. doi: 10.1021/i160021a021
- Salice, P., Fenaroli, D., De Filippo, C., Menna, E., Gasparini, G., and Maggini, M. (2012). Efficient functionalization of carbon nanotubes: An opportunity enabled by flow chemistry. *Chimica Oggi*, 30, 37–39.
- Tian, Y., Demirel, S. E., Hasan, M. M. F., and Pistikopoulos, E. N. (2018). An overview of process systems engineering approaches for process intensification: State of the art. *Chemical Engineering and Processing - Process Intensification*, 133, 160–210. doi: 10.1016/j.cep.2018.07.014
- Toftgaard Pedersen, A., de Carvalho, T. M., Sutherland, E., Rehn, G., Ashe, R., and Woodley, J. M. (2017). Characterization of a continuous agitated cell reactor for oxygen dependent biocatalysis: Biocatalytic Oxidation in a Continuous Agitated Cell Reactor. *Biotechnology and Bioengineering*, 114(6), 1222–1230. doi: 10.1002/bit.26267
- Tong, X., Ma, Y., and Li, Y. (2010). Biomass into chemicals: Conversion of sugars to furan derivatives by catalytic processes. *Applied Catalysis A: General*, 385(1), 1–13. doi: 10.1016/j.apcata.2010.06.049
- Trambouze, P. J., and Piret, E. L. (1960). Extractive reaction: Batch- and continuous-flow chemical-reaction systems, concentrated case. *AIChE Journal*, 6(4), 574–578. doi: 10.1002/aic.690060414
- Wan Azelee, N. I., Mahdi, H. I., Cheng, Y.-S., Nordin, N., Illias, R. M., Rahman, R. A., Shaarani, S. M., Bhatt, P., Yadav, S., Chang, S. W., Ravindran, B., and Ashokkumar, V. (2023). Biomass degradation: Challenges and strategies in extraction and fractionation of hemicellulose. *Fuel*, 339, 126982. doi: 10.1016/j.fuel.2022.126982
- Weingarten, R., Cho, J., Conner, J., Wm. Curtis, and Huber, G. W. (2010). Kinetics of furfural production by dehydration of xylose in a biphasic reactor with microwave heating. *Green Chemistry : An International Journal and Green Chemistry Resource : GC*, 12(8), Article 8. doi: 10.1039/c003459b

# CARLA-based digital twin via ROS for hybrid mobile robot testing<sup>\*</sup>

Charlotte Stubenvoll<sup>\*,\*\*</sup> Tauno Tepsa<sup>\*\*</sup> Tommi Kokko<sup>\*\*</sup>  
Petri Hannula<sup>\*\*</sup> Heli Väättäjä<sup>\*\*</sup>

<sup>\*</sup> *charlotte.stubenvoll@lapinamk.fi*

<sup>\*\*</sup> *Lapland University of Applied Sciences, Rovaniemi, Finland*

**Abstract:** We developed the data connection via ROS (Robot Operating System) between a mobile robot and its digital twin in a CARLA-based autonomous driving simulator, which simulates realistic arctic winter weather conditions for safer, faster and less expensive autonomous vehicles testing. In our test setup we tested the hybrid case, where both robot twins were moving in the simulation and the real-world test area at the same time. Verifying our digital twin in regards to delays and applicability proved the communication via ROS to be occurring in almost real-time and the digital testing ground to profit from additional inbuilt reference points.

**Keywords:** Simulation, Robotics, digital twin, CARLA, ROS, multimaster.fkie, winter conditions, arctic, mobile robot, ATV

## 1. INTRODUCTION

Autonomously driving vehicles and robots that drive in public environments need to be safe and reliable under all weather conditions, including arctic winter conditions. Digital twins provide an opportunity to test autonomous vehicles in a safer, faster, and less expensive environment than carrying out tests in real-life conditions.

To our knowledge, the CARLA simulator for autonomous driving research is seldom, if at all, used for vehicles types not commonly found on public streets and for mobile robots other than autonomous cars. Mobile work robots, for example for deliveries, cleaning the sidewalks or for snow work, have to navigate public urban environments with pedestrians, bicycles or occasional traffic and can profit from digital twins that are capable to simulate those.

A data connection via ROS (Robot Operating System) between a mobile robot and its digital twin was developed. This allows for almost real-time exchange of commands, information, and sensor data between the twins. The digital twins of the robot and the testing ground are constructed in the WinterSIM, a CARLA-based autonomous driving simulator, which adds realistic arctic winter weather conditions to the simulation (Tepsa et al. (2021), Dosovitskiy et al. (2017)).

The digital twin design was informed by the intended future use cases: Testing, optimizing, controlling, and monitoring autonomous driving and snow cleaning functions first with the digital twin, then in hybrid approaches. We explored the applicability of our digital twin for those use cases in experiments assessing delays, verifying our digital twin setup and trying a hybrid obstacle avoidance scenario.

The backdrop for this research is the idea to use one or multiple mobile robots for autonomous snow cleaning of

<sup>\*</sup> Lapland Robotics and AI.R projects, Lapland University of Applied Sciences

the public outside areas of the campus. One of the robots build from scratch for this purpose is the small mobile robot called miniATV, shown in Fig. 1 with its sensors and actuators. It has proven to be able to clean off freshly fallen snow with a snow blower attachment on the front while manually controlled.

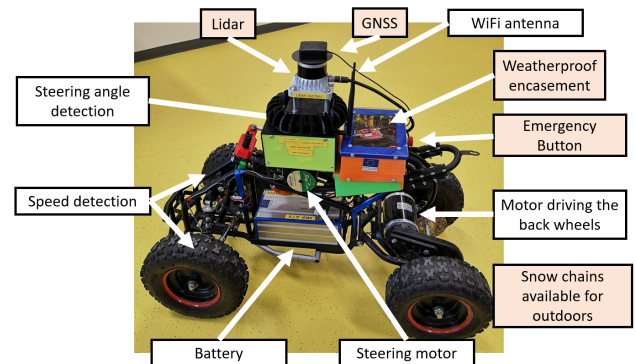


Fig. 1. The equipment, sensors and actuators of the miniATV

Automating the snow cleaning robot requires safe driving in the public outdoor spaces of the campus based on autonomous way-finding algorithms and robust obstacle avoidance algorithms, which is safer and more efficient to train within or in combination with a simulation. Simulations specifically for training autonomous snow cleaning robots have not been yet developed to our knowledge, so we used a simulation for autonomous driving for our purpose instead, which is described more in the next section on related work.

## 2. RELATED WORK

In the development of intelligent and automated traffic, it is particularly important to consider the impact of weather

conditions, especially in Finland and northern conditions. The goal of the WinterSIM project was to produce research and data on the performance of the most common vehicle sensors in winter conditions (FrostBit). In the WinterSIM project, a virtual reality (VR) simulation environment was implemented using a game engine, where weather conditions can be changed to demonstrate the impact of changing conditions on sensor data, particularly Lidar (Tepsa et al. (2021)). Generally, the produced simulation environments are utilized first, and then the driving data is transferred to the real world (Hu et al. (2023)).

The difference between the real world and its virtual counterpart is described by the concept of the "reality gap" (RG). This has been addressed in the publication by Hu et al. (2023). The study presents three different ways to reduce this gap: 1) transferring knowledge from simulation to reality, 2) learning in digital twins, and 3) learning by parallel intelligence technologies. (Hu et al. (2023)). The publication presents typical sim2real models for autonomous driving. In connection with the CARLA simulator, the KITTI vision benchmark suite is commonly used to bring features such as Lidar sensor data into the VR implementation. The KITTI benchmark suite was produced at the Karlsruhe Institute of Technology (KIT) and is freely available under a CC license. The problem of images produced by depth cameras in mixed-reality environments is particularly addressed by Argui et al. in their two publications (Argui et al. (2023) and Imane et al. (2023)). These publications present a method for combining depth camera images from the virtual world and the real world into a single view. The final conclusion of the publications is that this combination is potentially useful, but issues such as delays in combining camera data with the virtual world realistically could be problematic.

Bai et al. (2023) present an article on the Cyber mobility mirror architecture (CMM) to support Cooperative Driving Automation (CDA) research and development, and develop a CARLA-based co-simulation platform prototype. Additionally, it provides CARLA-based 3D object detection data and presents a case study demonstrating the necessity and functionality of lidar-based vehicle detection for CDA algorithm development.

The problem of accurately perceiving the environment is a very essential part of autonomous vehicles moving on public roads and in traffic systems. Cooperative perception (CP) and Vehicle-to-Everything (V2X) involve the sharing of data between autonomous vehicles and other road users, thereby extending traffic awareness beyond the sensory range of a single vehicle. This improves traffic flow and safety. Development towards this goal has been rapid in recent years, and a good overview can be found in the publication by Huang et al. (2024).

Deep and machine learning for decision-making and route planning in autonomous vehicles have benefited from developments in mapping and sensor technology. Examples of demanding decision-making include lane changing, where location data and lane markings, as well as tracking the movements of other vehicles, are utilized. Wang et al. (2023) present in their publication a deep learning method that improves the average success rate of lane changes compared to traditional planning control algorithms and

reinforcement learning methods. Practical tests were conducted using the CARLA simulator. The reinforcement learning method and its decision-making architecture are addressed by Al-Sharman et al. (2023) in their publication.

The unreliability of simulators in Autonomous Driving Systems (ADS) can lead to inconsistent test results. This is addressed by Amini et al. (2024) in their publication. The publication seeks to answer two research questions: (1) How do flaky ADS simulations affect automated testing based on random algorithms? and (2) Can machine learning (ML) effectively identify flaky ADS tests while reducing the number of required test repetitions? The publication concludes that unreliability is common and poses a real problem, but the presented method can effectively identify flaky ADS tests.

To minimize the influence of the reality gap in testing of autonomous vehicles and to take their real-life physics better into account, X-in-the-Loop testing as presented by Moten et al. (2018) is now considered state of the art. Drechsler et al. (2022) propose a similar approach called Dynamic Vehicle-in-the-Loop for testing automated driving functions. They feed a automated vehicle on a test track with simulated sensor information from the same track in the CARLA simulator that includes obstacles and pedestrians crossing the street. Our future robot testing use case is inspired by those approaches and lead is to explore the suitability of our digital twin setup for such robot testing approaches in an experiment described in the next section.

### 3. METHODOLOGY

#### 3.1 Software Architecture

The software architecture connects the digital and physical twins via ROS and in consequence ensuring updates from both sides fast enough for safety and to allow for deployment in use cases where both twins are used in parallel. Figure 2 shows the components of the software architecture and their interaction.

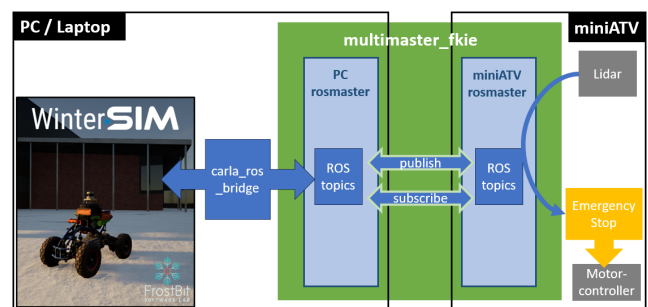


Fig. 2. Software architecture of the connection between the WinterSIM (in the PC or laptop) and the physical miniATV. The figure is explained from left to right.

Due to constraints in computation power, the simulation does not run on the robot itself, but on a separate machine, depicted in Fig. 2 as two black frames. Also any path-finding algorithm, object detection or test script is operating on the PC. This separation allows for cheaper, more compact and lightweight hardware architecture of

the robot, makes the robot easily interchangeable and the setup scalable to multiple robots.

Central to the software architecture is the *multimaster\_fkie* ROS package developed by the Fraunhofer FKIE (Tiderko et al. (2016)), here depicted in green. It allows to run separate ROS masters on each machine while being able to publish and subscribe to the topics of the other machine. The more common approach to run one ROS master for both machines was discouraged by the robot's rosserial not transmitting and receiving messages on the ROS simulation time the simulation requires.

In the PC, the *carla\_ros\_bridge* (CARLA Simulator) allows the two-way communication of the CARLA-based WinterSIM simulation with ROS over ROS topics. The WinterSIM contains a model of the robot and a map of the campus area which is used as digital and physical testing ground. On top of the weather simulated in carla, it offers a simulation of realistic arctic winter conditions and its influence on the sensor output (Tepsa et al. (2021)).

The mobile robot "miniATV" runs its own ROS master to which the sensor data of the robots sensors is published. Any incoming command is first filtered through the robot's lidar based emergency stop which checks if any obstacle is closer than a threshold. If an obstacle is detected, the original command is replaced with a stop command, if no obstacle is detected too close, the steering command is waved through. This improves safety in testing of the physical robot. Over rosserial, the command is sent to a micro controller controlling the motors for driving and steering.

Since the software architecture consists of a variety of different components, it was necessary to automate the startup of those in the right order.

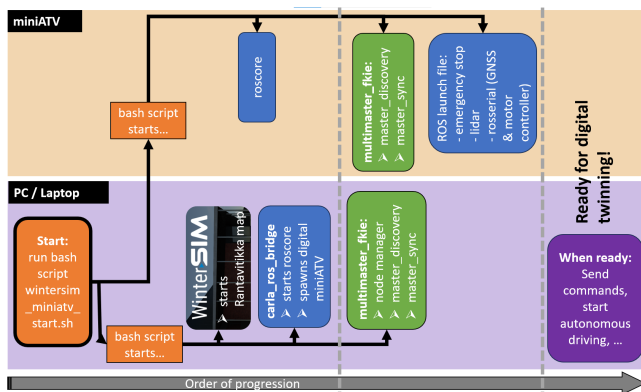


Fig. 3. The order of starting the software components for the digital twin

The components often rely on other software to be already running and cause errors if that is not the case. Especially for testing the setup it is important to have reliability and consistency in the startup process to minimize the human error. The startup script in the PC triggers the mobile robot's own startup sequence and starts first the WinterSIM and the *carla\_ros\_bridge* which also starts a roscore. Only after a ROS master is running in both the robot and the PC, the *multimaster\_fkie* can be started on each, discover the other rosmaster in the same network and synchronize the topics between both machines. In

the robot, the sensors and functionally necessary scripts are started then. After those are running and the data exchange via the *multimaster\_fkie* is established between the physical robot and the simulation, any testing scripts or path-finding algorithms can be started.

### 3.2 Experiments

*Delays:* The usability of the setup as a digital twin for testing the digital and physical twins in parallel testing scenarios like Dynamic Vehicle-in-the Loop testing (Drechsler et al. (2022)) depends highly on the communication happening in almost real-time.

In a test, both the real and the digital mobile robot receive the same steering commands. The ROS topics are recorded and then the time of arrival of the command in the robots' steering command topic and the reaction time of the robots are tracked. The rosbag time is the point in time, when the ROS node for the recording became aware of a ROS message and recorded it. It can vary from the individual timestamp that is solidified when the ROS message is sent out. Since the rosmaster in the PC is running on ROS simulation time and the robot on non-simulation time due to the rosserial connection to the motor controlling micro controller, using the rosbag time instead simplifies the timestamp evaluation. To identify if the network connection quality has a noticeable influence on delays in the system, the connection was changed from a mobile phone hotspot to a WiFi router after half the test runs. Measurements on the network speed were taken before each half with iperf3.

A testing script was developed to give the same commands to both twins at the same points in time in each test run, see Fig. 4. First, a zero speed command is sent for 30 seconds to both the physical and the digital twin. The long initial standstill was meant to improve the GNSS positioning on the initial position, but since the GNSS suffered from electric interference from the Lidar, the GNSS data could not be evaluated. After the initial standstill, a maximum speed command was given to both twins for 10 seconds, followed again by a 30 seconds stop command. This results in over 10 seconds where the speed is greater than zero, since accelerating and decelerating are both included. The measurement of delays utilizes those points in time where the speed measurement in the ROS messages of the robot's status topic changes from or back to zero.

Fourteen test runs have been conducted, each orchestrated by the same commands from the test script. The network connection was changed after half the test runs as described above. Each test run was recorded as a rosbag recording with all ROS topics and messages. The evaluation relies on the rosbag recording timestamps to detect delays and their magnitude, since the robot and the machine running the simulation require different clocks.

*Alignment* A pre-existing map of the testing area on campus was used in the WinterSIM. It was created for sensor research in the area of the campus where the physical robot testing ground is located. Around the robot testing area, the buildings and details are modeled precisely because the sensor research also took place in



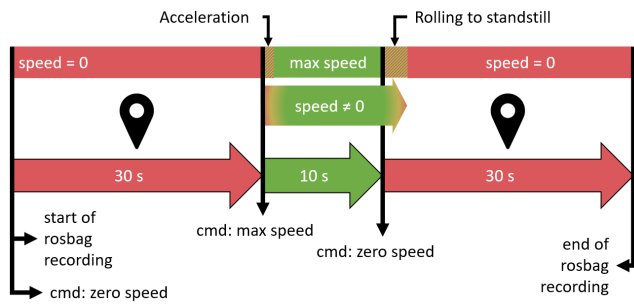


Fig. 4. Timing and course of each test run

that area. Further from this area used in testing, the map is sparse with details that were not conducive to the original purpose. The map was created based on a scan by a Leica BLK360 imaging laser scanner in winter and the surface height, which was offset due to snow covering, was later corrected with height data from the National Land Survey of Finland (NLS). Still, inconsistencies between the surface height of the real testing ground and the map were noted. The spatial relationship between buildings and details such as lampposts however is modeled realistically for the original sensor research purpose of the map. This test verifies the surface height of the digital map utilizing pre-existing GNSS reference points as ground truth in the physical space. Four exemplary points were chosen where the mobile robots would be driving later during tests. Since those reference points are not modeled in the map, triangulation to landmarks consisting of the precisely modeled details and buildings present in both the physical space and the digital map of the space was undertaken. Due to that, the number of reference points is chosen that small because the overlap between the map area detailed enough for triangulation, the area in which the robots would be roaming and the available reference points did not allow for more. The reference points are not modeled in the map of the robot testing area. To display them in the WinterSIM map, the coordinates of the reference points are transformed from the ETRS-TM35FIN/N2000 reference coordinate system to geodetic coordinates with the Paikkatietoikkuna portal maintained by the National Land Survey of Finland (NLS). The geodetic coordinates are then translated into CARLA's own coordinate system and made visible in the map of the CARLA-based WinterSIM with spawning a mark at this location. The marks consist of three axis in x, y, and z direction intersecting at the marked point for easy visibility of the reference point and precise placement of the mark at the landmarks' locations.

Any discrepancy in surface level height is measured by placing another mark on the projection of the reference point onto the surface and calculating the distance. To judge the alignment by triangulation, if the reference point coordinate is located at a similar spot in the digital map then in the real robot testing ground, the distance to the chosen landmarks in the digital map is measured with a mark similarly. Three landmarks were chosen per reference point out of prominent details and structures existing both in the digital and real map nearby the reference points. In the real testing ground, the distance between reference points and their landmarks was measured with a measuring tape. Unevenness of the surface, overgrown edges

and even the grass on the lawn however added imprecision that suggests that the millimeter accuracies that scans can achieve will not be reflected in the comparison between the tape measurements and the distances in the map.

*Object Avoidance* The digital twin was created for robot testing and one of the intended use cases is a Robot-in-the-Loop test after the Dynamic Vehicle-in-the-Loop approach Drechsler et al. (2022) proposed. The digital twin we created would be fit for application in such a way if the real robot's reactions can be made dependent on the sensor input from the digital twin. If the sensors of the digital twin show an obstacle being too close, both the real robot and its digital twin should take action not to hit the object only present in the simulation. A prerequisite is a functioning communication between the twins in almost real-time, which was tested in the first experiment on delays above. This experiment examines the applicability of our setup for hybrid Robot-in-the-Loop testing scenarios.

Figure 5 shows the data flow for this experiment. In order not to interfere with the mobile robot's own internal lidar-based emergency stop in case of unexpected physical obstacles in the test area of the physical twin, the physical twin is not using the digital sensor data directly instead of its own sensor data. Instead, a similar lidar based emergency stop is created which receives the only the digital lidar data from the simulation and forwards the steering commands only if no digital obstacle is too close and otherwise sends a stop command. This forwarded steering command is then input to both the digital and physical robots. The physical robot then first checks in its emergency stop if no real physical obstacle is too close and only then puts this incoming command to action. The experiment on delays examines if this extra step can cause a delay.

Ten test runs have been conducted for this experiment orchestrated by a test script. Both the digital and the real robot are placed at the same location in the digital and real testing ground at the start of each test run. A digital obstacle is spawned at the same location for each test run only in the simulation in the digital robot's way. No obstacle is in the physical robot's way. The initial command that is then filtered through the robots' emergency stops according to Fig. 5 is then send out, the command is for maximum speed for ten seconds. During the execution of the command the simulated lidar data in the digital twin will then show the obstacle that soon is closer as the chosen threshold of 1.2 m. The digital twins emergency stop should then send a stop command instead of forwarding the maximum speed command to itself and the physical robot. Both twins stop propelling themselves forward and roll to a standstill since neither twin has brakes.

## 4. RESULTS

### 4.1 Experiment on Delays

Table 1 shows the measured network connection bandwidths between the command giving machine where the simulation and the test scripts were running and the real



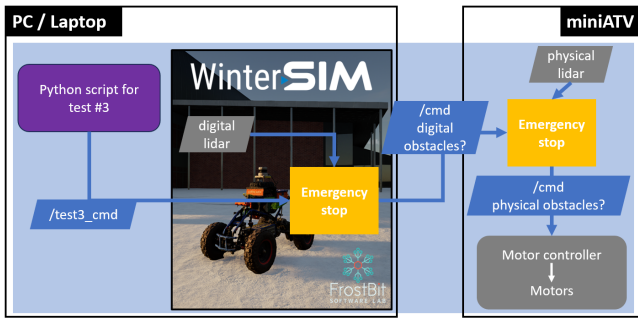


Fig. 5. How the cmd runs through the setup

Table 1. Bandwidth acquired with iperf3 for both the phone hotspot and the WiFi box

Network	From →To	Bandwidth sender [Mbits/sec]	Retry	Bandwidth receiver [Mbits/sec]
Phone hotspot	PC →miniATV	21.0	7	18.2
Phone hotspot	miniATV →PC	18.9	0	18.4
WiFi box	PC →miniATV	66.6	74	63.8
WiFi box	miniATV →PC	31.5	0	29.9

mobile robot after completing the automated startup sequence and before each half of the test runs. As expected, the WiFi has a better bandwidth compared to the phone hotspot, but an unexpected high number of retries of sending data packages. The quality of the network connection does not seem to have any noticeable influence on the delays though.

Figure 6 depicts the average delays in the command flow, both for the maximum speed command and the zero speed command combined. Only in the last step the depiction of the delays was divided since starting to drive after the first maximum speed commands and rolling to a standstill after the zero speed command are mechanically different and produced different delays.

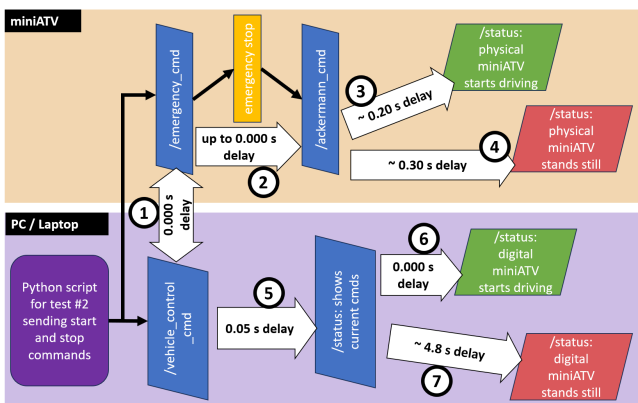


Fig. 6. Average over the delays in the command flow

The differences between the timestamps for each step in the flow of commands has been computed for each of the two commands (maximum speed and zero speed) for each of the fourteen test runs and then averaged. The only exception is delay number two (see Fig. 6). Only when sending the maximum speed command, never the zero speed command, the delay was up to 3.4 s. But the next delay in the flow, delay number three then showed a delay of over -3 s. This suggests that the delay was

not in the command flow itself, but in the recording of the topic. The command on which the computation of delay two is based has been recorded in the rosbag sooner than the previous command. The delay calculation is based not on the timestamp when the ROS message originated since the simulation and the robot require vastly different clocks, but on the timestamp of when the message became available for the recording ROS node.

As a side note, no worsening of the delays has been found over the course of the fourteen test runs, which indicates that exposure of the equipment to temperatures around 0 °Celsius did not influence the delays.

4.2 Experiment on Alingment

Table 2. Difference in distances between the reference points and the landmarks between the digital and the physical twin

Reference Point	Landmark	Absolute difference [m]
1	Landmark1	0.349
	Landmark2	1.852
	Landmark3	0.151
2	Landmark1	3.195
	Landmark2	5.241
	Landmark3	1.010
3	Landmark1	0.475
	Landmark2	0.689
	Landmark3	2.132
4	Landmark1	1.533
	Landmark2	0.823
	Landmark3	0.031

The surface height of 3 out of 4 reference points is aligning within 10 cm. On the other reference point, the difference between the map surface and the coordinates of the reference point translated into a map position is 78 cm. This occurs in an especially uneven area of the testing ground.

For the x and y axis, the distances between the reference points and the landmarks varies greatly between the real testing ground and the digital map of the same area and does not reflect the millimeter or centimeter level accuracies that could be possible for a map based on a scan with this technique and equipment. Only half of the differences in distances to the 12 landmarks lie under one meter with 0.031 m as the lowest difference and only one under 10 cm. The highest difference is 5.241 m, suggesting that either the detail used as landmark was placed wrongly or a wrong detail has been taken for the landmark in the evaluation.

*Experiment on Obstacle Avoidance* Out of the ten test runs, the digital robot stopped propelling itself forward in all of them once the obstacle came closer than the threshold. Unfortunately, since neither the real nor the digital robot possess breaks, the digital robot did not roll to a standstill in time not to collide with the obstacle. This behavior warrants an adjustment in parameters of the digital twin model of the robot. CARLA offers a variety of parameters to adjust a simulated vehicles behavior, but they are tailored to vehicles typically found in public

traffic, not to self-build smaller mobile robots. Therefore, the adjustment is not something that could be done on the fly once the problem became apparent in the test, but needs multiple rounds of try and error.

The physical robot in the real testing ground only stopped in eight out of ten test runs. In one of those, the real robot stopped only after approximately ten seconds of driving with maximum speed, in the other the real robot did not even move from the starting position. Since the digital robot drove and reacted to the obstacle in those two test runs. One explanation for the first case would be that the physical robot lost connection and regained it after those approximately ten seconds. The test script would have ended after ten seconds as well, but it is probably a coincidence that the robot regained connection around the same time. If the real robot had not regained the connection to receive a stop command, it's own emergency stop would have prevented it from crashing into any obstacles. In the second case, when the physical robot would not start, the most likely explanation is that the startup sequence was not yet fully finished and that the test script was started by the experimenter before the respective components were ready in the physical robot. In the eight out of ten test runs in which the physical robot stopped due to the digital obstacle perceived by its digital twin, it almost immediately came to a standstill without crashing into the digital object.

## 5. SUMMARY AND DISCUSSIONS

The Experiment on delays finds that the communication via ROS between the digital twins is indeed in almost real-time with delays under 0.4 seconds both for command transmission and maximal speed changes, with the exception of the rolling to a standstill behaviour of the digital robot that has to be adjusted to be realistic. Therefore the setup is suitable for application in the intended robot training and testing scenarios. This experiment also uncovered a steep discrepancy between the rolling to a standstill behaviour between the digital and physical robot twins, which was discovered in the experiment on obstacles as well.

The Alignment experiment shows that the surface height of the map in the simulation is within 10 cm for three of the four reference points. The triangulation to verify the position of the reference points however did not provide clear indications if the alignment in the x and y axis is sufficient, half of the distances to landmarks differ more than one meter from the measurements in the real testing ground. Uneven ground, vegetation and differences in the level of detail likely added to inaccuracies in the measurement. Adding GNSS reference points to such maps in the simulation in future, for example by using targets measured with a GNSS total station on reference points, would not just improve the accuracies of the simulation but also make the verification process easier.

The Experiment on avoiding obstacles that exist only in the simulation suggests adjustment of the rolling parameters of the digital robot twin, so that the digital robot also rolls to a standstill almost immediately like the physical robot. Otherwise this difference in behavior could cause a

worsening of the reality gap between training algorithms in the digital twin compared to the physical twin.

The physical robots behaviour during the test suggests adding a check to the startup sequence if all required components are running and responding before giving a clear for starting further scripts. Also monitoring if the connection to the command giving machine is still intact (and stopping if the connection is lost) would prevent the need to rely on the real robot's emergency stop and thus increasing safety in testing if the connection is lost.

With the described adjustments to the digital twin setup and model parameters, the digital twin is ready to be deployed in the intended robot training and testing scenarios. We will also introduce other robot into the digital twin for testing, such as an autonomous car or robots for heavier snow work. The testing ground will be equipped with 5G for a faster and more stable connection. Following the doubts about the sufficient accuracy of the map of the campus area that our tests showed, the map is currently remade from a drone scan with 1 - 2 cm accuracy and GNSS data. Additionally, a second test ground has been scanned with a Leica BLK 360 laser scanner and GNSS reference points to create a map for robots to train in an industrial environment with an overall accuracy of 5 mm. Since ROS 2 offers new features that are interesting for autonomous vehicles, we are currently moving our platforms to away from ROS 1.

## ACKNOWLEDGEMENTS

This paper was made possible by the Lapland Robotics and AI.R projects of the Lapland University of Applied Sciences. Many thanks also to the creators of the WinterSIM (by the FrostBit Software Lab of the Lapland University of Applied Sciences) and especially Tuomas Herranen.

## REFERENCES

- Al-Sharman, M., Dempster, R., Daoud, M.A., Nasr, M., Rayside, D., and Melek, W. (2023). Self-learned autonomous driving at unsignalized intersections: A hierarchical reinforced learning approach for feasible decision-making. *IEEE Transactions on Intelligent Transportation Systems*, 24(11), 12345–12356. doi:10.1109/TITS.2023.3285440.
- Amini, M.H., Naseri, S., and Nejati, S. (2024). Evaluating the impact of flaky simulators on testing autonomous driving systems. *Empirical Software Engineering*, 29(2), 47. doi:10.1007/s10664-023-10433-5.
- Argui, I., Guérliau, M., and Ainouz, S. (2023). A mixed-reality framework based on depth camera for safety testing of autonomous navigation systems. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2050–2055. doi:10.1109/ITSC57777.2023.10421982.
- Bai, Z., Wu, G., Qi, X., Liu, Y., Oguchi, K., and Barth, M.J. (2023). Cyber mobility mirror for enabling cooperative driving automation in mixed traffic: A co-simulation platform. *IEEE Intelligent Transportation Systems Magazine*, 15(2), 251–265. doi:10.1109/itsmag.2023.10421982.

- 2022.3203662. doi:10.1109/MITS.2022.3203662.
- CARLA Simulator, *ROS Bridge Documentation*. URL <https://carla.readthedocs.io/projects/ros-bridge/en/latest/>. Online. Accessed on 12.12.2023.
- Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017). CARLA: An open urban driving simulator. In *Proceedings of the 1st Annual Conference on Robot Learning*, 1–16.
- Drechsler, M., Sharma, V., Reway, F., Sch”
- Huber, W. (2022). Dynamic vehicle-in-the-loop: A novel method for testing automated driving functions. *SAE International Journal of Connected and Automated Vehicles*, 5, 1–14. doi:10.4271/12-05-04-0029.
- FrostBit, *WinterSIM*. URL <https://www.frostbit.fi/en/portfolio/wintersim-2/>. Accessed on 10.06.2024.
- Hu, X., Li, S., Huang, T., Tang, B., Huai, R., and Chen, L. (2023). How simulation helps autonomous driving: a survey of sim2real, digital twins, and parallel intelligence.
- Huang, T., Liu, J., Zhou, X., Nguyen, D.C., Azghadi, M.R., Xia, Y., Han, Q.L., and Sun, S. (2024). V2x cooperative perception for autonomous driving: Recent advances and challenges.
- Imane, A., Guérliau, M., and Ainouz, S. (2023). Building a vision-based mixed-reality framework for autonomous driving navigation. doi:10.1109/CoDIT58514.2023.10284251.
- Moten, S., Celiberti, F., Grottole, M., van der Heide, A., and Lemmens, Y. (2018). X-in-the-loop advanced driving simulation platform for the design, development, testing and validation of adas. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, 1–6. doi:10.1109/IVS.2018.8500409.
- Tepsa, T., Korhonen, M., Paananen, R., Narkilahti, A., and Herranen, T. (2021). Towards autonomous vehicle winter simulations utilising carla and real-world sensor data. In *27th ITS World Congress, Hamburg, Germany, 11-15 October 2021. Paper ID 895.*, 2330 – 2338.
- Tiderko, A., Hoeller, F., and Röhling, T. (2016). *The ROS Multimaster Extension for Simplified Deployment of Multi-Robot Systems*, volume 625, 629–650. doi:10.1007/978-3-319-26054-9\_24.
- Wang, J., Chu, L., Zhang, Y., Mao, Y., and Guo, C. (2023). Intelligent vehicle decision-making and trajectory planning method based on deep reinforcement learning in the frenet space. *Sensors*, 23, 9819. doi:10.3390/s23249819.

# Identifiability and Kalman Filter Parameter Estimation Applied to Biomolecular Controller Motifs

Eivind S. Haus, Malin Harr Overland, Damiano Rotondo,  
Kristian Thorsen, Tormod Drengstig

*Department of Electrical Engineering and Computer Science,  
University of Stavanger, Norway, (eivind.haus@uis.no)*

---

**Abstract:** In this paper we apply Augmented Extended Kalman filters (AEKFs) to perform parameter estimation in two different biological controller motifs under both noise-free and noisy conditions. Based on measurements of the two states of the controller motifs, we show that under both noise conditions it is possible to estimate all 5 and 6 parameters, respectively, which is in accordance with previously published results that investigated the theoretical concept of *structural identifiability*. We further investigate how the level of process/measurement noise and the initial estimates of both the parameters and states in the AEKFs affect the estimation performance, and the results indicate that the degree of non-linearity affects filter performance.

*Keywords:* controller motifs, observability, identifiability, Augmented Extended Kalman filter

---

## 1. INTRODUCTION

Mathematical models are widely used to understand and predict the behavior of real-world systems, both human-engineered and natural. Within the field of systems biology, complex models of metabolic pathways and entire cells have been created in recent years. In this context, *controller motifs* are simple biomolecular reaction networks that can explain how regulatory function is achieved through negative feedback, see e.g. Thorsen et al. (2013); Briat et al. (2016); Tang and McMillen (2016), and thus, they are useful as building blocks in larger biological models as shown in Agafonov et al. (2016).

One of the main purposes of a mathematical model is to make predictions of the modeled system's behavior. The accuracy of model predictions is heavily reliant on parameter values, and finding accurate parameter values is often a difficult and time-consuming process that requires experiments. *Identifiability* and *observability* are two closely related concepts that are helpful tools when evaluating the usefulness of a model, as they provide information on whether the model can be parameterized with the available measurements. Identifiability describes whether unknown model parameters can be determined based on knowledge of the input and output of the model. On the other hand, observability implies that it is possible to infer the values of unmeasured states by combining together the measurements and the available model.

In practice, the estimation of unmeasured states is achieved by designing a special type of dynamical system called *state observer*. Among several possible structures proposed for models described by nonlinear differential equations, the extended Kalman filter (EKF) is the *de facto* standard in many applications, such as navigation systems and GPS, see e.g. Huang et al. (2009); Loron and Laliberte (1993); Böhler et al. (2021); Narayanan et al.

(2020). The EKF merges two sources of information: the mathematical model, used to compute future prediction of the state based on the current estimate, and the measurements, which introduce a feedback mechanism for addressing possible model-reality mismatch and the presence of unmodeled disturbances. This information fusion is done by weighting the different sources of information according to their reliability (for example, a noisy measurement is weighted less than the data coming from a precise sensor). The EKF addresses the inherent nonlinearities of the model by applying a linearization based on first-order Taylor expansions, thus approximating the nonlinearities as linear functions in small regions around the current state estimate. Although this linearization introduces some error, motivating the search for more efficient estimators (see, e.g., Julier and Uhlmann (1997); Huang et al. (2020); Sarmavuori and Sarkka (2011); Liu and Guo (2021); Rotondo (2023)), the EKF is still considered to be effective in many practical scenarios.

In Haus et al. (2023), the *structural identifiability* of a set of controller motifs was investigated using a symbolic approach. Structural identifiability is a theoretical concept that is fully determined by the structure of the model and the chosen outputs, and assumes that measurements are noise-free and sufficiently informative, see e.g. Ljung and Glad (1994); Villaverde et al. (2018). As it is well known that available measurements from biological systems are limited, practical identifiability limitations may occur for models that have been found to be structurally identifiable. In this paper, we investigate identifiability of controller motifs in a practical setting using Augmented Extended Kalman filters (AEKFs) with varying degree of process/measurement noise. We further investigate how the initial conditions of the parameters and state variables in the filters affect the performance.

## 2. CONTROLLER MOTIFS

Controller motifs are simple molecular reaction networks where negative feedback is achieved through signaling between the species. In this paper, we consider motif 1 and 3, shown in Fig. 1, from the set of 8 basic two component controller motifs presented in Drenth et al. (2012).

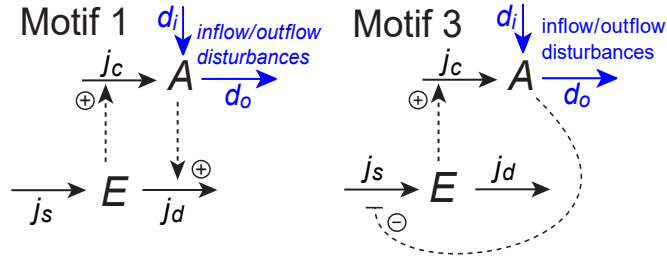


Fig. 1. Basic controller motif 1 and 3.

The general state equations for the motifs are given as:

$$\dot{A}(t) = d_i(t) - d_o(t) + j_c(t) \quad (1)$$

$$\dot{E}(t) = j_s(t) - j_d(t), \quad (2)$$

where  $A$  is the controlled species,  $E$  is the controller species,  $d_i$  and  $d_o$  are input and output disturbances, respectively,  $j_c$  is the compensatory flow, and  $j_s$  and  $j_d$  are the synthesis and degradation flows of  $E$ , respectively.

The purpose of the compensatory flow  $j_c$  is to maintain the level of  $A$  by adding  $A$  in the presence of a dominating outflow disturbance  $d_o$ , and both motifs are therefore called *inflow* controllers. Moreover,  $j_c$  is for both motifs *activated* by  $E$  as increased level of  $E$  results in increased inflow through  $j_c$  (indicated by a  $\oplus$  sign). The difference between the two motifs is the signaling from  $A$  to  $E$ , where motif 1 is activating and motif 3 is inhibiting (higher level of  $A$  results in lesser flow). We assume that activation is modeled as first order kinetics and inhibition as saturable kinetics. Furthermore, inflows (synthesis) are modeled as zero order, while outflows (degradation) are modeled as first order with respect to its own state variable. The state equations for motifs 1 and 3 are given in Eqs. (3)–(6):

Motif 1:

$$\dot{A}(t) = k_i - k_o \cdot A(t) + k_c \cdot E(t) \quad (3)$$

$$\dot{E}(t) = k_s - k_d \cdot A(t) \cdot E(t), \quad (4)$$

Motif 3:

$$\dot{A}(t) = k_i - k_o \cdot A(t) + k_c \cdot E(t) \quad (5)$$

$$\dot{E}(t) = k_s \cdot \frac{K_i^A}{K_i^A + A(t)} - k_d \cdot E(t), \quad (6)$$

where  $k_i$ ,  $k_o$ ,  $k_c$ ,  $k_s$ , and  $k_d$  are rate constants, and  $K_i^A$  is an inhibition constant.

## 3. OBSERVABILITY AND IDENTIFIABILITY

In order to introduce the concepts of observability and identifiability we consider a general nonlinear state space model:

$$\dot{x}(t) = f(x(t), p, u(t)) \quad (7)$$

$$y(t) = g(x(t), p) + v(t) \quad (8)$$

$$x_0 = x(t_0, p), \quad (9)$$

where  $x(t) \in \mathbb{R}^{n_x}$  is the state vector,  $u(t) \in \mathbb{R}^{n_u}$  is the input vector,  $y(t) \in \mathbb{R}^{n_y}$  is the output vector,  $v(t) \in \mathbb{R}^{n_v}$  is the measurement noise vector, and  $p \in \mathbb{R}^{n_p}$  is a vector of system parameters assumed to be constant. Furthermore,  $x_0 \in \mathbb{R}^{n_x}$  denotes the initial conditions, and the nonlinear functions  $f(\cdot)$  and  $g(\cdot)$  define the state and output equation, respectively.

## 3.1 Observability

Biological systems are typically only partially observable due to experimental limitations, and the available system outputs may be a function of several states and/or parameters (Raue et al., 2009). Thus, to determine the unmeasurable states they must be inferred from the available system outputs, which is possible only if the system is *observable* (Kalman, 1960).

*Observability:* Given an initial state  $x_0$  and an admissible control  $u(t)$ , if the current system state  $x(t)$  can be determined only through the system output  $y(t)$  in a finite time, the system is said to be *observable*. (Miao et al., 2011)

Observability is a theoretical concept determined by the system structure and the chosen outputs and typically assumes noise-free measurements of the system output  $y$  and that the parameters  $p$  of the model are known. However, the parameters of a biological system are rarely known, as these typically represent biochemical processes inside the cells that are impossible to measure directly. Thus, the parameter values must be estimated, typically based on experiments, which often is associated with high cost. In order to minimize the potential cost of experiments, *identifiability* analysis is a useful tool in finding which measurements are necessary to fully parameterize a model.

## 3.2 Identifiability

*Identifiability:* The dynamic system given by Eqs. (7)–(9) is *identifiable* if  $p$  can be uniquely determined from the given system input  $u(t)$  and the measurable system output  $y(t)$ ; otherwise, it is said to be *unidentifiable*, (Miao et al., 2011).

If a model is identifiable, it is possible to uniquely determine the value of all its system parameters based on the chosen model structure and outputs. Thus, through performing identifiability analysis for different measurement combinations, the smallest, or easiest to perform, set of measurements that allows the model to be accurately parameterized can be found. However, identifiability is, as observability, a theoretical concept, and *practical* identifiability limitations such as noisy measurements, restrictions on admissible inputs, or limited time resolution of outputs may occur.

Observability and identifiability both imply a strong connection between inputs, states and outputs, and identifiability can be considered a particular case of observability where system parameters are treated as states with zero dynamics, see e.g. Villaverde et al. (2016); Villaverde (2019). Consequently, the parameters  $p$  are included as part of the state vector  $x$  and methods originally developed for investigating observability can also be used for parameter identifiability as shown in Villaverde et al. (2016). This



approach is similar to an Augmented Kalman filter, which we introduce in the next section.

#### 4. KALMAN FILTERING

Let us consider the discrete-time nonlinear system:

$$z(k+1) = \phi(z(k), u(k)) + w(k) \quad (10)$$

$$y(k) = \gamma(z(k)) + v(k) \quad (11)$$

where  $z(k) \in \mathbb{R}^{n_z}$  denotes the (unknown) state vector,  $u(k) \in \mathbb{R}^{n_u}$  denotes the (known) input vector,  $y(k) \in \mathbb{R}^{n_y}$  denotes the (known) output vector,  $w(k) \in \mathbb{R}^{n_x}$  is the (unknown) process noise,  $v(k) \in \mathbb{R}^{n_v}$  is the (unknown) measurement noise, and  $\phi(\cdot)$  and  $\gamma(\cdot)$  denote known functions assumed to be differentiable with respect to  $z$ . It is also assumed that  $w(k)$  and  $v(k)$  are zero mean multivariate Gaussian noise with (known) covariance matrices  $Q \in \mathbb{S}^{n_x}$  and  $R \in \mathbb{S}^{n_h}$ , respectively, where  $\mathbb{S}^n$  denotes (real) symmetric matrices of order  $n$ .

##### 4.1 Extended Kalman filter (EKF)

The EKF comprises alternating phases of model-based prediction and measurement-based update, forming an ongoing (theoretically infinite) cycle involving the following variables and matrices, which are internal to the EKF:

- the *a priori estimate*  $\bar{z}(k) \in \mathbb{R}^{n_z}$ ;
- the *a priori covariance matrix*  $M(k) = E \left[ (z(k) - \bar{z}(k))(z(k) - \bar{z}(k))^T \right] \in \mathbb{S}^{n_z}$ ;
- the *a posteriori estimate*  $\hat{z}(k) \in \mathbb{R}^{n_z}$ ;
- the *a posteriori covariance matrix*  $P(k) = E \left[ (z(k) - \hat{z}(k))(z(k) - \hat{z}(k))^T \right] \in \mathbb{S}^{n_z}$ ;

During the model-based prediction, the EKF computes the *a priori estimate* based on the state equation, as follows:

$$\bar{z}(k) = \phi(\hat{z}(k-1), u(k-1)) \quad (12)$$

At the same time, the *a priori covariance matrix* is updated according to the following equation:

$$M(k) = \mathcal{A}(k)P(k-1)\mathcal{A}(k)^T + Q \quad (13)$$

where the matrix  $\mathcal{A}(k)$  is obtained through a first-order Taylor approximation of the nonlinear function  $\phi$  at the most recent state estimate:

$$\mathcal{A}(k) = \left. \frac{\partial \phi}{\partial z} \right|_{\bar{z}(k-1), u(k-1)} \quad (14)$$

where the operator  $\partial/\partial$  is meant in the Jacobian matrix sense.

During the measurement-based update, the EKF computes the Kalman gain  $K(k)$  to be used to account for the so-called *innovation*, i.e., the measurement-estimate mismatch, according to:

$$K(k) = M(k)\mathcal{C}(k)^T [\mathcal{C}(k)M(k)\mathcal{C}(k)^T + R]^{-1} \quad (15)$$

where the matrix  $\mathcal{C}(k)$  is obtained through a first-order Taylor approximation of the nonlinear function  $\gamma$  at the most recent state estimate:

$$\mathcal{C}(k) = \left. \frac{\partial \gamma}{\partial z} \right|_{\bar{z}(k)} \quad (16)$$

The Kalman gain essentially determines how much weight is given to the current measurement, and helps strike a

balance between trusting the model and incorporating new measurements  $y(k)$ , according to the following equation:

$$\hat{z}(k) = \bar{z}(k) + K(k) [y(k) - \gamma(\bar{z}(k))] \quad (17)$$

Finally, the current value of the a posteriori covariance matrix is computed using:

$$P(k) = [I - K(k)\mathcal{C}(k)] M(k) \quad (18)$$

##### 4.2 Augmented Extended Kalman filter (AEKF)

Following an approach that is well consolidated in the field of fault diagnosis, see e.g. Patton and Klinkhieo (2009); Zhang et al. (2020); Rotondo et al. (2021), it is possible to use state observers, such as the above described EKF, to obtain a real-time estimate of the unknown parameter vector  $p$  appearing in Eq. (7). The first required step is a discretization of Eq. (7), which can be done using a variety of methods (Franklin et al., 1998), with forward-Euler being the most common due to its simplicity, thus obtaining:

$$x(k+1) = x(k) + T_s \cdot f(x(k), p, u(k)) \quad (19)$$

$$y(k) = g(x(k), p) + v(k) \quad (20)$$

$$x_0 = x(k_0, p), \quad (21)$$

where  $T_s$  denotes the sampling time. Then, a description of the dynamical behavior of the parameters to be estimated is introduced, which in the case of constant parameters reads as follows:

$$p(k+1) = p(k) \quad (22)$$

At this point, it is possible to define an augmented state vector as  $z(k) = [x(k)^T, p(k)^T]^T = [z_1(k)^T, z_2(k)^T]^T$  and, by combining together Eqs. (19)–(22), obtain an augmented state-space model that fits the form of Eqs. (10)–(11), with:

$$\phi(z(k), u(k)) = \begin{bmatrix} z_1(k) + f(z_1(k), z_2(k), u(k)) \\ z_2(k) \end{bmatrix} \quad (23)$$

$$\gamma(z(k)) = g(z_1(k), z_2(k)) \quad (24)$$

Clearly, an EKF implemented on the augmented model would return estimates  $\bar{z}(k)$  and  $\hat{z}(k)$  which correspond to the joint state-parameter estimates  $\bar{x}(k)$ ,  $\bar{p}(k)$  and  $\hat{x}(k)$ ,  $\hat{p}(k)$ , respectively.

To this end, one can define the matrix  $Q$  to account for disturbances acting on the state  $x$ , the possible time-varyingness of  $p$ , or to incorporate information about the model uncertainty into the estimation.

##### 4.3 Implementation

The models in Eqs. (3)–(4) and Eqs. (5)–(6), respectively, are the basis for both the process and the AEKF for motif 1 and 3, where the augmented state vectors are given as

$$\text{motif 1: } z(k) = [A, E, k_i, k_o, k_c, k_s, k_d] \quad (25)$$

$$\text{motif 3: } z(k) = [A, E, k_i, k_o, k_c, k_s, k_d, K_i^A]. \quad (26)$$

For easier reference to the parameters, we refer to them as

$$[k_i, k_o, k_c, k_s, k_d, K_i^A] = [k_1, k_2, k_3, k_4, k_5, k_6], \quad (27)$$

or generally as  $k_n \forall n \in \{1, \dots, N\}$  where  $N=5$  for motif 1 and  $N=6$  for motif 3. As there is no external input  $u(k)$ , the motifs are *autonomous* systems (Haus et al., 2023), and the estimation is solely based on measurements of the state variables.

## 5. RESULTS

One of the main findings in Haus et al. (2023) was that noise free measurements of the state variables  $A$  and  $E$  of motif 1 and 3 proved sufficient for structural identifiability, i.e. in theory it is possible to uniquely estimate all the parameters. Hence, the results presented here are focused on reproducing these results using AEKF, and also to investigate the robustness with respect to noise level and initial filter conditions of parameters and state variables.

## 5.1 Simulation setup

For simplicity, the values of all the parameters in the process are set to 1, and we term these parameters as  $k_{n,\text{process}}$ . The steady state values of  $A$  and  $E$  for each motif are calculated using Eqs. (3)–(4) and Eqs. (5)–(6), respectively. As there are no external input to provide excitation, the initial conditions of  $A$  and  $E$  in both the process and the AEKF are increased with 30% from the calculated steady state values. The initial conditions of the unknown parameters in the AEKF are assigned a uniformly distributed random number within 1% deviation from the value of  $k_{n,\text{process}}=1$ . This small deviation is considered sufficient to demonstrate whether the AEKF produce similar results as the structural identifiability from Haus et al. (2023).

For the noisy conditions, both the process noise  $w(k)$  and the measurement noise  $v(k)$  are modeled as bandwidth limited zero mean white noise. Simulations are run for 10 seconds with a stepsize of  $T_s=0.0002$  seconds. As the AEKFs typically converge within 2 seconds, only the first 3 seconds are shown in Figs. 2–4. In order to quantitatively compare the Kalman filter performances, we calculate the following *parameter error measure* for each individual parameter  $k_n$ ,

$$\Delta e_n = \frac{1}{5000} \sum_{i=45000}^{50000} |k_{n,\text{process}}(i) - k_{n,\text{est}}(i)| \quad (28)$$

where  $k_{n,\text{est}}$  is the estimated parameter value. Thus,  $\Delta e_n$  is the mean of the difference between the process value and the estimated value of parameter  $k_n$  during the last second of the simulation, i.e. the last 5000 samples. To get a performance measure for the entire motif, we calculate an *overall motif error measure* as follows,

$$\Delta e_{\text{motif}} = \frac{1}{N} \sum_{n=1}^N \Delta e_n \quad (29)$$

where  $N=5$  for motif 1 and  $N=6$  for motif 3.

All state and parameter values are either in arbitrary units or without units.

## 5.2 Motif 1

The simulation results for the noise free motif 1 are shown in Fig. 2, where we observe that all parameters quickly converge with an overall *motif error measure* of  $\Delta e_{\text{motif}}=0.0002$ . These results are in accordance with the findings in Haus et al. (2023). In order to investigate the effect of noise, we added both process and measurement noise to  $A$  and  $E$ , and increased the noise power gradually. A simulation example of this is shown in Fig. 3, where the

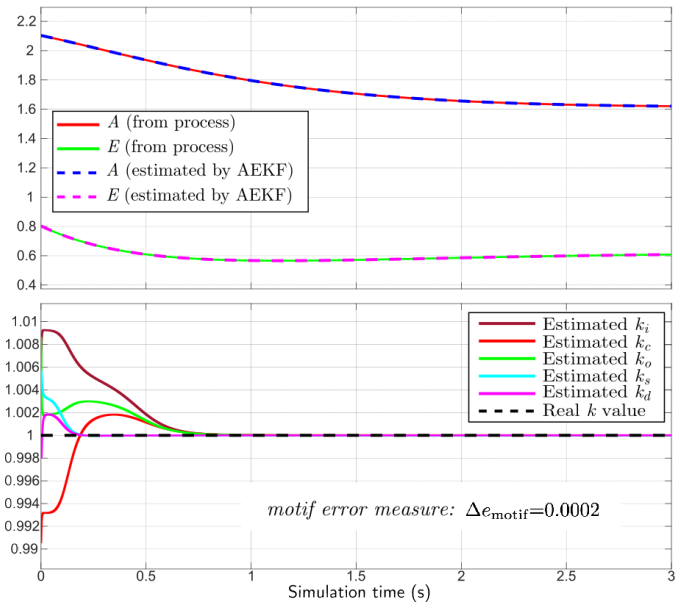


Fig. 2. Simulation results for motif 1 with no noise. The upper panel shows the states  $A$  and  $E$ , whereas the lower panel shows the parameter estimates.

noise power is  $1 \cdot 10^{-6}$  for both  $w(k)$  and  $v(k)$ , and where all parameters converge fairly quickly with an overall *motif error measure* of  $\Delta e_{\text{motif}}=0.059$ .

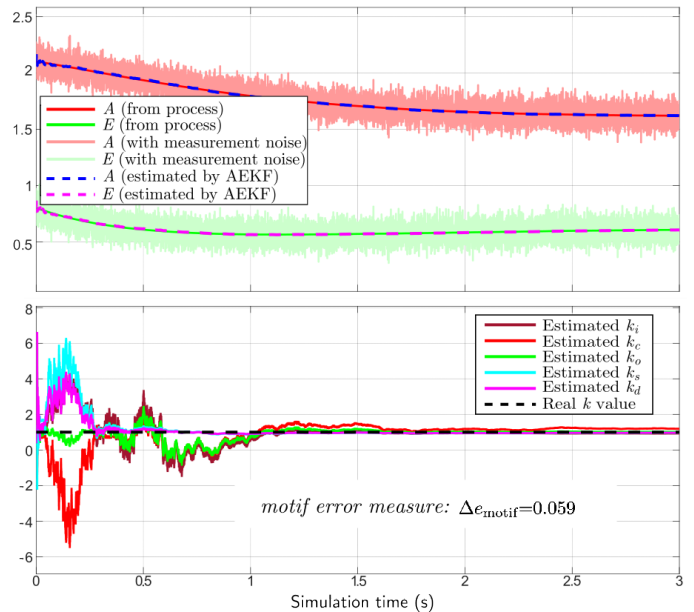


Fig. 3. Simulation results for motif 1 with process and measurement noise. The upper panel shows the states  $A$  and  $E$ , whereas the lower panel shows the parameter estimates.

The upper panel of Fig. 3 shows that the noise is quite substantial, whereas the lower panel demonstrate that even though several parameters deviate up to 500% from the initial condition during the first 0.5 seconds, all parameters still converge to within 10% of the process value at steady state. Note that negative parameter values implies that the species flow in the motif is reversed. To summarize, we

conclude that the AEKF for motif 1 is quite robust against process and measurement noise.

### 5.3 Motif 3

Similarly to motif 1, we performed simulations with and without noise for motif 3. The noise free simulations (not shown) revealed slightly slower dynamics compared to motif 1, but with the same overall *motif error measure* of  $\Delta e_{\text{motif}}=0.0002$ . On the other hand, when both process noise and measurement noise were added, the AEKF showed poor performance already at a noise power level of  $1 \cdot 10^{-11}$ , and estimation breaks down at a noise power of  $1 \cdot 10^{-10}$  with an overall *motif error measure* of  $\Delta e_{\text{motif}}=0.5$ , see Fig. 4. Thus, the AEKF is more sensitive to noise for motif 3 compared to motif 1, and we attribute this to the increased non-linearity from the saturable inhibition term  $\frac{K_i^A}{K_i^A + A}$  in Eq. (6).

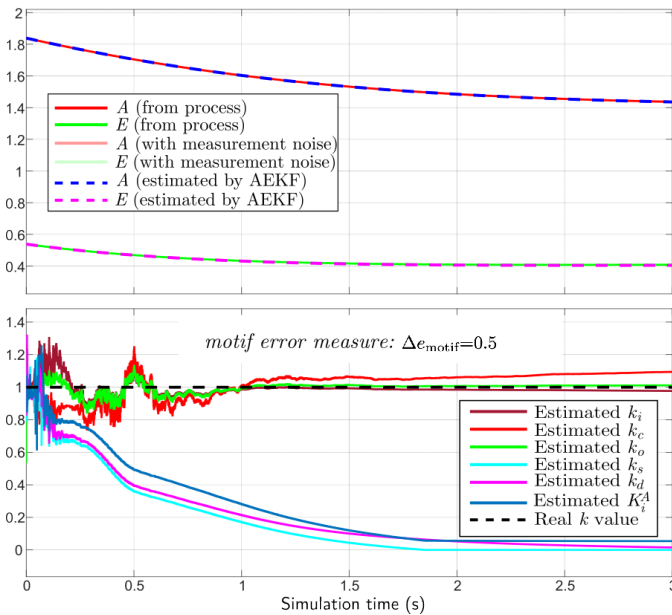


Fig. 4. Simulation results for motif 3 with process and measurement noise. The upper panel shows the states  $A$  and  $E$ , whereas the lower panel shows the parameter estimates.

### 5.4 Noise and initial condition of parameters

To investigate the effect of noise in the performance of the AEKF in more detail, we simulated both motifs while the noise power level was increased from  $1 \cdot 10^{-15}$  to  $1 \cdot 10^{-3}$ , in steps of factor 10. To compensate for the randomly chosen initial conditions for the parameters (still within 1% deviation of  $k_{n,process}$ ), we averaged the results over 10 simulations for each noise power value. We performed simulations with *i*) only process noise, *ii*) only measurement noise, and *iii*) both types of noise. As it turned out that the results were more or less similar for all three cases, we present in Fig. 5 the results for case *iii*) only.

Panels A and B show the *averaged parameter error measure* for motif 1 and 3, respectively, calculated as

$$\overline{\Delta e_n} = \frac{1}{10} \sum_{i=1}^{10} \Delta e_{n,i} \quad (30)$$

where  $\Delta e_{n,i} \forall i \in 1, \dots, 10$  is based on the *parameter error measure* from Eq. (28) calculated during the 10 simulations. The results clearly show that motif 3 is more sensitive to noise than motif 1. Moreover, the profile of each  $\overline{\Delta e_n}$  curve (for each parameter) is more consistent for motif 1, whereas the profiles for motif 3 are more diverse where we see that the parameters  $k_s$ ,  $k_d$ , and  $K_i^A$  are more prone to low noise levels, while  $k_c$  and  $k_o$  are more prone to high noise levels.

We also investigated how the initial conditions of the parameter values in the AEKF affected the overall *motif error measure*  $\Delta e_{\text{motif}}$  from Eq. (29) in noise free simulations. Thus, we performed simulations where the initial condition for each parameter  $k_n$  in the AEKF were increased in steps from the true value of 1 up to 10. These results are shown in panels C and D in Fig. 5 for motif 1 and motif 3, respectively, and we see that the AEKF for motif 1 is relatively unaffected by the increasing initial conditions. Actually, an initial condition 10 times the real parameter value results in  $\Delta e_n \approx 10^{-2}$  for most of the parameters. Interestingly, the curve for parameter  $k_c$  have a surprisingly odd shape where  $\Delta e_n$  have a distinct decreased value for an initial condition of 5 times the real parameter value. On the other hand, and in accordance with panel B, the results for motif 3 show relatively poor performance for increased initial conditions. The performance drops significantly already at a small deviation of only a few percent from the real value and becomes increasingly worse as the initial conditions increase.

To summarize, the AEKF show better performance for motif 1 compared to motif 3, both for increasing noise levels and increasing initial conditions for the parameters. The cause of this is likely that motif 3 is more non-linear than motif 1, as although the added noise is initially zero-mean white noise, it is not guaranteed to be zero-mean after being processed through the nonlinear system. Thus, the noise creates a bias which increases with the degree of non-linearity. Furthermore, as the AEKF use a linearized  $\mathcal{A}(k)$  matrix based on the current state estimate together with the parameter values with initial conditions far away from the real values, it is most likely to introduce a bias which the Kalman filter is unable to compensate for.

### 5.5 Initial conditions of A and E

Finally, we examined how the level of excitation of the process affected the performance of the AEKF, i.e. the level of the initial conditions of the states  $A$  and  $E$  in both the process and the filter. We adjusted the initial conditions for both states simultaneously with a factor  $\alpha$  between  $0.2 < \alpha < 2$ , and simulated both without noise and with a noise power of  $1 \cdot 10^{-6}$  in both the process and measurement noise. The results are shown in Fig. 6 for motif 1, where  $\Delta e_{\text{motif}}$  is plotted as a function of  $\alpha$  for both noise free and noisy conditions.

The blue curve in Fig. 6 show that under noise-free conditions only a small excitation in the initial conditions of  $A$  and  $E$  is necessary to reduce the value of  $\Delta e_{\text{motif}}$ .

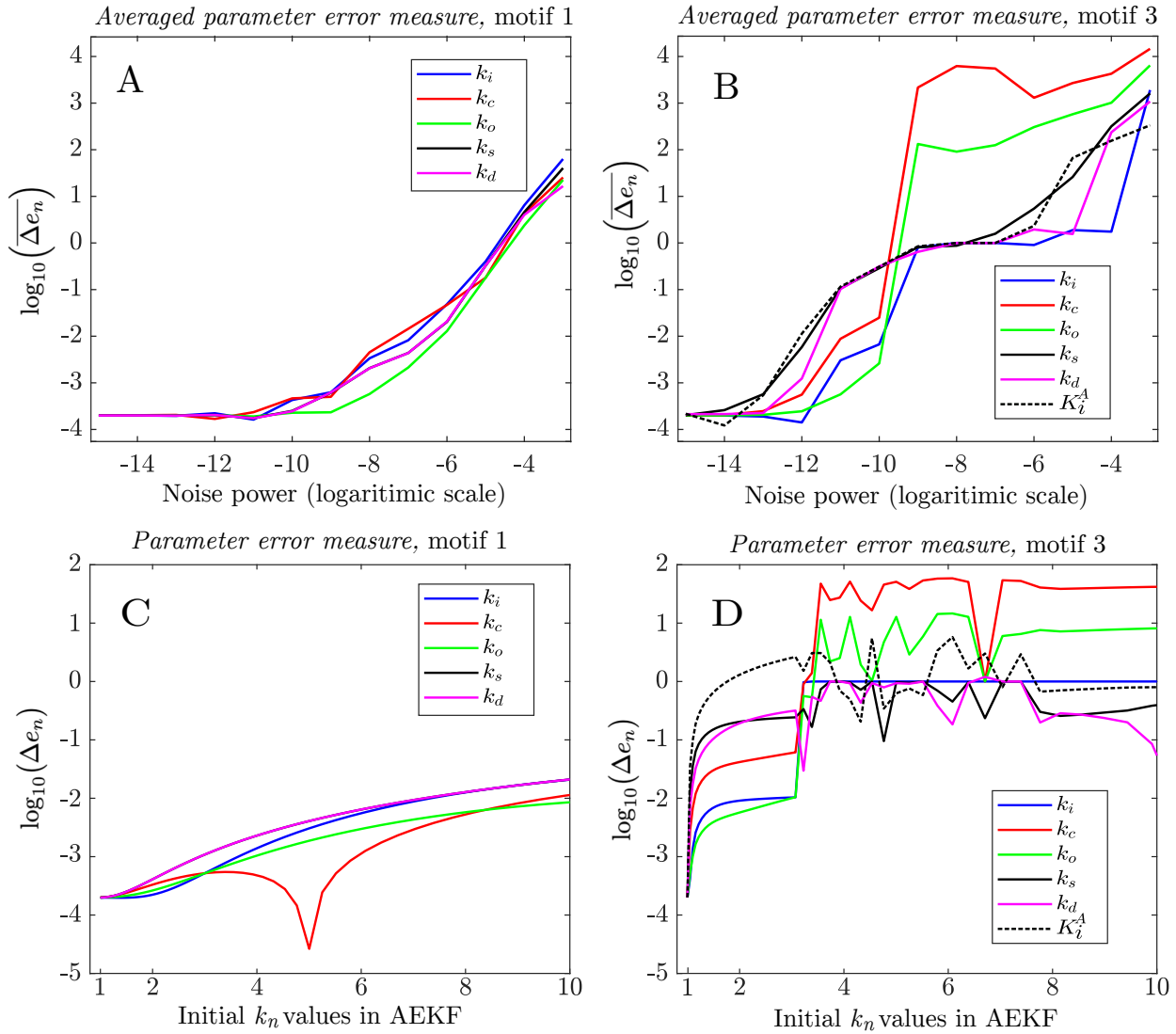


Fig. 5. Panels A/B show the averaged parameter error measure  $\overline{\Delta e_n}$  from Eq. (30) as a function of noise power for motif 1 / motif 3. Panels C/D show the parameter error measure  $\Delta e_n$  from Eq. (28) as a function of initial conditions for the parameters in the AEKFs for motif 1 / motif 3.

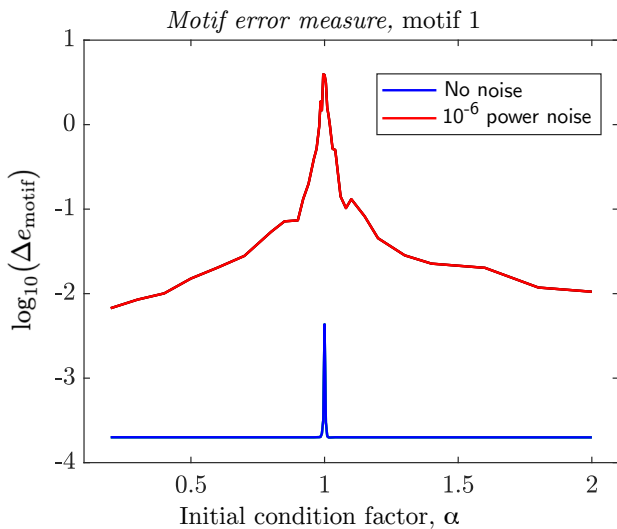


Fig. 6. Plot of the overall motif error measure  $\Delta e_{\text{motif}}$  as a function of initial condition factor  $\alpha$  for both noise free and noisy conditions.

In fact, an increase in  $\alpha$  from  $\alpha=1$  to  $\alpha=1.02$  reduces  $\Delta e_{\text{motif}}$  from 0.007 to approximately 0.0002. The noisy situation illustrated by the red curve show the same underlying mechanisms where increased excitation in initial conditions of  $A$  and  $E$  reduces the value of  $\Delta e_{\text{motif}}$ , but the decrease is not as prominent and the overall level is higher than for the noise free situation. The results are reasonable since decreased excitation decreases the dynamics of  $A$  and  $E$ , leading to an increased noise-to-signal ratio. Thus, at low degrees of excitation the information provided to the Kalman filter by the measurements is predominantly noise.

## 6. CONCLUSIONS

We have implemented an Augmented and Extended Kalman filter (AEKF) able to estimate all unknown parameters for the basic controller motifs 1 and 3 when both states are measured. Under noise free conditions all the parameters were accurately estimated, which corresponds well with previous results showing that these motifs are structurally identifiable with the same measurements. We

also investigated the performance of the AEKF with respect to noise in both the process and measurements, initial conditions of the unknown parameters of AEKF, and finally, the initial conditions of the states  $A$  and  $E$ . We found that the Kalman filter generally performed better on motif 1 than on the more nonlinear motif 3. This suggests that a state observer better suited for nonlinear models, such as the *Unscented Kalman Filter* (Julier and Uhlmann, 1997), may be more appropriate to use for the more nonlinear controller motifs.

## REFERENCES

- Agafonov, O., Selstø, C.H., Thorsen, K., Xu, X.M., Drenstig, T., and Ruoff, P. (2016). The Organization of Controller Motifs Leading to Robust Plant Iron Homeostasis. *PLoS ONE*, 11(1), e0147120.
- Böhler, L., Ritzberger, D., Hametner, C., and Jakubek, S. (2021). Constrained extended kalman filter design and application for on-line state estimation of high-order polymer electrolyte membrane fuel cell systems. *international journal of hydrogen energy*, 46(35), 18604–18614.
- Briat, C., Gupta, A., and Khammash, M. (2016). Anti-thetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Bimolecular Networks. *Cell Systems*, 2(1), 15–26.
- Drenstig, T., Jolma, I.W., Ni, X.Y., Thorsen, K., Xu, X.M., and Ruoff, P. (2012). A Basic Set of Homeostatic Controller Motifs. *Biophysical Journal*, 103(9), 2000–2010.
- Franklin, G.F., Powell, J.D., Workman, M.L. (1998). *Digital control of dynamic systems*, Third edition, Addison-Wesley Reading, MA.
- Haus, E.S., Drenstig, T., and Thorsen, K. (2023). Structural identifiability of biomolecular controller motifs with and without flow measurements as model output. *PLoS Computational Biology*, 19(8), e1011398.
- Huang, Y., Zhu, F., Jia, G., and Zhang, Y. (2020). A slide window variational adaptive Kalman filter. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(12), 3552–3556.
- Huang, Z., Du, P., Kosterev, D., and Yang, B. (2009). Application of extended kalman filter techniques for dynamic model parameter calibration. In *2009 IEEE Power & Energy Society General Meeting*, 1–8. IEEE.
- Julier, S.J. and Uhlmann, J.K. (1997). New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, volume 3068, 182–193. Spie.
- Kalman, R. (1960). On the General Theory of Control Systems. *IFAC Proceedings Volumes*, 1(1), 491–502.
- Liu, J. and Guo, G. (2021). Vehicle localization during GPS outages with extended Kalman filter and deep learning. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–10.
- Ljung, L. and Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2), 265–276.
- Loron, L. and Laliberte, G. (1993). Application of the extended kalman filter to parameters estimation of induction motors. In *1993 Fifth European Conference on Power Electronics and Applications*, 85–90. IET.
- Miao, H., Xia, X., Perelson, A.S., and Wu, H. (2011). On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. *SIAM Review*, 53(1), 3–39.
- Narayanan, H., Behle, L., Luna, M.F., Sokolov, M., Guillén-Gosálbez, G., Morbidelli, M., and Butté, A. (2020). Hybrid-ekf: Hybrid model coupled with extended kalman filter for real-time monitoring and control of mammalian cell culture. *Biotechnology and Bioengineering*, 117(9), 2703–2714.
- Patton, R.J. and Klinkhieo, S. (2009). Actuator fault estimation and compensation based on an augmented state observer approach. In *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, 8482–8487. IEEE.
- Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15), 1923–1929.
- Rotondo, D. (2023). The Weighted Kalman Filter. *IFAC-PapersOnLine*, 56(2), 8869–8874.
- Rotondo, D., Buciakowski, M., and Witczak, M. (2021). Simultaneous state and process fault estimation in linear parameter varying systems using robust quadratic parameter varying observers. *International Journal of Robust and Nonlinear Control*, 31(17), 8390–8407.
- Sarmavuori, J. and Sarkka, S. (2011). Fourier-hermite Kalman filter. *IEEE Transactions on Automatic Control*, 57(6), 1511–1515.
- Tang, Z.F. and McMillen, D.R. (2016). Design principles for the analysis and construction of robustly homeostatic biological networks. *Journal of Theoretical Biology*, 408, 274–289.
- Thorsen, K., Ruoff, P., and Drenstig, T. (2013). Control theoretic properties of physiological controller motifs. *ICSSE 2013 - IEEE International Conference on System Science and Engineering, Proceedings*, 165–170.
- Villaverde, A.F. (2019). Observability and Structural Identifiability of Nonlinear Biological Systems. *Complexity*, 2019.
- Villaverde, A.F., Barreiro, A., and Papachristodoulou, A. (2016). Structural Identifiability of Dynamic Systems Biology Models. *PLoS Computational Biology*, 12(10), e1005153.
- Villaverde, A.F., Evans, N.D., Chappell, M.J., and Banga, J.R. (2018). Sufficiently Exciting Inputs for Structurally Identifiable Systems Biology Models. *IFAC-PapersOnLine*, 51(19), 16–19.
- Zhang, W., Wang, Z., Raïssi, T., Wang, Y., and Shen, Y. (2020). A state augmentation approach to interval fault estimation for descriptor systems. *European Journal of Control*, 51, 19–29.



# A Novel Approach to Simulating the Performance of Autonomous Inflow Control Devices

Ismail Hossain Rafi, Ali Moradi, Soheila Taghavi, Britt Margrethe Emilie Moldestad

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway.

{[ismailrafi16@gmail.com](mailto:ismailrafi16@gmail.com), [ali.moradi@usn.no](mailto:ali.moradi@usn.no), [soheila.t.hosnaroudi@usn.no](mailto:soheila.t.hosnaroudi@usn.no), [britt.moldestad@usn.no](mailto:britt.moldestad@usn.no)}

**Abstract:** Improving the efficiency of oil recovery is a crucial necessity in the current energy landscape. The widespread adoption of advanced wells, equipped with Autonomous Inflow Control Devices (AICDs), represents a leading strategy for this purpose. However, the absence of a predefined and straightforward option for modeling advanced wells in dynamic multiphase flow simulators like OLGA® poses a significant challenge. To address the issue, this paper proposes a novel approach based on developing a mathematical model derived from experimental data characterizing the AICD behavior. The Algebraic Controller option in OLGA is then leveraged to integrate the AICD effects into the simulation seamlessly. The proposed methodology undergoes rigorous testing on the PUNQ-S3 reservoir model as a benchmark case study with Water Alternating Gas (WAG) injection. Results demonstrate that AICD has a better water reduction rate of 36.3% and 3.7% compared to OPENHOLE and ICD. This result also indicates the accurate modeling and simulation of AICD performance in the software, showcasing the effectiveness of the developed mathematical model. Comparative analyses of advanced wells with different Flow Control Devices (FCDs) underscore the conclusion that AICDs significantly enhance oil recovery efficiency, thereby maximizing profit and minimizing the carbon footprint.

**Keywords:** OLGA, Near-well simulation, Advanced wells, AICD, Algebraic controller.

## 1. INTRODUCTION

Oil has been the most consumed energy among all the energy sources. Oil recovery has to be maximized considering the economic and environmental effects. Preventing early water and gas breakthroughs in horizontal wells is a major challenge in the oil industry. Inflow control technology like inflow control devices (ICD) and autonomous inflow control devices (AICD) were invented to minimize the issue of breakthroughs of unwanted fluid. AICD valve opening control is based on the properties of different fluids. Accurate modeling of the AICD behavior using a dynamic multiphase flow simulator like OLGA is important but challenging. Previous researchers tried to use the PID controller and Table controller to model the behavior of AICD. PID controller acts on a fixed setpoint, and Table controller is applicable for two-phase fluid mixtures like oil with gas or oil with water. Another approach is needed to consider the three-phase fluid mixture, a solution is to use the algebraic controller feature in the OLGA simulator. A logical or mathematical equation can be used as an expression form to control the valve opening output of AICD.

## 2. INFLOW CONTROL TECHNOLOGIES

### 2.1 Inflow Control Devices (ICD)

ICD was invented in early 1990 by Norsk Hydro on the horizontal well section in the Troll field (Al-Khelaiwi and D.R., 2007). Due to reservoir heterogeneity, early water or gas breakthroughs can occur in the heel section or in high permeability zones. ICDs are mounted on the production tube as shown in Fig. 1. ICDs minimize the fluid flow with an

additional pressure drop to create an even flow distribution along the horizontal well.

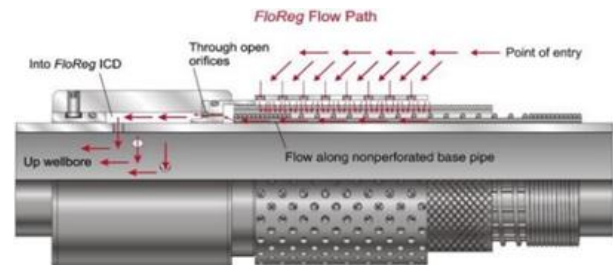


Fig. 1. Orifice-type ICD setup and flow pattern of fluids (Birchenko et al., 2010).

However, ICD has a disadvantage as it cannot choke back the water after a breakthrough into the production pipe has occurred. For this reason, the water cut rises more than the capacity of the separation facilities can handle. The whole well needs to be choked to avoid this higher water cut. Choking the well results in minimizing oil production (Moradi and Moldestad, 2020). Many ICDs are mounted on a horizontal well. The pressure drop is the function of the flow rate, ICD geometry, and fluid density. In most cases, the orifice-type ICD is used, and the mathematical equation is:

$$\dot{Q} = C_D A \sqrt{\frac{2\Delta P}{\rho}}, \quad (1)$$

Where  $\dot{Q}$  is the volumetric flow rate of the fluid through the ICD,  $\Delta P$  is the pressure drop over the ICD,  $\rho$  is the fluid density,  $A$  is the cross-sectional area of the ICD nozzle,  $C_D$  is the discharge coefficient.

## 2.2 Autonomous Inflow Control Devices (AICD)

An uneven production flow can still be observed from the toe to the heel section of a horizontal well even when using ICDS. This occurs due to frictional pressure drop and permeability variations in different regions. ICD can delay the breakthrough of unwanted fluids but cannot stop the breakthrough. Among all types of AICD, the AICD developed by Statoil is the most commonly used. A schematic of an RCP version of AICD is shown in Fig 2. AICD is functional with the viscosity differentials of the fluids. The RCP valve reduces the flow of low-viscous fluids like water and gas and is fully open for high-viscous fluids like oil (Mathiesen et al., 2011).

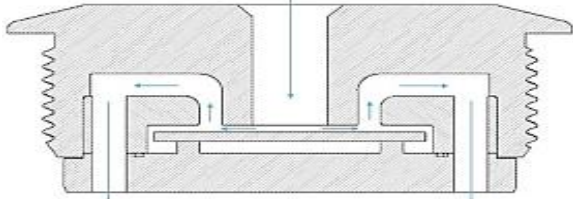


Fig. 2. Schematic of the RCP valve developed by Statoil (Mathiesen et al., 2011).

Taking the function of fluid properties and volume flow empirical equation of differential pressure is:

$$\Delta P = f(\rho, \mu) \cdot a_{AICD} \cdot \dot{Q}^x, \quad (2)$$

$$f(\rho, \mu) = \left( \frac{\rho_{mix}^2}{\rho_{cal}} \right) \cdot \left( \frac{\mu_{cal}}{\mu_{mix}} \right)^y, \quad (3)$$

Where,  $\dot{Q}$  is the volumetric flow rate of the fluid through the RCP,  $\Delta P$  is the pressure drop over the AICD.  $a_{AICD}$ ,  $x$ , and  $y$  are the parameters specified by the user depending on the fluid properties and the RCP design criteria.  $f(\rho, \mu)$  is the function of the density and viscosity in which  $\rho_{cal}$  and  $\mu_{cal}$  is the calibrated density and viscosity respectively. The equations for mixture density and viscosity are as follows:

$$\rho_{mix} = \alpha_{oil}\rho_{oil} + \alpha_{water}\rho_{water} + \alpha_{gas}, \quad (4)$$

$$\mu_{mix} = \alpha_{oil}\mu_{oil} + \alpha_{water}\mu_{water} + \alpha_{gas}\mu_{gas}, \quad (5)$$

Where  $\alpha_{oil}$ ,  $\alpha_{water}$ , and  $\alpha_{gas}$  is the volume fraction of oil, water, and gas in the mixture respectively.

## 3. PUNQ-S3 RESERVOIR MODEL

For this study, the PUNQ-S3 (Production forecasting with Uncertainty Quantification, variant 3) synthetic reservoir model is used and designed in ECLIPSE. Elf Exploration Production implemented this model in the real field according to the reservoir engineering study. It is a three-dimensional dome-shaped heterogeneous reservoir containing a total of 2660 grid blocks, of which 1761 blocks are active. The dimensions of the reservoir are given in Table 1. Corner point geometry and the Carter-Tracey aquifer were used to design this reservoir (Hutahaean, 2017). It has a Bottom Hole Pressure (BHP) of 220 bar with a maximum liquid production rate of 4000 m<sup>3</sup>/day.

Table 1: PUNQ-S3 reservoir grid dimensions.

Direction	No. of blocks	Length (m)/dip angle
x	19	19×180
y	28	28×180
z	5	2355/1.5°

The production well and the four injectors are designed by trial-and-error method for more oil production and to minimize early water and gas breakthroughs. Water and CO<sub>2</sub> are simultaneously injected by the four injectors at a regular time interval. Injectors 1, 2, 3, and 4 are placed at the depth of 2390 m, 2375 m, 2370 m, and 2370 m respectively. Figure 3 represents the positioning of the injectors and production pipe with top face depth. The length of the horizontal well is 3240 m which is designed in the OLG simulator.

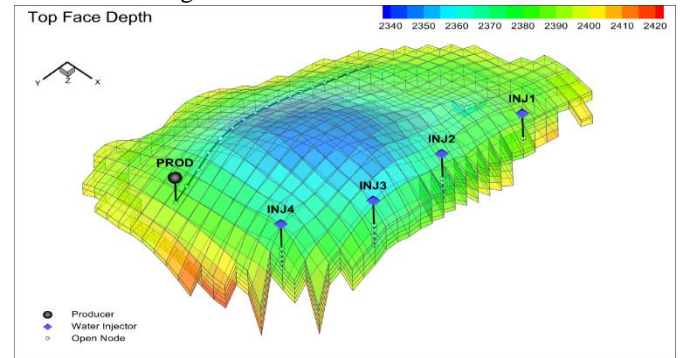


Fig. 3. Production pipe and injectors topology.

Table 2 shows the rock and fluid properties of the reservoir for the simulation cases and Figure 4 shows the porosity and permeability in different directions.

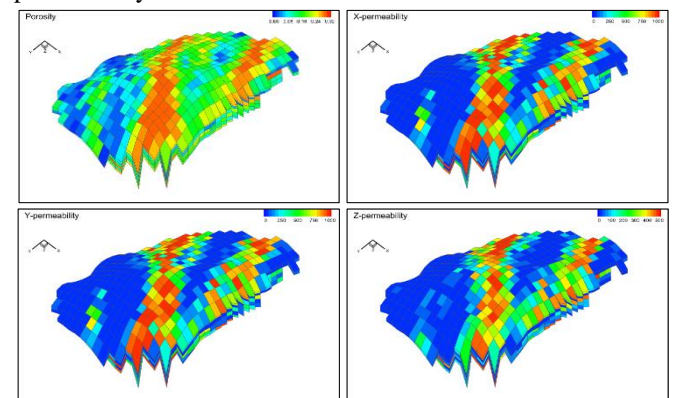


Fig. 4. Porosity and permeability.

Table 2. Rock and fluid properties of the PUNQ-S3 reservoir

Parameter	Value
Oil density	912 kg/m <sup>3</sup>
Water density	1000 kg/m <sup>3</sup>
Gas density	0.8266 kg/m <sup>3</sup>
GOR	74 Sm <sup>3</sup> / Sm <sup>3</sup>
Reservoir pressure	234.5 bar
Temperature	105 °C
Water viscosity (reservoir condition)	0.5 cP
Oil viscosity (reservoir condition)	1.46 cP
Gas viscosity (reservoir condition)	0.0133 cP
Porosity	0.1 – 0.3
Mean porosity	0.14
Rock compressibility	0.000451/bar

#### 4. WELL MODEL IN OLGA

For the advanced horizontal well model, the length is specified as 3240 m in OLGA. The wellbore and production pipe have a diameter of 0.2159 m and 0.1397 m respectively. The production well has 18 valves to divide the production pipe into 18 zones. Each zone is 180 m long and separated by two packers. In reality, each section is about 12 m and consists of one flow control device (FCD). For 180 m 15 FCDs are required. In this study, each FCD is considered equivalent to 15 real FCDs. The equivalent diameter is taken as 0.0078 m for both ICD and AICD considering the discharge coefficient (CD) as 0.85. Figure 5 shows a simplified sketch of a single production zone in a production pipe containing packers, a fluid flow path, and an inflow control device. Packers prevent fluid flow from an adjacent zone through the annulus. Near-well source is the connecting component between OLGA and ECLIPSE. Through section I fluid enters into the wellbore and then passes through the inflow control devices. After that fluid passes through the Leak into the production pipe in section II. This method was proposed by Haavard Akre in 2012 and is much used in simulation studies (Moradi and Moldestad, 2020).

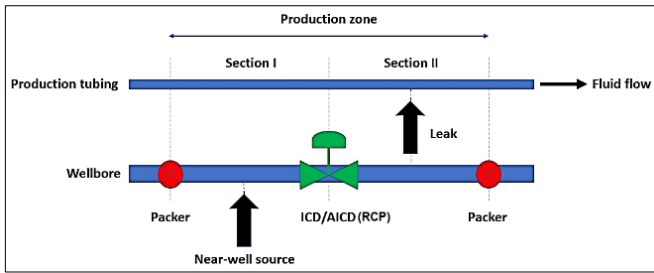


Fig. 5. Schematic of a single zone in a production pipe (Moradi and Moldestad, 2020).

#### 5. ALGEBRAIC CONTROLLER

The algebraic controller is a feature of OLGA for implementing algebraic equations or logical expressions to manipulate input signals for a desired output. In this study, the algebraic controller is used to control the valve opening of AICD. A mathematical equation is derived as the input signal for the AICD valve considering the oil, water, and gas volume fractions.

For ICD the pressure differential and flow rate can be written as:

$$\Delta P_{ICD} = \hat{C}_u \frac{\rho_{mix} \dot{Q}_{ICD}^2}{2\gamma^2 A^2 C_D^2}, \quad (6)$$

$$\dot{Q}_{ICD} = \gamma A C_D \sqrt{\frac{2\Delta P_{ICD}}{\rho_{mix} \hat{C}_u}}, \quad (7)$$

The pressure differential was derived from the available experimental data in (Halvorsen Martin et al., 2016) for AICD in the PUNQ-S3 reservoir model with similar fluid properties. Both linear and non-linear regression method was used to develop the mathematical model which is expressed by:

$$\Delta P_{AICD} = a_{AICD} \cdot \frac{\rho_{mix}^2}{1000} \cdot \left(\frac{1}{\mu_{mix}}\right)^y \cdot \dot{Q}_{AICD}^x, \quad (8)$$

For AICD flow rate can be expressed by:

$$\dot{Q}_{AICD} = \left(\frac{1000 \cdot \Delta P_{AICD} \cdot \mu_{mix}^y}{a_{AICD} \cdot \rho_{mix}^2}\right)^{\frac{1}{x}}, \quad (9)$$

Here  $\dot{Q}$  is the volumetric flow rate,  $\Delta P$  is the pressure drops,  $A$  is the cross-sectional area of the fluid flow,  $C_D$  is the discharge coefficient,  $\hat{C}_u$  is the unit conversion value, and  $\gamma$  is the valve opening. Three-phase fluid mixture density and viscosity can be written as (4) and (5).

$\alpha_{oil}$ ,  $\alpha_{water}$ ,  $\alpha_{gas}$  are the volume fractions of oil, water, and gas in the mixture and the sum of the fractions is:

$$\alpha_{oil} + \alpha_{water} + \alpha_{gas} = 1$$

Now, matching  $\Delta P - \dot{Q}$  curves of ICD and AICD at  $\Delta P_{match}$  and  $\dot{Q}_{match}$  it can be assumed that

$$\Delta P_{ICD} = \Delta P_{AICD} = \Delta P_{mix} \quad \text{and} \quad \dot{Q}_{ICD} = \dot{Q}_{AICD}$$

Considering  $\dot{Q}_{ICD} = \dot{Q}_{AICD}$  the valve opening can be expressed as:

$$\gamma = \frac{\left(\frac{1000 \cdot \Delta P_{mix}}{a_{AICD}}\right)^{\frac{1}{x}}}{A C_D \sqrt{\frac{2\Delta P_{mix}}{\hat{C}_u}}} \cdot \mu_{mix}^{\frac{y}{x}} \cdot \rho_{mix}^{\frac{x-4}{2x}}, \quad (10)$$

Alternately it can be expressed as:

$$\gamma = \beta \cdot \left\{ \alpha_{oil} \mu_{oil} + \alpha_{water} \mu_{water} + \alpha_{gas} \mu_{gas} \right\}^{\frac{y}{x}} \cdot \left\{ \alpha_{oil} \rho_{oil} + \alpha_{water} \rho_{water} + \alpha_{gas} \rho_{gas} \right\}^{\frac{x-4}{2x}}, \quad (11)$$

where

$$\beta = \frac{\left(\frac{1000 \cdot \Delta P_{mix}}{a_{AICD}}\right)^{\frac{1}{x}}}{A C_D \sqrt{\frac{2\Delta P_{mix}}{\hat{C}_u}}}, \quad (12)$$

and

$$\left\{ \alpha_{oil} \mu_{oil} + \alpha_{water} \mu_{water} + \alpha_{gas} \mu_{gas} \right\}^{\frac{y}{x}} = \mu_{mix}^{\frac{y}{x}} \quad \left\{ \alpha_{oil} \rho_{oil} + \alpha_{water} \rho_{water} + \alpha_{gas} \rho_{gas} \right\}^{\frac{x-4}{2x}} = \rho_{mix}^{\frac{x-4}{2x}}, \quad (13)$$

In the OLGA model, two transmitters are used to take the values of the volume fraction of oil ( $\alpha_{oil}$ ) and water ( $\alpha_{water}$ ) as an input variable from the wellbore. Equation (8) is put as an expression option in the algebraic controller in OLGA. To implement  $\alpha_{oil}$ ,  $\alpha_{water}$  as input variables into (14), they are introduced as unknown variables X1 and X2 in the algebraic controller. The expression in the algebraic controller is as follows:

$$\gamma = \beta \cdot \left\{ X1 \cdot \mu_{oil} + X2 \cdot \mu_{water} + (1 - X1 - X2) \cdot \mu_{gas} \right\}^{\frac{y}{x}} \cdot \left\{ X1 \cdot \rho_{oil} + X2 \cdot \rho_{water} + (1 - X1 - X2) \cdot \rho_{gas} \right\}^{\frac{x-4}{2x}}, \quad (14)$$

Here  $\alpha_{oil} = X1$ ,  $\alpha_{water} = X2$  and  $\alpha_{gas} = 1 - X1 - X2$

Experimental data were used to find the parameters in (12) using multivariable nonlinear regression. The values in Appendix A are based on that. Figure 6 shows the controller behavior as the valve opening control for AICD according to different phases of the fluids.



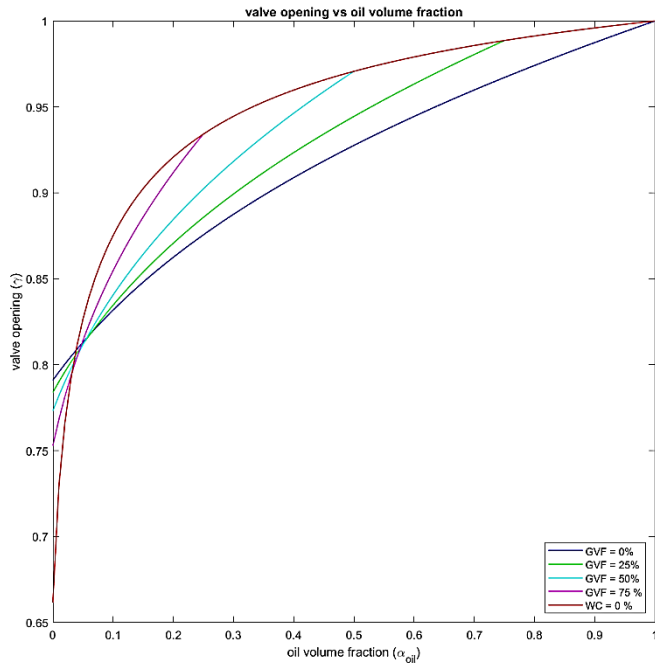


Fig. 6. Valve opening vs oil volume fraction for the algebraic controller in different phases of fluid.

### 6. RESULTS AND DISCUSSION

In this chapter, the obtained simulation results from the OLGA/ECLIPSE model are shown and discussed. Performance analysis of ICD and AICD are shown and compared to the OPENHOLE case for improving oil recovery and reducing water cut. The functionality of the algebraic controller controlling the valve opening of AICD is analyzed considering water cut and gas volume fraction (GVF). The pressure drop for the cases was considered 15 bar with a constrained liquid production rate of 4000m<sup>3</sup>/day for ICD.

#### 6.1 Total oil and water production

The total flow of oil and water are two of the most important parameters to analyze for the performance of the FCDs. Figure 7 illustrates the total oil and water production for OPENHOLE and FCDs. OPENHOLE has a larger cross-sectional inlet area and there are no restrictions for liquid production so more water and oil can be produced compared to FCDs. ICD and AICD had the same amount of oil production, and both of the curves overlapped (blue over black), and this can happen because of the recovery of low viscous oil. The functionality of AICD can be observed with higher oil production if the simulation time was more than 3500 days. It is very important to have less water production for a better economy and environmental impact. In this case, FCDs showed better performance producing less water than OPENHOLE. In the cases of ICD and AICD, the water production is reduced by 33.8% and 36.3% respectively compared to the OPENHOLE case. The AICD reduced the accumulated water production by 3.7% compared to the ICD. This indicates that AICD has a better choking effect on low viscous fluid like water. Less water production means less production cost in the oil processing step after recovery.

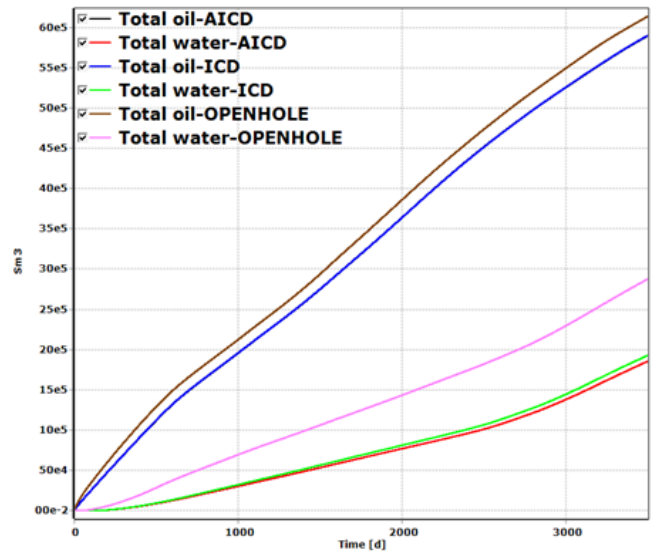


Fig. 7. Total oil and water production.

#### 6.2 Oil and water production rate

OPENHOLE has a larger cross-sectional area to produce more liquid than FCDs. Figure 8 illustrates the oil and water production rate for 3500 days of simulation.

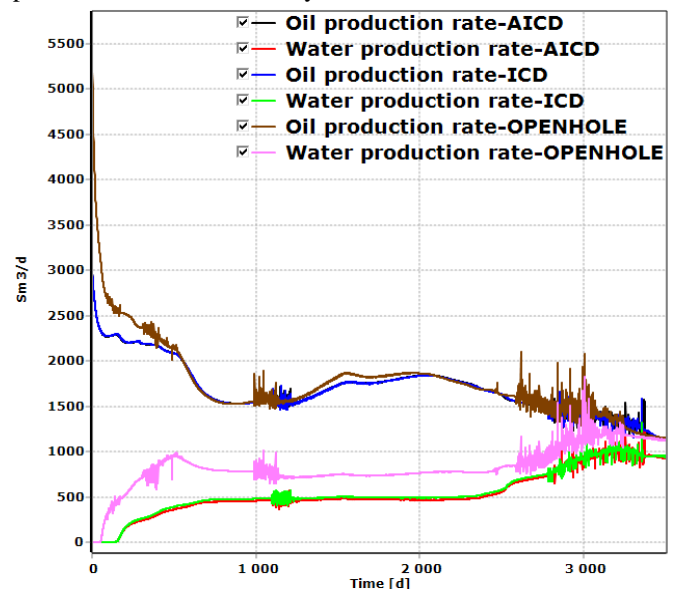


Fig. 8. Oil and water production rate.

The OPENHOLE case has a larger cross-sectional area which is exposed to production. For this reason, OPENHOLE has a higher fluid production rate. In reality, there is a limitation for maximum fluid production with a regulating valve on top side, but this is not considered for this study. After 1000 days the oil production rate increased because of gas injection. It can be observed that after 3500 days FCDs are showing a tendency to produce more oil than OPENHOLE. So, simulation for more than 3500 days can result in a higher oil production rate for the FCDs. ICD and AICD almost have the exact amount of oil showing the blue curve overlapping over the black curve. But ICD has a higher production rate of water than AICD. This is because ICDs cannot prevent further production after the water enters the well whereas AICD is partially closed when water

enters through the inlet. This also proves the choking ability of AICD to the low viscous fluids.

### 6.3 Algebraic controller behavior to WC

In this study algebraic controller is used to control the valve opening of the AICD. Equation (12) was implemented by expression form in the algebraic controller. Figure 9 shows the valve opening control according to the WC. A total of 18 controllers were used in the well model. Among those, controller 1 and controller 18 are chosen for the toe and heel respectively. A maximum water cut of 0.64 and 0.82 was found for the toe and heel sections. At the toe section, the minimum valve opening was 90% fully open and at the heel section, the minimum valve opening was 85%. It can be observed that the more the water cut increased the more the valve was closing. From this observation, it can be said that the algebraic controller is showing the choking effect on water production.

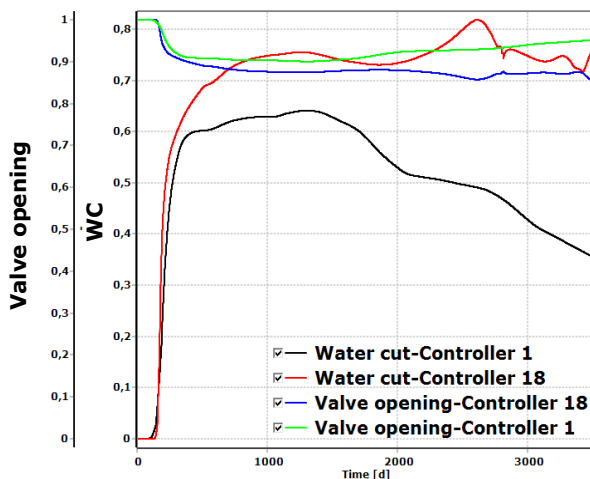


Fig. 9. Valve opening with algebraic controller in toe and heel section for water cut.

### 6.4 Algebraic controller behavior to GVF

To observe the functionality of valve opening of the algebraic controller according to the GVF, controllers 8 and 9 at the middle of the horizontal well were selected. Figure 10 shows the behavior of the algebraic controller for the GVF.

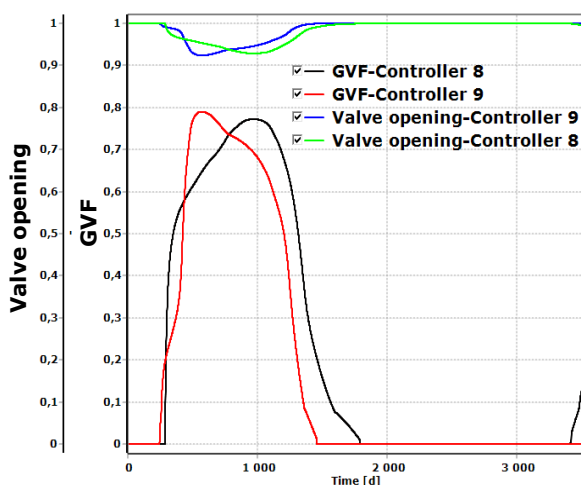


Fig. 10. Valve opening of the algebraic controller according to GVF.

A maximum GVF of 0.77 and 0.79 was found for the controllers 8 and 9 respectively. For controller 8 the valve opening was up to 0.93 and for controller 9 valve opening was up to 0.92. It can be observed that the more the GVF increased the more the valve was getting closed. From this observation, it can be said that the algebraic controller is showing the choking effect on gas volume fraction.

## 7. CONCLUSIONS

According to the findings from the simulations, it can be concluded that FCDs show a better impact on the WAG injection oil recovery process in heterogeneous reservoir. ICD and AICD have reduced water production by 33.8% and 36.3% respectively compared to OPENHOLE. The most important part of this study was to implement and investigate the performance of the algebraic controller in terms of controlling the valve opening of AICD. Though the oil production rate was the same for both ICD and AICD, water production was 3.7% lower for AICD compared to ICD. This satisfies the main purpose of using AICD to minimize water production. It also indicates the performance of the algebraic controller that can be implemented for AICD valve opening in the OLGA simulator. It showed better performance in controlling valve opening with increasing WC and GVF. Using the transmitters for getting input of oil and water volume fractions to the controller and manipulating the valve opening from a logical mathematical expression was also successful.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the economic support from the Research Council of Norway and Equinor through Research Council Project No. 308817, "Digital Wells for Optimal Production and Drainage" (DigiWell).

## REFERENCES

- Al-Khelaiwi, F.T., D.R., D. (2007). Inflow Control Devices: Application and Value Quantification of a Developing Technology. International Oil Conference and Exhibition. doi:10.2118/108700-MS
- Birchenko, V.M., Muradov, K.M., Davies, D.R. (2010). Reduction of the horizontal well's heel-toe effect with inflow control devices. J Pet Sci Eng 75, 244–250. doi:10.1016/j.petrol.2010.11.013
- Halvorsen, Martin, Madsen, Martin, Vikøren, Mo Mathias, Isma Mohd, Ismail, Green, Annabel (2016). Enhanced Oil Recovery On Troll Field By Implementing Autonomous Inflow Control Device. Annabel. doi:10.2118/180037-MS
- Hutahaean, J.J.J. (2017). Multi-Objective Methods for History Matching, Uncertainty Prediction and Optimisation in Reservoir Modelling. Heriot-Watt University.
- Mathiesen, V., Aakre, H., Werswick, B., Elseth, G. (2011). The Autonomous RCP Valve-New Technology for Inflow Control In Horizontal Wells, in: SPE Offshore Europe Oil and Gas Conference and Exhibition. Aberdeen, UK. doi:10.2118/145737-MS
- Moradi, A., Moldestad, B.M.E. (2020). Near-well simulation of oil production from a horizontal well with ICD and AICD completions in the Johan Sverdrup field using



OLGA/ROCX, in: Proceedings of The 61st SIMS Conference on Simulation and Modelling SIMS 2020, September 22-24, Virtual Conference, Finland. Linköping University Electronic Press, pp. 249–256. doi:10.3384/ecp20176249

## Appendix A

Parameter	Value
$C_D$	0.85
$\dot{C}_u$	$1.34 \cdot e^{-15}$
$a_{AICD}$	$3.41 \cdot e^{-6}$
$\Delta P_{mix}$	20 bar
$A = \frac{\pi}{4} D^2$	$\frac{\pi}{4} (0.002)^2 = 3.2687 \cdot e^{-5}$
$x$	3.35
$y$	0.4
$\mu_{oil}$	2.7
$\mu_{water}$	0.45
$\mu_{gas}$	0.02
$\rho_{oil}$	890 kg/m <sup>3</sup>
$\rho_{water}$	1100 kg/m <sup>3</sup>
$\mu_{gas}$	150 kg/m <sup>3</sup>

# Integration of Dynamic Multiphase Flow and Reservoir Models for Improved Oil Recovery Simulation

Charith Rajapaksha, Ismail Hossain Rafi, Nirajan Raut, Ali Moradi, Soheila Taghavi,  
Britt M. E. Moldestad

Department of Process and Environmental Technology, University of South-Eastern Norway, Norway.  
{[charithpriya@gmail.com](mailto:charithpriya@gmail.com), [ismailrafi16@gmail.com](mailto:ismailrafi16@gmail.com),  
[nirajanraut8@gmail.com](mailto:nirajanraut8@gmail.com), [ali.moradi@usn.no](mailto:ali.moradi@usn.no),  
[soheila.t.hosnaroudi@usn.no](mailto:soheila.t.hosnaroudi@usn.no) , [britt.moldestad@usn.no](mailto:britt.moldestad@usn.no)}

**Abstract:** The utilization of advanced multilateral wells to enhance well-reservoir contact, coupled with water injection, stands out as a common approach to boost oil extraction efficiency. It is imperative to develop precise, fully integrated, dynamic, well-reservoir models tailored for this type of oil recovery to enhance the design of advanced multilateral well completions. This study addresses the challenge by constructing a well model using OLGA®, which is, a dynamic multiphase flow simulator, and a reservoir model using Eclipse™, a reservoir simulator. Subsequently, these models are seamlessly integrated to perform comprehensive simulations. The proposed approach is tested on a case study involving oil recovery through an advanced multilateral well completed with various Flow Control Devices (FCDs) supported by water injection. Results from the simulations demonstrate the success of the integration approach, offering a reliable method for accurately modelling oil recovery from advanced multilateral wells to improve oil recovery. Notably, according to this study, wells completed with Autonomous Inflow Control Valves (AICVs) exhibit superior performance, optimizing oil recovery with a reduced carbon footprint.

**Keywords:** Advanced multilateral wells, Well-reservoir model, Autonomous Inflow Control Valve, Autonomous Inflow Control Devices, Inflow Control Devices.

## 1. INTRODUCTION

Even though the world is moving towards renewable energy, crude oil is still significantly contributing to the world's energy demand. Despite advances, a considerable portion of oil remains unrecovered due to traditional technologies.

The multilateral well model can be implemented to increase the oil recovery, a well-completion technique more suitable for horizontal drilling. This method has several advantages, such as increased reservoir exposure, reduced water and gas coning, accelerated production, connecting high permeability areas, and lower capital costs than constructing single-well systems, although with disadvantages like early water and gas breakthrough (Elyasi, 2016). Figure 1 illustrates the main types of multilateral wells used in industry.

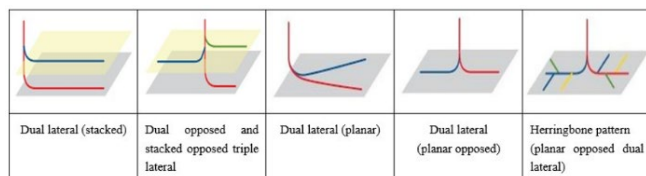


Fig. 1. Types of different multilateral wells (Flatern, 2021).

Methods like polymer and zonal control with flow control devices (FCDs) tackle early breakthroughs and increase oil recovery. Inflow control devices (ICDs), Autonomous Inflow Control Devices (AICDs), and Autonomous Inflow Control

Valves (AICVs) are mainly used as FCDs in the industry, which makes the oil well an advanced well (Aakre *et al.*, 2014).

To design and maintain advanced multilateral wells, a proper simulation and modelling are to be done to decide the parameter values for optimum production. Oil recovery through advanced multilateral wells is a transient process, and the simulation model must capture the transient interaction between the reservoir and the well. Therefore, a dynamic, fully coupled, well-reservoir model is required to simulate oil recovery accurately through advanced wells. Researchers widely use the multisegmented well (MSW) model to simulate advanced wells, but it is a homogeneous model and is not very accurate due to simplifications. Coupling the well model in a dynamic multiphase simulator with a reservoir model can be used to overcome the inaccuracies. This study was conducted to study the coupling well model in the OLGA multiphase simulator with a reservoir model done in ECLIPSE.

## 2. MULTISEGMENT WELL MODEL

The multisegmented well model in ECLIPSE is an advanced extension for accurately modelling fluid behaviour in advanced wells. It divides the production tubing into multiple one-dimensional segments with independent variables to describe the conditions. These variables are determined by solving material balance equations for each phase and

component using different pressures. This approach allows for precisely modelling the relationship between pressure gradients and fluid composition changes in advanced wells (Anuththara *et al.*, 2023).

Figure 2 presents a schematic of an MSW model for an advanced horizontal well. This model treats the production tubing and wellbore as distinct branches of specific segments. Additionally, segments can be designed to simulate Flow Control Devices (FCDs). These FCDs connect the wellbore and production tubing, as depicted in Fig. 2. Fluids enter the wellbore through its segments, pass through FCD valve segments into the production tubing, and then flow to the production outlet via the production tubing segments (Anuththara *et al.*, 2023).

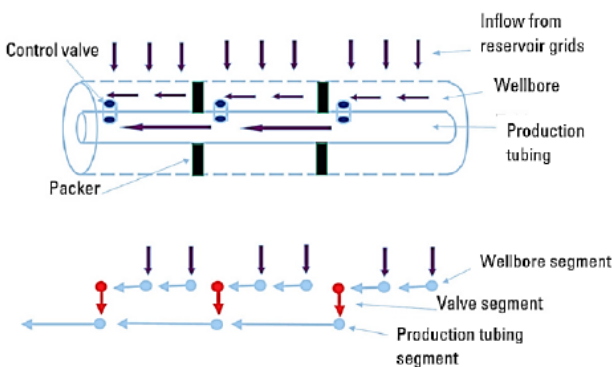


Fig. 2. Schematic of a multisegmented well model (Moradi *et al.*, 2022a).

### 3. FLOW CONTROL TECHNOLOGY

Flow control devices (FCDs) are applied to prevent early gas and water breakthroughs in a well, making it an advanced well. The inflow control technology is essential for improving oil recovery and expanding reservoir lifespan, making oil production economical.

#### 3.1 Inflow control device (ICD)

ICD is a passive inflow control device without any active part that can adjust the flow with the conditions. It limits the flow by giving an additional pressure drop to achieve a distributed flow profile along the horizontal well with the pre-determined design, as shown in Fig. 3,

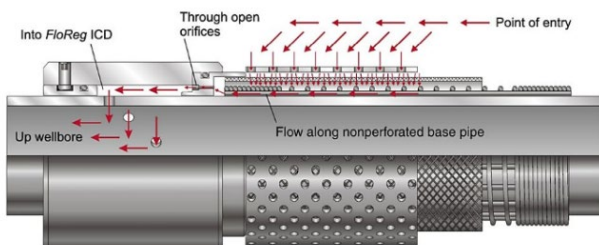


Fig. 3. Orifice-type ICD and its flow path (Mathiesen *et al.*, 2014).

Several ICDs are usually placed along the well tubing, and with an even production rate along the well, the water and gas breakthroughs can be delayed. However, ICDs cannot control the flow after the breakthrough and choke the low viscous

fluids into the production tubing. The governing equation for orifice-type ICD is mentioned as follows (Moradi *et al.*, 2022b),

$$\dot{Q} = C_D A \sqrt{\frac{1}{1-\beta^4}} \sqrt{\frac{2\Delta P}{\rho}} \tag{1}$$

Where,

$\dot{Q}$  is the volume flow rate of the fluid through the ICD

$\Delta P$  is the pressure drop over the ICD

$\rho$  is the fluid density

$A$  is the cross-sectional area of the ICD nozzle

$C_D$  is the discharge coefficient, which depends on the ICD design

#### 3.2 Autonomous inflow control device (AICD)

AICD is an improved version of ICDs that can delay the water and gas breakthroughs and partially be close to low-viscosity fluids like water and gases. AICD has active and passive control elements to produce a pressure drop and control the flow autonomously.

Figure 4 presents a schematic of the rate control production (RCP) type AICD, which consists of a free-floating disc, an outer seat, and an inner seat (Anuththara *et al.*, 2023). According to the pressure, forces acting on the disc will move to control the flow accordingly.

When a low viscous fluid compared to oil flows through the valve, a low pressure will be created in the inlet area due to low friction force, according to Bernoulli's equation. This pressure reduction creates a force that pulls the moving plate toward the inlet, partially closing the valve. This mechanism enables these valves to autonomously reduce the flow rate of unwanted fluids such as water or gas (Moradi *et al.*, 2022b).

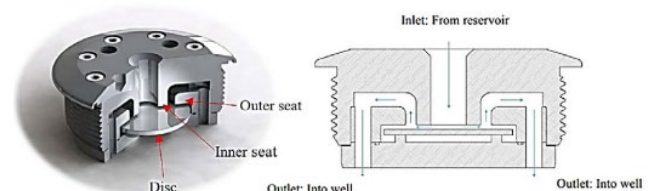


Fig. 4. Schematic diagram of Statoil's RCP valve (Anuththara *et al.*, 2023).

The empirical equation pressure drop across an RCP-type AICD is as follows,

$$\Delta P = a_{AICD} \cdot \left(\frac{\rho_{mix}}{\rho_{cat}}\right) \cdot \left(\frac{\mu_{cat}}{\mu_{mix}}\right)^y \cdot \dot{Q}^x \tag{2}$$

where,

$\dot{Q}$  is the volume flow rate of the fluid through the AICD

$\Delta P$  is the pressure drop over the AICD

$\rho_{mix}$  is the density of the fluid mixture

$\mu_{mix}$  is the viscosity of the fluid mixture

$a_{AICD}$ ,  $x$  and  $y$  are user input parameters that depend on the AICD design and the fluid properties, while  $\rho_{cat}$  and  $\mu_{cat}$  are calibrating parameters.  $\rho_{mix}$  and  $\mu_{mix}$  are calculated as follows,

$$\rho_{mix} = \alpha_{oil}\rho_{oil} + \alpha_{water}\rho_{water} + \alpha_{gas}\rho_{gas} \tag{3}$$

$$\mu_{mix} = \alpha_{oil}\mu_{oil} + \alpha_{water}\mu_{water} + \alpha_{gas}\mu_{gas} , \quad (4)$$

where,

$\alpha_{oil}$  is the volume fraction of oil in the mixture

$\alpha_{water}$  is the volume fraction of water in the mixture

$\alpha_{gas}$  is the volume fraction of gas in the mixture

### 3.3 Autonomous inflow control valve (AICV)

AICV is the latest inflow control device developed by InflowControl AS, which has been claimed to have better choking performance than AICD and ICD. AICV can almost entirely close its opening to unwanted fluids with low viscosities, such as water and gas.

AICV consists of two restrictors where: one is a laminar restrictor, and the other is a turbulent restrictor. Figure 5 illustrates a schematic diagram of an AICV and how the pressure gradients stack where oil gets and lower pressure drops, allowing it to pass while the valve chokes water and gas (Moradi *et al.*, 2022b).

The pressure drops in the laminar and turbulent restrictors are calculated as follows,

$$\Delta P_{Laminar} = \frac{32\mu\rho vL}{D^2} , \quad (5)$$

$$\Delta P_{Turbulent} = \frac{k\rho v^2}{2} , \quad (6)$$

Where,

$\Delta P$  is the pressure drop over the restrictor

$\mu$  is the fluid viscosity

$\rho$  is the fluid density

$v$  is the fluid velocity

$L$  is the laminar restrictor length

$D$  is the laminar restrictor diameter

$k$  is the geometrical constant

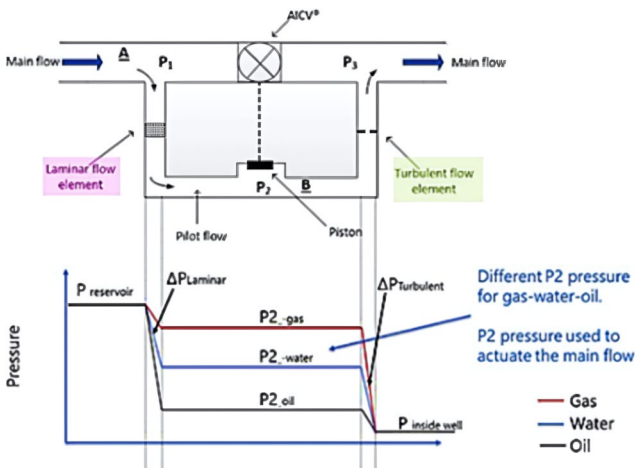


Fig. 5. A simplified sketch of the flow paths in AICV and pressure changes inside for different fluids (Anuththara *et al.* 2023).

Equation 5 explains how the pressure drops in the laminar restrictor depend on the fluid density and viscosity. Therefore, when low viscous fluid like water and gases pass through the restrictor, it gets a low-pressure drop compared to a high viscous fluid such as heavy oil. Because of this low-pressure drop, low viscous fluids have a higher pressure in the chamber between the restrictors, leading to higher velocity before

passing through the turbulent restrictor. As Equation 6 mentions, low viscous fluids experience a higher pressure drop than oil across the turbulent restrictor, which allows the AICV to remain open for oil while it is almost closed for water and gas (Moradi *et al.*, 2022b).

The performance criteria for AICV and AICD mentioned in Fig. 6 were used in this study where AICV closes its opening 60% when the water cut reaches 1, while AICD can close nearly 20% of its opening. The plots in Fig. 6 were obtained based on the experimental results of AICD and AICV for a fluid with the properties discussed in Chapter 4.1.

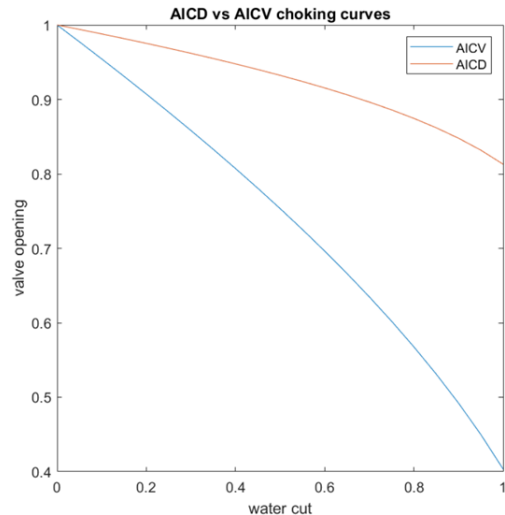


Fig. 6. Choking functionality comparison between AICV and AICD.

## 4. RESERVOIR MODEL IN ECLIPSE

### 4.1 Reservoir fluid and rock properties

The reservoir conditions and fluid properties used in this study were similar to the Troll field in the North sea which also used to obtain AICV/AICD test data. Table 1 presents the reservoir and fluid properties used for the simulations. (Anuththara *et al.*, 2023).

Table 1. Reservoir properties and rock properties

Parameter	Value
Oil density	890 kg/m <sup>3</sup>
Water density	1000 kg/m <sup>3</sup>
Gas density	0.67 kg/m <sup>3</sup>
GOR	50 Sm <sup>3</sup> / Sm <sup>3</sup>
Reservoir Pressure	130 bara
Water viscosity	0.45 cp
Oil viscosity	2.7 cp
Porosity	0.15 – 0.27
Temperature	68 °C

### 4.2 Reservoir grid

The dimensions of the reservoir are illustrated in Table 2. Figure 7 presents the 3D image of the reservoir is presented. The reservoir has two layers separated by a shale layer with a thickness of 50 m. From the 10<sup>th</sup> cell, the water is injected into the reservoir.

Table 2: Dimensions of the reservoir

Length of the reservoir (x)	1250 m
Width of the reservoir (y)	500 m
Height of the reservoir (z)	140 m

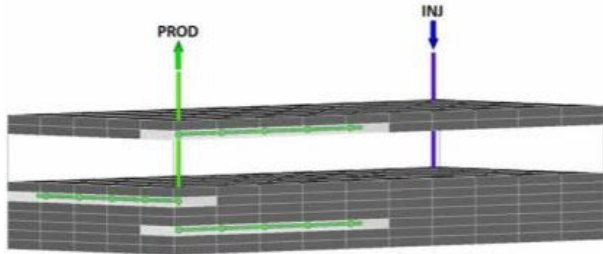


Fig. 7. Topology of the reservoir.

#### 4.3 Reservoir permeability

The reservoir is considered a homogeneous reservoir with the same porosity and permeability. The long-normal absolute permeability is assumed in the 100-500 mD range.

Figure 8 represents the relative permeability curves for water ( $k_{rw}$ ) and oil ( $k_{ro}$ ) with the water saturation ( $S_w$ ) used in this study, which was developed using Corey model.

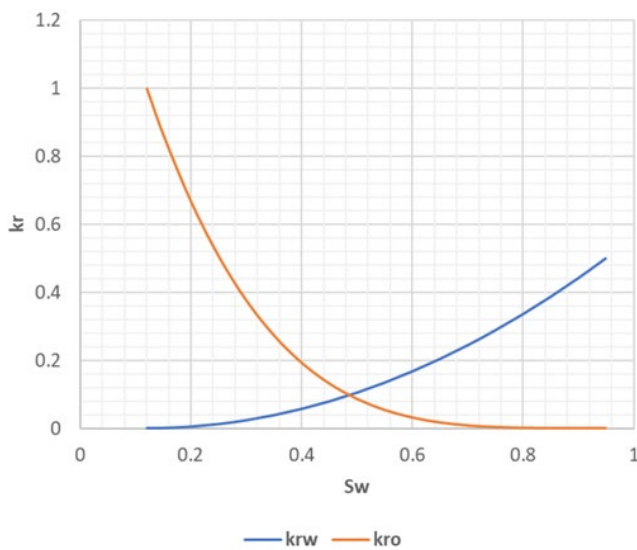


Fig. 8. Relative permeability curves.

#### 4.4 Initial conditions

The reservoir developed with ECLIPSE has three different zones. Therefore, the pressure, oil, water, and gas saturation change along the depth.

Initially, the reservoir pressure is 130 bar at 68°C, increasing as the depth rises. The depth of oil-water contact (OWC) and gas-oil contact (GOC) are 2300 m and 2010m, respectively. Therefore, the pressure increases as the depth increases per hydrostatic pressure, as shown in Fig. 9.

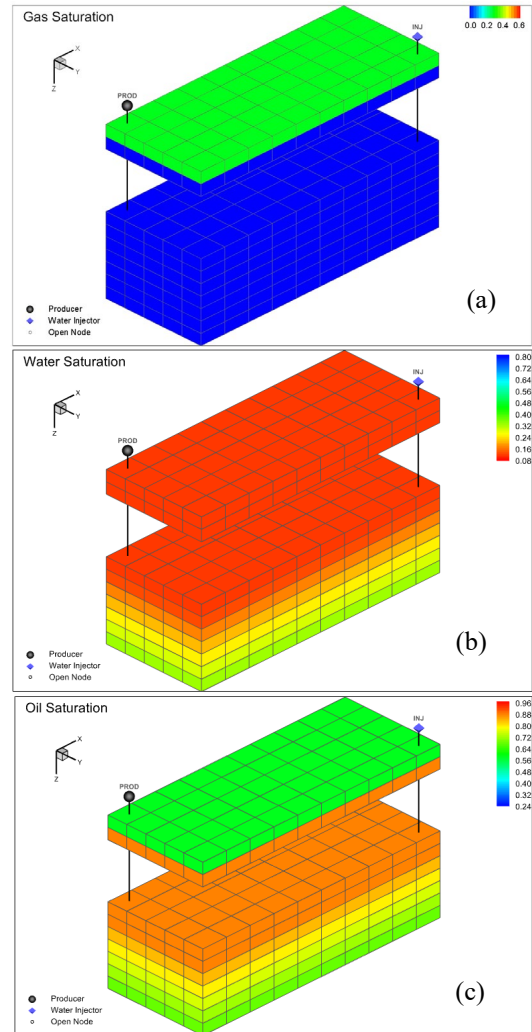


Fig. 9. Initial (a) water, (b) oil, and (c) gas saturation.

## 5. WELL MODEL IN OLGA

Figure 10 illustrates the pipe in the horizontal annulus. The annulus is the gap between the wellbore and the surface. Since OLGA does not have a method to simulate the flow through the annulus and inflow control devices, the OLGA model is developed to separate pipelines called wellbore and production tubing (Anuththara *et al.*, 2023).

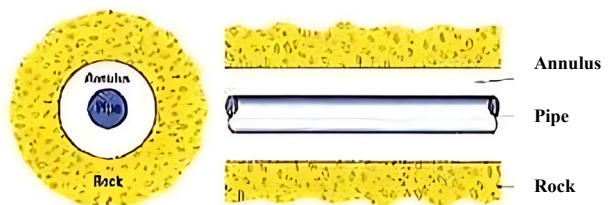


Fig. 10. Pipe in the horizontal annulus.

### 5.1 Compositional settings

The three black oil components (gas, oil, and water) are defined for the simulations. No gas has been injected into the reservoir; only oil and water feed rates are given. Compositions of the feeds are as in Table 3.



Table 3. Oil and water feed components.

Feed	Gas fraction	Water cut
Oil	50 Sm <sup>3</sup> / Sm <sup>3</sup> (GOR)	0.0001
Water	0.0001 Sm <sup>3</sup> /Sm <sup>3</sup> (GLR)	0.99

5.3 Flow component settings

Since the three laterals are modelled and studied for oil production, each lateral consists of a wellbore and production tubing. The length and diameter of each wellbore are 625 m and 0.2159 m, respectively. The top and bottom production tubing are the same length as the wellbore, while the middle length is 500 m. The junction node connects the three laterals into one production pipe.

The description of each lateral is illustrated in Table 4 below.

Table 4: Technical description of the wellbore and production tubing

SN	Pipe name	Diameter (m)	Roughness (µm)	Elevation (m)
1	Wellbore top	0.2159	15	
	Production tubing	0.1397	15	10
2	Wellbore top	0.2159	15	
	Production tubing	0.1397	15	80
3	Wellbore top	0.2159	15	
	Production tubing	0.1397	15	110
4	Outlet production tubing	0.1397	15	10

The oil is assumed to be produced from 5 zones with one inflow control device. Two packers isolate the wellbore to stop the fluid from flowing to different zones. Near well source uses the data given by ECLIPSE to connect OLGAs with ECLIPSE. Then, the fluid enters the wellbore after passing the inflow control device in the first section. The reservoir fluids enter the production tubing through a leak in the second section, as shown in Fig. 11.

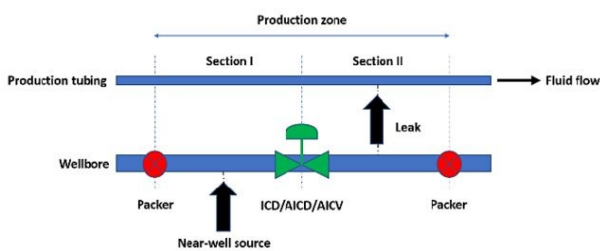


Fig. 11. Layout for one production zone.

6. RESULTS AND DISCUSSION

This chapter shows and discusses the results obtained from the OLGAs-ECLIPSE model. The functionality of ICD, AICD, and AICV was discussed, and the results obtained for 500 days were compared. AICV showed better results within the given circumstances in this case, as expected.

6.1 Oil production

Figure 12 illustrates the oil production rates for the FCDs studied in this work for 500 days. Initially, oil production for all the FCDs showed similar rates because the water

breakthrough had not yet occurred. However, the oil production rate decreased with the water breakthrough at around 160 - 180 days. As seen in Fig. 12, AICV shows a lower value for oil production, which is undesirable for AICV because AICV chokes the flow when it consists of low viscous fluids like water. However, the simulation was done for only 500 days, insufficiently covering the whole lifetime of the reservoir. However, AICV obtained a better oil fraction despite low oil production.

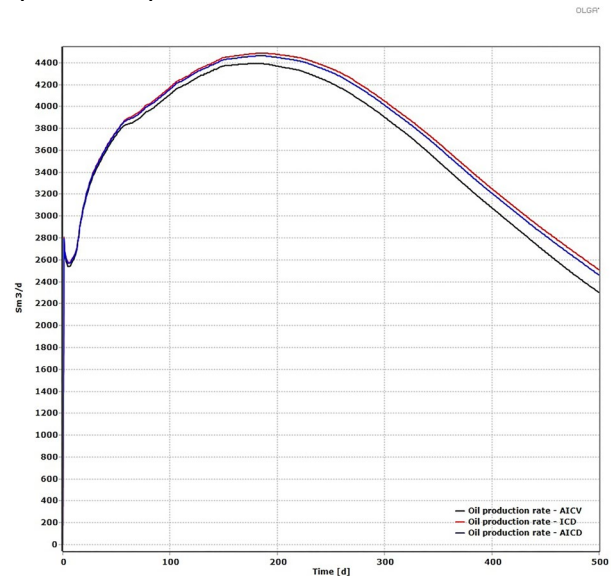


Fig. 12. Oil production rates for ICD, AICD, and AICV models.

Similarly, Figure 13 presents the accumulated oil production for each device over 500 days. AICV showed low oil production primarily due to its choking function when there is more water.

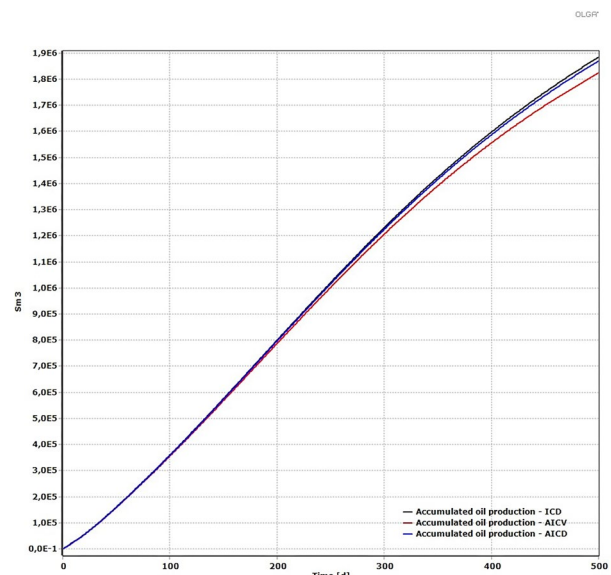


Fig. 13. Accumulated oil production over 500 days for ICD, AICD, and AICV.

6.2 Water production

The water production rates for 500 days for each control device are illustrated in Fig. 14. A small amount of water is produced even before the water breakthrough because of the water from the bottom lateral. However, the water production

increased significantly for each FCD with the water breakthrough. Nevertheless, after the water breaks through, the production rate increases significantly, and AICV showed a low water production rate increase compared to other FCDs.

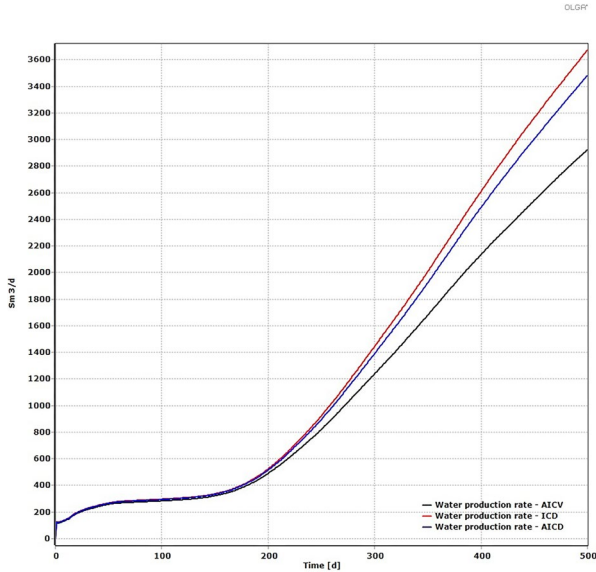


Fig. 14. Water production rates for 500 days for ICD, AICD, and AICV.

Figure 15 illustrates water accumulation over 500 days, and AICV showed lower water accumulation than AICD and ICD.

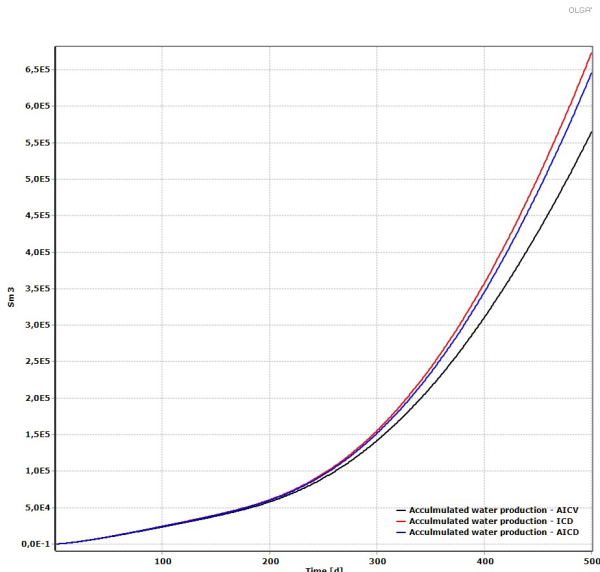


Fig. 15. Accumulated water production for ICD, AICD, and AICV for 500 days.

### 6.3 Total liquid production

With no liquid production cap defined, AICV showed the lowest liquid production, as shown in Fig. 16 and 17. Figure 16 represents the total liquid production rate, and Figure 17 shows the total accumulated liquid production. AICV shows low production because, after the water breakthrough, the water amount in the mixture is higher.

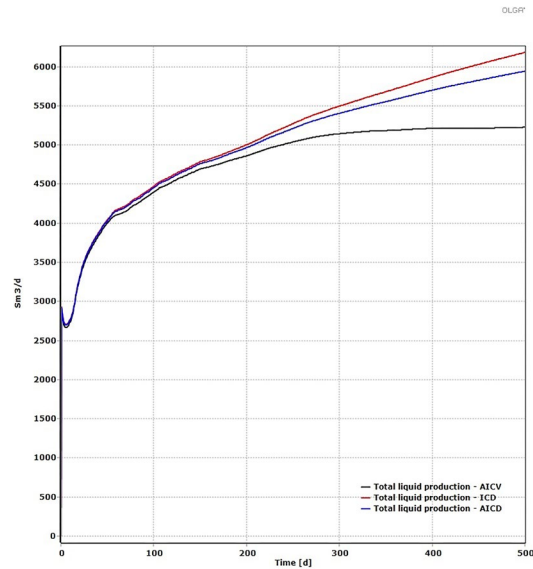


Fig. 16. Total liquid production rates for 500 days for ICD, AICD, and AICV.

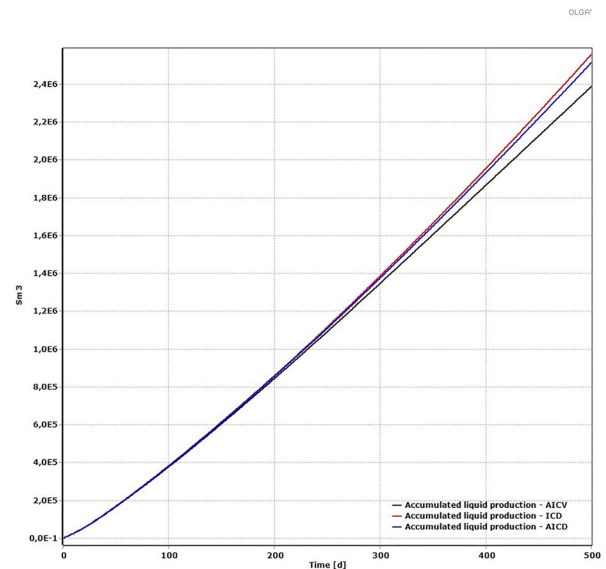


Fig. 17. Total accumulated liquid production for 500 days for ICD, AICV, and AICV.

### 6.4 Water cut variations

Keeping the water cut at its minimum is essential in the oil and gas industry as it impacts the overall economy and environmental sustainability. Figure 18 illustrates the water cut variation in the outlet for each FCD during 500 days of production. According to the Fig. 18, AICV, represented by solid colours, shows the lowest water cut along three laterals at almost every point compared with other FCDs. AICD represents continuous dotted lines showing the second lowest water cut along the tubes, while the well with ICD shows the highest water cut throughout all the production tubes because AICV has a higher choking ability with the water cut than AICD, while ICD does not have a choking ability.

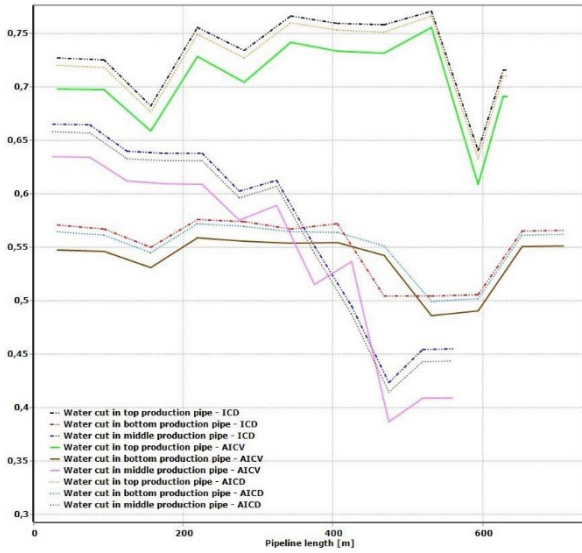


Fig. 18. Water cut (WC) variation along all the laterals for ICD, AICD, and AICV.

6.5 Fluid saturations

Figure 19 shows the oil saturation after 500 days in the reservoir simulated in ECLIPSE, while Fig. 20 shows the oil saturation along the horizontal plain of the three laterals. Compared to Fig. 9, the oil levels seem very low in the grid cells after 500 days of operation due to oil recovery.

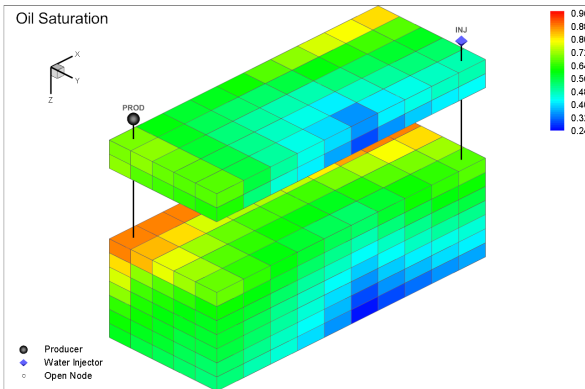
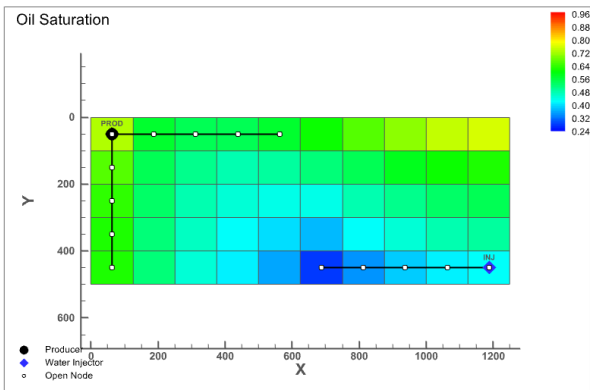
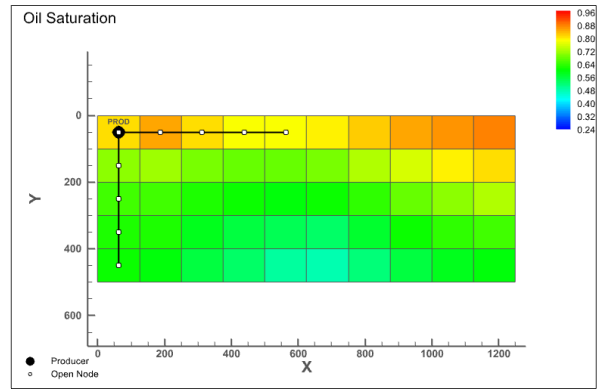


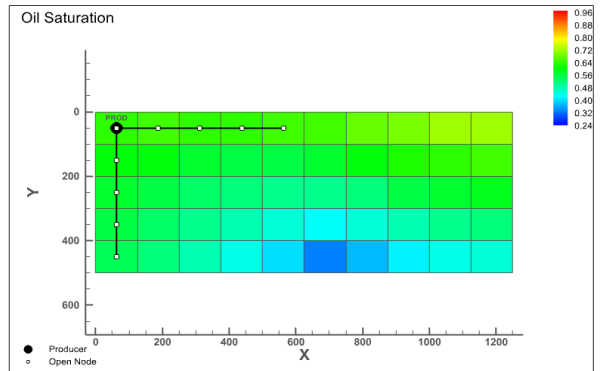
Fig. 19. Oil saturation in the reservoir after 500 days.



(a)



(b)



(c)

Fig. 20. Oil saturation in the horizontal layer around the (a) top, (b) middle, and (c) bottom laterals after 500 days.

Similarly, Figure 21 represents the water saturation after 500 days in the reservoir in three dimensions, while Fig. 22 shows the water saturation along the horizontal plain of the three laterals. Compared to Fig. 9, the water levels seem very high in the grid cells after 500 days of operation, mainly because of the water injection.

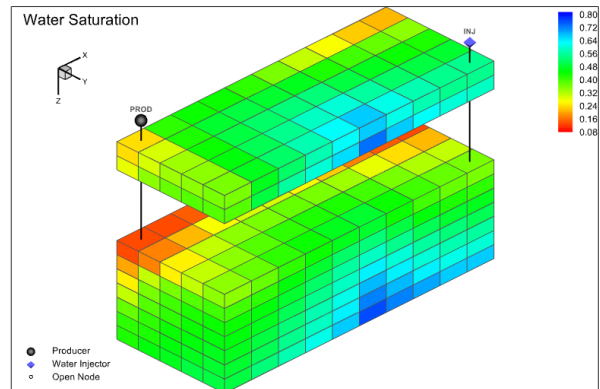
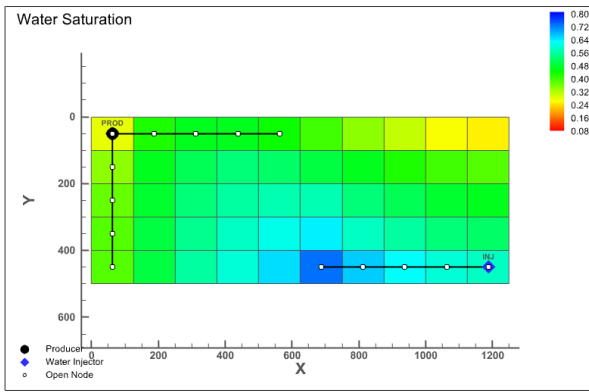
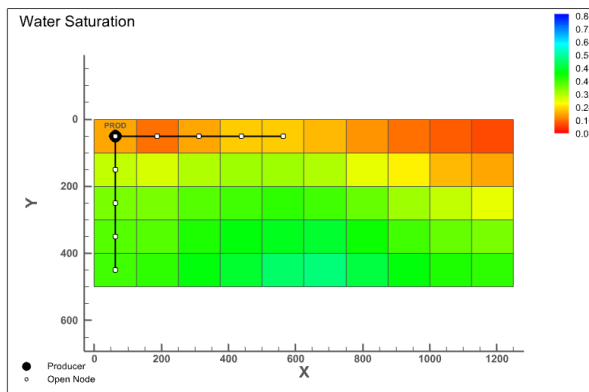


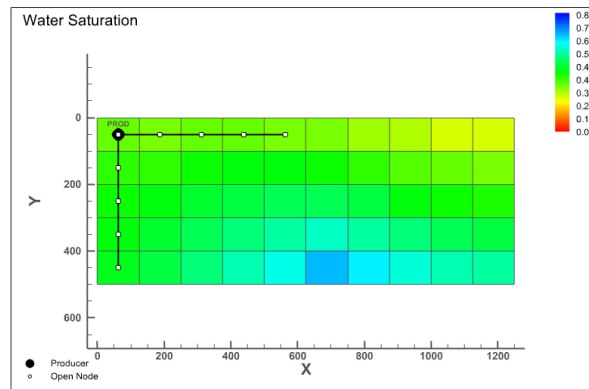
Fig. 21. Water saturation in the reservoir after 500 days.



(a)



(b)

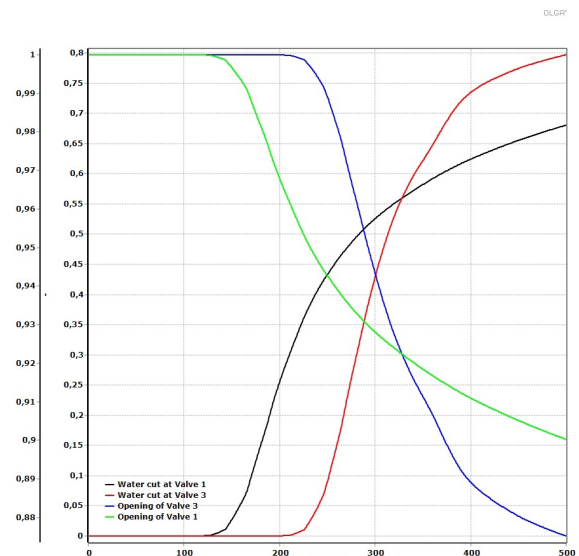


(c)

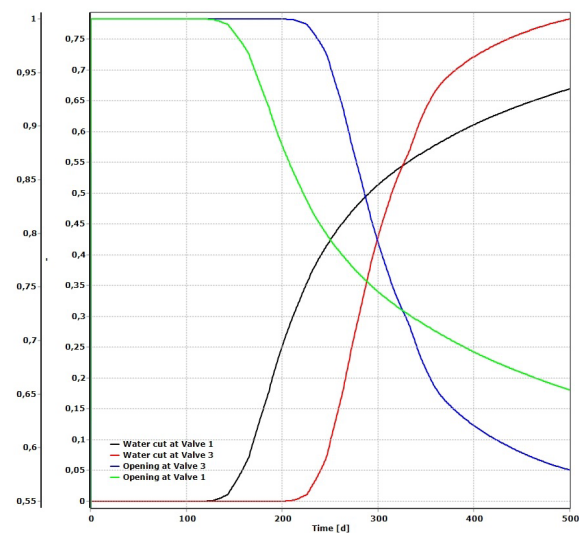
Fig. 22. Water saturation in the horizontal layer around the (a) top, (b) middle, and (c) bottom laterals after 500 days.

### 6.6 Chocking Effects of FCDs

Figure 23 shows how FCDs in similar locations at the top lateral (first and third FCD) for AICD and AICV perform with the water cut, and it shows that AICV has closed its opening up to 0.4 and AICD up to 0.12 for a water cut around 0.8 near third valve. Furthermore, near first valve, for a water cut around 0.67, AICD closed its opening up to 0.12 and AICV up to 0.35. As Fig. 5 mentions, AICV chokes its opening more than AICD does when more water is present. Therefore, the results obtained in Fig. 23 agree with the FCD functionalities.



(a)



(b)

Fig. 23. Choking effect on (a) AICD and (b) AICV with the relevant water cuts.

## 7. CONCLUSIONS

Water breakthroughs are a significant challenge in the oil and gas industry, and various inflow control devices (ICDs, AICDs, and AICVs) are designed to mitigate this issue. ICDs balance drawdown pressure and fluid flow but cannot block water once it enters the well. AICDs and AICVs can choke water entry, reducing water production and delaying water breakthrough. AICVs are more effective than AICDs in choking low-viscosity fluids.

The study simulated a simple reservoir and multilateral well for 500 days, observing the effects of ICDs, AICDs, and AICVs on oil and water production rates and water cuts. Results showed that AICV had the lowest water cut (0.539) compared to AICD (0.566) and ICD (0.573) and the lowest production rates with 2302 m<sup>3</sup>/day for oil and 2919 m<sup>3</sup>/day for water. Although AICVs performed unexpectedly due to the

short simulation period, they are expected to perform better over a more extended period with production limitations.

The study concluded that AICVs are the most effective in reducing water production, making them economically and environmentally viable. It recommended simulating reservoirs for at least 2000-3000 days for better results and proposed further work on gas injection and actual case simulations to benefit the industry. The primary objective of coupling OLGA and ECLIPSE was to model an advanced multilateral well and simulate the reservoir, and the well was successfully achieved.

#### ACKNOWLEDGMENTS

We sincerely thank the Research Council of Norway and Equinor for their financial support through Research Council Project No. 308817, Digital Wells for Optimal Production and Drainage (DigiWell). We also appreciate the University of South-Eastern Norway for providing the essential software arrangements for this work.

#### REFERENCES

- Elyasi, S. (2016). Assessment and evaluation of degree of multilateral well's performance for determination of their role in oil recovery at a fractured reservoir in Iran, *Egyptian Journal of Petroleum*, 25(1), pp. 1–14. doi: 10.1016/j.ejpe.2015.06.006.
- Von Flatern, R., (2021). Defining Multilateral Wells [WWW Document]. URL <https://www.slb.com/resource-library/oilfield-review/defining-series/defining-multilateral-wells> (accessed 04.03.24).
- Aakre, H., Halvorsen, B., Werswick, B., and Mathiesen, V. (2014). Autonomous Inflow Control Valve for Heavy and Extra-Heavy Oil, *SPE Heavy and Extra Heavy Oil Conference: Latin America*, OnePetro. doi: 10.2118/171141-MS.
- Anuththara, M., Moradi, A., Kumara, A.S., and Moldestad, B.M.E. (2023). Simulation of Oil Recovery Through Advanced Wells Using a Transient Fully Coupled Well-Reservoir Model, *Scandinavian Simulation Society*, pp. 86–93. doi: 10.3384/ecp200012.
- Moradi, A., Tavakolifaradonbe, J., and Moldestad, B.M.E. (2022a). Data-Driven Proxy Models for Improving Advanced Well Completion Design under Uncertainty, *Energies*, 15(20), p. 7484. doi: 10.3390/en15207484.
- Mathiesen, V., Werswick, B., and Aakre, H. (2014). The Next Generation Inflow Control, the Next Step to Increase Oil Recovery on the Norwegian Continental Shelf, *SPE Bergen One Day Seminar*, OnePetro. doi: 10.2118/169233-MS.
- Moradi, A., Samani, N.A., Kumara, A.S., and Moldestad, B.M.E. (2022b). Evaluating the performance of advanced wells in heavy oil reservoirs under uncertainty in permeability parameters, *Energy Reports*, 8, pp. 8605–8617. doi: 10.1016/j.egy.2022.06.077.



# Integration of Optimization Methods into Simulation Technology for Manufacturing via Warehouse Optimization

H. Hakalahti\*. A. Ala-Huikka\*\*. T. Luomanmäki\*\*\*. J. Hirvonen\*\*\*\*.

\**Digital Factory research group, Seinäjoki University of Applied Sciences, Finland (e-mail: [hannu.hakalahti@seamk.fi](mailto:hannu.hakalahti@seamk.fi)).*

\*\**Digital Factory research group, Seinäjoki University of Applied Sciences, Finland (e-mail: [alisa.ala-huikka@seamk.fi](mailto:alisa.ala-huikka@seamk.fi)).*

\*\*\**Digital Factory research group, Seinäjoki University of Applied Sciences, Finland (e-mail: [toni.luomanmaki@seamk.fi](mailto:toni.luomanmaki@seamk.fi)).*

\*\*\*\**Digital Factory research group, Seinäjoki University of Applied Sciences, Finland (e-mail: [juha.hirvonen@seamk.fi](mailto:juha.hirvonen@seamk.fi)).*

---

**Abstract:** The manufacturing industry is in a strong transition towards digital, intelligent, and sustainable manufacturing. However, small and medium-sized enterprises (SME) in the manufacturing industry often lack the resources and know-how to utilize digital tools as part of their research and development (R&D) activities. Thus, there is a need for concrete examples to show the benefits of these tools. This paper discusses a demonstration of warehouse optimization where a genetic algorithm is applied to optimize pallet transfers. The simulation model of a flexible manufacturing system (FMS) cell includes a warehouse with nine Euro pallets and a stacker crane. Visual Components simulation software was used for the simulation and an external Python application for the algorithm. As a result of the optimization, the total duration of the transfers was reduced by approximately 20 seconds (8.1 %). The demonstration has been used to showcase the integration of optimization methods into simulation technology and has ignited longer-term collaboration with the local industry on the same theme.

*Keywords:* 3D simulation, genetic algorithms, high bay rack, optimization, discrete event simulation

---

## 1. INTRODUCTION

Simulation and optimization are useful tools to understand, design, and develop production processes. For example, simulation can be used to model and analyze processes before they are implemented in the real production environment. Production processes can be optimized based on real-time or historical data. For example, several data analysis and computation models can be used to quickly locate equipment or quality problems and improve resource efficiency in production.

The combination of simulation and optimization enables a comprehensive analysis of production processes and efficient production planning. They can save costs, reduce production risks, and enhance product quality.

Bojic et al. (2023) did a case study using discrete-event simulation (DES) and genetic algorithm (GA) to assist operational production planning and optimization in the textile industry. A simulation model was created using Tecnomatix Plant Simulation software for a textile factory that produces over 300,000 shirts per year. The simulation model considered changes in customer demand, production times, available resources, and batch sizes. GA optimization improved production efficiency and reduced work in process (WIP) inventory levels.

Ernst et al. (2017) developed an optimization tool called Adv:ProcessOptimizer for multi-objective chemical process optimization. A specific GA was customized and developed for this tool. The tool integrates established methods with new concepts that work with simulation tools like Aspen Plus and ChemCad. The effectiveness of the tool was validated by optimizing an industrial styrene process. The results showed a well-distributed Pareto front, leading to savings in investment and operating costs compared to traditional methods. This confirmed the capability of Adv-tool to improve process efficiency and its support for decision making in process design.

Howard et al. (2023) developed a method to investigate energy flexibility in process cooling systems. A case study was performed on a Danish plant that uses process cooling for canned meat production. They used a combination of multi-agent, discrete-event, and system dynamics simulations to model the process. The results showed that significant savings in operational costs and reduction in CO<sub>2</sub> emissions can be achieved by optimizing the schedule of the refrigeration units based on forecasts of weather conditions, electricity prices and CO<sub>2</sub> emissions. This method provided insights on how to improve the energy performance of process cooling systems in food production without compromising the product quality and the production rate, through a weeklong simulation scenario.

Xie et al. (2015) used the genetic algorithm to schedule the single overhead crane so that its transport and shuffling operations are completed in the shortest possible time. The study found that the developed genetic algorithm provided good and quick solutions to the crane scheduling problem.

These case studies demonstrate how the integration of simulation and optimization tools works in various manufacturing industries. They can be useful for different kinds of applications, leading to more efficient, cost-effective, and high-quality production. However, SMEs in the manufacturing industry often lack resources and know-how to utilize these tools as part of their R&D activities or even information of their existence. Concrete pilots that demonstrate solving common problems in manufacturing industry are required to increase awareness of the possibilities of new simulation and optimization tools, and their integrations.

Warehouse operations provide an ideal demonstration environment for manufacturing SMEs due to their complexity, scalability, and the critical need for efficiency. It has been estimated that approximately 55 % of warehouse operating costs are caused by picking tasks (Bartholdi and Hackman, 2019, p. 25). Thus, increasing the picking and placing efficiency can generate remarkable savings.

Genetic algorithms are often used in optimization problems due to their robust search capabilities and flexibility as shown by two of the examples mentioned earlier in this section. The interest in genetic algorithms in the field of logistics has as well increased in the recent years among the researchers, and the number of publications has doubled between the years 2016 and 2020 (Grznár et al, 2021). In warehouse optimization problems reported in the literature, the goal often is to improve a forklift route in a warehouse described by a 2D layout consisting of shelves and corridors (e.g. Avdekins and Savrasovs, 2019; Kordos et al, 2020). Grznár et al (2021) worked with a 3D simulation of a conveyor system with workers sorting the goods coming to and leaving the warehouse. However, this work did not involve the actual warehouse structure.

This paper discusses optimizing an automated warehouse of a flexible manufacturing cell by utilizing a genetic algorithm. The high-bay rack and the stacker crane of the model add complexity due to vertical space utilization and dynamic movement, which are not typically addressed in simpler, 2D warehouse models. Also, the 3D simulation provides a more realistic and comprehensive testbed for evaluating the performance of GAs. Moreover, the FMS cell as a sample environment makes the demonstration interesting and accessible for the metal industry companies that are locally abundant. Finally, the demonstration involves cooperation of simulation software and an external optimization library, which gives a good example of the expansion potential of the applicable tools.

The paper is organized as follows: Section 2 describes the main methods applied in this study, discrete event-based

simulation and genetic algorithms, and the software used. Section 3 discusses the experiments performed and the results gained. The conclusion is drawn in the last section.

## 2. METHODS AND SOFTWARE

### 2.1 Discrete Event-Based Simulation

According to Banks et al (2019, p. 3) a simulation replicates the functioning of a real-world process or system as it evolves over time. Whether the simulation is conducted manually or utilizing a computer, it entails creating an artificial history of a system and observing that history to make inferences about the real system's operating characteristics. Shannon (1998, p. 1) defines simulation as a process of modeling a real system and taking experiments with the model in order to gain insights about the behavior of the real-world system and furthermore, to evaluate different operational strategies for the system.

Banks et al (2019, p. 9) categorize systems as discrete or continuous. Choi and Kang (2013, p. 8) adds quantum class to enhance the system's classification. They agree together that rarely any process is purely a certain one, but more of a combination of two or more, however, some class describes more the system behavior than the other ones. Banks et al (2019, p. 9) define the discrete system in a way where state variable(s) change only at a discrete set of points in time. Furthermore, Choi and Kang (2013, p. 9) defines the discrete-event simulation as a computer evaluation of a discrete-event dynamic system model. In the model, the operation of the simulated system is defined as a chronological sequence of events. As the pilot case was a discrete manufacturing process, the simulation method was chosen accordingly.

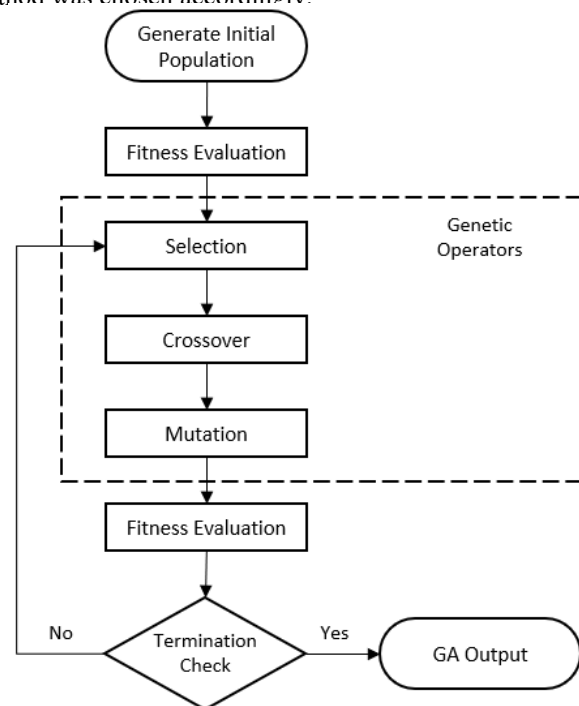


Fig. 1. Flowchart of the genetic algorithm.

## 2.2 Genetic Algorithms

### 2.2.1 Basics

A genetic algorithm developed by John Holland (Holland, 1975) is an optimization and search technique that is based on the principles of natural selection and genetics. In the genetic algorithm, a population of chromosomes evolve towards a better solution over each consecutive generation by using parent selection, crossover, and mutation. The genetic algorithms belong to the larger class of the evolutionary algorithms (EA) (Townsend, 2003).

The basic flowchart of the genetic algorithm is presented in Fig. 1.

### 2.2.2 Population

GA is an iterative process which starts with the creation of randomly chosen initial population of individuals (solution candidates for a given problem), which are represented by finite linear string of symbols, known as chromosome. A gene is an element in the chromosome, as shown in Fig. 2, and the allele is the value of the gene (Townsend, 2003).

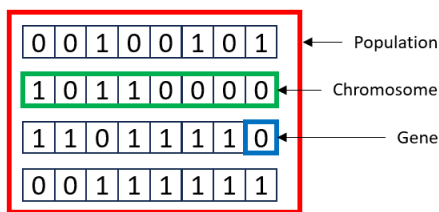


Fig. 2. Example of population, chromosome, and gene.

Determining the size of the population is important since choosing too small population might cause the GA to converge prematurely to a local minimum instead of global minimum due to lack of genetic variation in the population. On the other hand, too large population will require more computing time and thus makes the GA to run slower. The population size remains constant during the running of the genetic algorithm.

### 2.2.3 Genotype Representation

Genotype is a genetic composition of the chromosome (Haupt and Haupt, 2004).

A binary representation, where the chromosomes are represented as bits (a string of 1s and 0s), is the simplest and widely used representation. A floating-point representation is used for the continuous GA, and permutation representation is used for the cases where the order of genes matters (Fig. 3). Perhaps the most well-known use case of the permutation representation is the traveling salesman problem (TSP), where each city can be visited just once in some order.

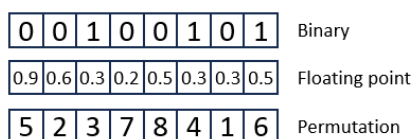


Fig. 3. Example of genotype representations in GA.

### 2.2.4 Fitness Function

At every evolutionary step, also known as generation, the current population is evaluated according to a fitness function set for a given problem. Because the fitness function calculates the fitness value for every chromosome on every generation, it greatly impacts the run time of the GA. Too computing heavy fitness function increases the run time of the GA (Townsend, 2003).

### 2.2.5 Selection

Parent selection operator selects the chromosomes in the population for reproduction. On every generation, the selected chromosomes are collected to a list known as mating pool. The better fitness value the chromosome has, the higher probability it has for being selected for the mating pool. Thus, the selection is based on the strategy of the survival-of-the-fittest (Townsend, 2003).

There are several different selection methods used in the genetic algorithms such as fitness-proportional selection, ranked selection, stochastic universal sampling, roulette wheel selection, truncation selection, and tournament selection (Townsend, 2003).

In tournament selection, a random number of individuals are selected from the population. Then, the best individual is selected from this group to be as a parent. This process is repeated until the mating pool is filled. Tournaments are often held between pairs of individuals (Goldberg & Deb, 1991).

### 2.2.6 Crossover

Crossover operator swaps the genetic material between two parent chromosomes to create new offspring for the next generation (Townsend, 2003).

The crossover between two good chromosomes does not necessarily create as fit or better offspring. However, because the parents are good, the probability of the offspring to be good is high. If the offspring happens to be a poor solution candidate, it will be removed from the population during the next generation.

In one-point or simple crossover, a random crossover point  $k$  is selected uniformly between 1 and the length of the parent chromosomes minus one  $[1, l - 1]$ . The genes after the crossover point  $k$  are then swapped between the parent chromosomes to create new offspring, as shown in Fig. 4 (Goldberg, 1989, p. 12).

In two-point crossover, two random crossover points are selected uniformly among the length of the parent chromosomes. The alternating segments of genes are then swapped between the parent chromosomes, as shown in Fig. 5. In general, the two-point crossover is better than one-point crossover to find solution more quickly (Townsend, 2003).

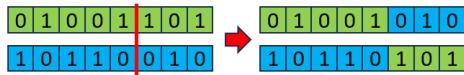


Fig. 4. Example of one-point crossover.

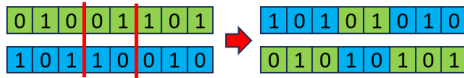


Fig. 5. Example of two-point crossover.

In uniform crossover, a randomly generated crossover mask is used to decide from which parent the offspring gets its genes. If there is a value 1 or 0 in the mask, the gene is copied from the first parent or second parent, respectively (Fig. 6). This procedure is repeated for the second offspring (Townsend, 2003).

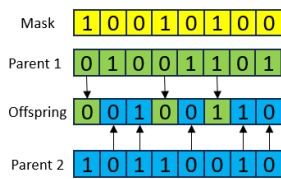


Fig. 6. Example of uniform crossover.

In permutation problems, standard crossover operators are not appropriate since each gene should be represented once and only once in the chromosome. One possible solution is to use a partially matched crossover (PMX). Under PMX, two random crossover points are chosen, and the genes are exchanged between these two points. The exchanged genes remain intact during the rest of the procedure. On the final step, the doubles (marked as yellow color in Fig. 7) are exchanged between the children to get correct permutations (Haupt and Haupt, 2004). Each child chromosome contains ordering information partially determined by each of its parents.

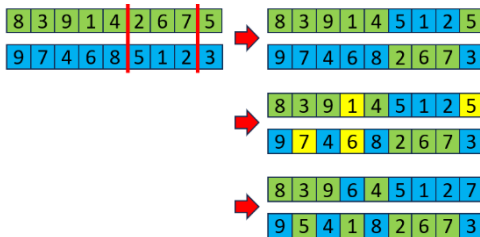


Fig. 7. Example of partially matched crossover (PMX).

Townsend (2003) gives a following summary of crossover methods: there is no more than 20 % difference in speed among the one-point, two-point, and uniform crossover. Uniform crossover or two-point crossover works better if the population is small or large, respectively, compared to the problem complexity.

### 2.2.7 Mutation

The mutation operator changes the value of one or more genes at randomly selected position in the chromosome. Mutation can take place at each position in the chromosome with some pre-defined probability, known as mutation rate, which is

usually small. The mutation operator makes the GA to find a near optimal solution to a given problem more easily by maintaining the genetic diversity in the population (Townsend, 2003).

The most common mutation operators are binary mutation, random resetting, swap mutation, scramble mutation, and inversion mutation.

In binary mutation, the value of the one or more genes is altered with a probability equal to the mutation rate (Larranaga, 1999). For example, the value of the third gene is changed from 1 to 0 in Fig. 8.



Fig. 8. Example of binary mutation.

In swap mutation (Fig. 9), the values of two randomly selected genes are interchanged. The swap mutation is commonly used in permutation-based representations (Larranaga, 1999).



Fig. 9. Example of swap mutation.

In scramble mutation (Fig. 10), a subset of the genes on the chromosome are selected and scrambled randomly (Larranaga, 1999).



Fig. 10. Example of scramble mutation.

In inversion mutation, two random points are chosen along the length of the chromosome. The genes between these points are then inverted as shown in Figure 11 (Goldberg, 1989, p. 166).



Fig. 11. Example of inversion mutation.

### 2.2.8 Elitism

Elitism means that the most fit chromosomes of the current generation are preserved for the next generation. Elitism prevents the population from losing its best solution due to crossover or mutation. The unwanted side effect is that there might be a super fit chromosome that causes the GA to converge prematurely (Townsend, 2003).

The elite size means the number of fit chromosomes preserved for the next generation.

### 2.2.9 Termination Condition

The termination condition of a genetic algorithm defines when to stop running the algorithm. Usually, the GA run is terminated when one of the following conditions is met: there is no improvement in the population for given number of consecutive generations, the maximum number of generations



is reached, or the objective function has reached a certain pre-defined value.

### 2.3 Visual Components

Visual Components is a 3D simulation and offline programming software that can be used for layout planning, feasibility analysis, virtual commissioning, and robot programming. The software has an extensive library of 3D models with 3000+ pre-defined and ready-to-use components including robots, conveyors, machines, resources, robot tools, factory facilities, and more. The user can also import self-made 3D models into the software (Visual Components, 2024).

Python programming language (version 2.7) can be utilized in components scripts. It is possible to add 3<sup>rd</sup> party Python package to the Visual Components, but this depends on the package. Alternatively, the user can communicate with an external application by using the TCP/IP sockets.

## 3. EXPERIMENTS AND RESULTS

### 3.1 Flexible Manufacturing Cell

There is a flexible manufacturing cell in the SeAMK's laboratory of the Machine and Production Technology. The cell consists of two Fanuc R2000iB/165F industrial robots, Kitamura HX500i machining center, workpiece positioner, and the storage system made by Fastems. The storage system has following components: high bay rack for storing Euro pallets and machining pallets, stacker crane for moving the pallets between storage shelves and workstations, material station for inserting and retrieving pallets in and out of the storage, loading station for moving the machining pallets between the stacker crane and robot, and two pallet banks next to the robots for holding the Euro pallets.

A simulation model (Figs. 12 and 13) of the FMS cell was made during one of the research and development projects. The robots, workpiece positioner, grill fences, storage selves, and Euro pallets are from the component library of the Visual Components. The stacker crane, pallet banks, loading station, material station, machining center, and machining pallet were designed in a 3D CAD software and imported into the Visual Components as STL files.

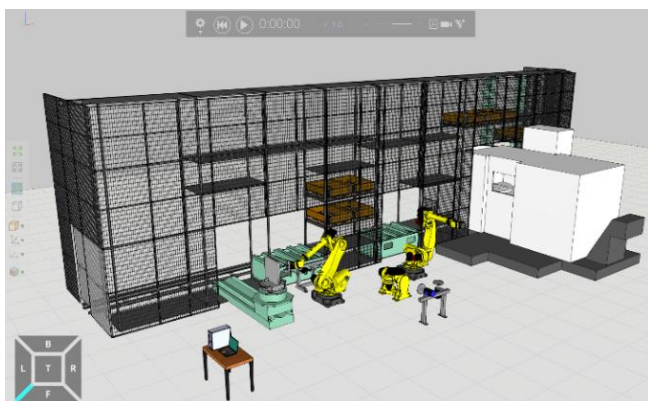


Fig. 12. Front view of the simulation model.

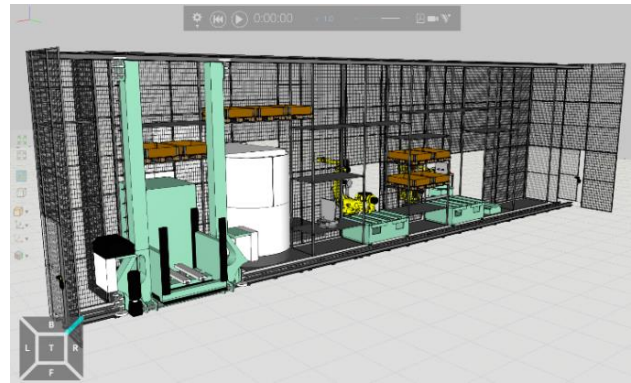


Fig. 13. Back view of the simulation model.

The simulation model of the FMS cell works as its real-world counterpart. The model does not, however, work together with the Fastems Manufacturing Management Software (MMS) that is used to control the real FMS cell. Instead, the simulation model relies on the component scripts to function properly.

### 3.2 Goal

The goal of the experiment was to utilize the genetic algorithm in the simulation model of the flexible manufacturing cell to minimize the time it takes to reshuffle the warehouse. In other words, the goal was to find the best permutation of Euro pallet transfers so that the stacker crane could move the pallets in the least amount of time. This kind of storage reshuffle process is common in the warehouses.

As the number of transfers increases, the number of different combinations of transfers also increases. For example, in the case of just nine transfers, the number of different combinations is  $9! = 362\,880$ .

Although the genetic algorithm can explore multiple solution candidates in parallel, it is not possible to compute the duration of every possible combination of transfers in a reasonable time frame. With the genetic algorithm, however, there is no need to go through every possible combination of transfers.

### 3.3 Pallet Transfers

The storage reshuffle process consists of nine transfers of Euro pallets. The transfers are randomly generated to eliminate the human bias for choosing transfers that one knows will benefit strongly from the optimization.

In the simulation model, the transfer of the Euro pallet is represented by a dictionary which contains the pallet ID, stock keeping unit (SKU) of the pallet, and source and destination shelf positions. For example, `{“pid”: 1, “sku”: “epallet”, “src”: 3, “dst”: 20}`, represents the transfer of the Euro pallet with id = 1 from a shelf position 3 to a shelf position 20. In the GA, these pallet transfers are represented as integers from 1 to 9, so when the fitness values are computed, one needs to decode the genes of the chromosomes into dictionaries.

In this experiment, the fitness function computes the total duration of the nine pallet transfers. Because the goal is to find the minimum duration, a shorter duration yields to smaller fitness value.



To speed up the computation of the fitness function, the transfer time from any given shelf position to any other shelf position was measured programmatically by computing the time difference between the moment when the pallet was picked up from the shelf and the moment when the pallet was placed onto the shelf. The transfer times were then hard coded into the component script of the stacker crane.

The process of measuring the transfer times was quite a tedious task. Alternatively, one could have made the simulation model so that the fitness value of the chromosome is computed by simulation. Based on some testing, the computation of fitness values by simulation took at least 30 minutes so it was decided to use hard coded transfer times instead.

The Euro pallet components are created to the source positions of the pallet transfers when the user starts the simulation.

### 3.4 Utilizing the GA in the Simulation Model

The Python code for the genetic algorithm was written into the PythonScript object of the stacker crane component in the simulation model. The genetic algorithm was implemented as a Python class whose input arguments are the direction of the optimization as an integer (0 for minimum and 1 for maximum), genes as a list of pallet IDs from 1 to 9, size of the population as integer, number of generations as integer, fitness function as a callable function, size of elite parents in the population as integer, and the mutation rate as a float.

The minimum direction was used since the goal was to find the order of the transfers in which the transfers are performed as quickly as possible, i.e. in minimum time. Based on some testing the population size was set to 1000 since with larger sizes the simulation ran significantly longer times without any major reduction in total transfer times of the pallets. The number of generations was set to 40 and the elite size to 10. According to Townsend (2003, p. 43) the probability of mutation is set to be inversely proportional to the size of the chromosome. Since each chromosome has 9 genes, the mutation rate was set to 11 % ( $1/9 = 0.111$ ). The parents of the next generation were selected using a tournament selection of 3 chromosomes. Because the goal was to find the best permutation of pallet transfers, PMX was used for the crossover operator. Finally, the swap mutation was used for the mutation operator. Below is a list of the values of the operators and input arguments of the GA.

- Direction: minimum
- Genes: [1, 2, 3, 4, 5, 6, 7, 8, 9]
- Population size: 1000
- Number of generations: 40
- Elite size: 10
- Mutation rate: 11 %
- Selection: tournament of 3 chromosomes
- Crossover: partially matched crossover (PMX)
- Mutation: swap.

### 3.5 Results

The simulation was run 50 times with and without the genetic algorithm to see how much the genetic algorithm reduces the total transfer time.

When the order of the transfers was optimized with the genetic algorithm, the total duration of the transfers was reduced by circa 20.5 seconds on average. This represents about an 8.1 % reduction in transfer time. The time reduction achieved with the GA in each run is shown in Fig. 14.

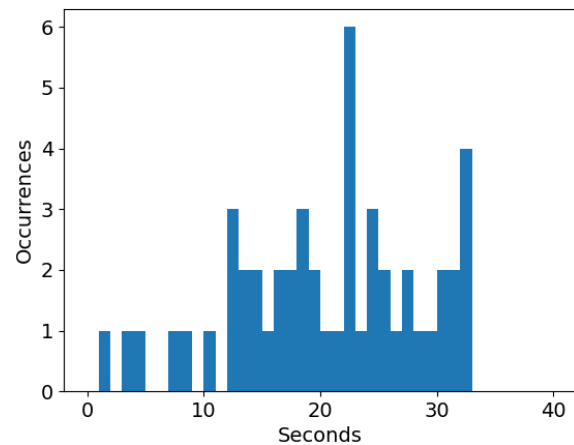


Fig. 14. Histogram of the time reduction achieved with the GA.

It was noticed that the fitness value did not improve that much between the first and last generation of the genetic algorithm. Figure 15 shows a typical progress of the fitness value during the generations of the GA.

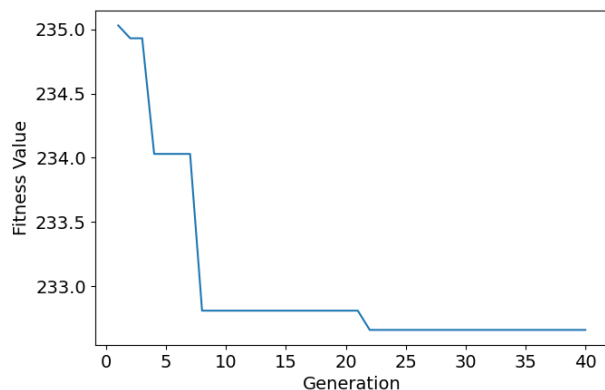


Fig. 15. The progress of the fitness value.

In many cases, the fitness value, i.e. the duration of the transfers, improved mere few seconds. The different combinations of input parameters of the genetic algorithm were tested but the results remained the same.

One possible explanation is the small population size which covers about 0.002 % of all possible solutions (362 880). Increasing the population size would, however, make the simulation to run significantly longer. Thus, the population size of 1000 was chosen.

Other optimization methods than genetic algorithm were not tested in this experiment so it is hard to say if some of them could perform better than GA.

#### 4. CONCLUSIONS

Manufacturing systems are complex environments with many interacting parts and variables in dynamic processes. These complexities often lead to challenges that are difficult to solve with traditional methods, especially as the demand for efficiency grows. To address these issues, combining optimization techniques with 3D simulation has proven to be highly effective. This approach allows us to explore a wide range of potential solutions, helping to find the best ways to improve system performance. By integrating these methods, we can make manufacturing processes more efficient and better equipped to handle the complexities of modern production.

This paper showed how to use an optimization method together with a simulation tool to find a solution that saves time and resources in production. The total time saving of 20 seconds (8.1 %) that was achieved in the demonstration is a significant improvement to the process. The automated warehouse that was utilized as the testbed is a practical and familiar example of a suitable scale for SMEs and thus makes the demonstration more illustrative and easier to catch. The 3D environment is a descriptive surrounding that facilitates showing the process in practice and makes the simulation more realistic. In addition, the high-bay rack and the stacker crane increase complexity because they involve using vertical space and dynamic movement, aspects usually not considered in basic, two-dimensional warehouse models.

The demonstration has been presented in one regional technology event and several smaller workshops for selected SMEs from the manufacturing industry. The reception has been good, and a closer collaboration on the same theme has started with one of the companies. Thus, the methods presented in this paper will be applied in other warehouses in real manufacturing surroundings shortly. Furthermore, the longer-term aim is to promote the integration possibilities of optimization methods and simulation technology from the product level to the development of production-level solutions. This supports the green transition as unnecessary work can be eliminated and processes can be streamlined.

#### ACKNOWLEDGEMENTS

This paper was written as a part of the project OPLITE (A80151), and the funding from the Regional Council of South Ostrobothnia is greatly appreciated.

#### REFERENCES

- Avdekins, A. and Savrasovs, M. (2019). Making warehouse logistics smart by effective placement strategy based on genetic algorithms. *Transport and Telecommunication Journal* 20(4):318-324. doi:10.2478/tjt-2019-0026.
- Banks, J. (2005). *Discrete-event system simulation*. 4th ed. Upper Saddle River, N.J: Pearson Prentice Hall.
- Bartholdi, J.J. and Hackman, S.T. (2019). *Warehouse and Distribution Science*, Georgia Institute of Technology. <https://www.warehouse-science.com/book/index.html> (accessed on 22 May 2024).
- Bojic, S., Maslaric, M., Mircetic, D., Nikolicic, S., and Todorovic, V. (2023). Simulation and Genetic Algorithm-based approach for multi-objective optimization of production planning: A case study in industry. *Advances in Production Engineering & Management*, 18(2), 250-262. doi: 10.14743/apem2023.2.471.
- Choi, B. K. and Kang, D. (2013). *Modeling and simulation of discrete-event systems*. Hoboken, N.J.: John Wiley & Sons Inc.
- Ernst, P., Zimmermann, K., and Fieg, G. (2017). Multi-Objective Optimization-Tool for the Universal Application in Chemical Process Design. *Chemical Engineering & Technology*, 40(10), 1867–1875. doi: 10.1002/ceat.201600734.
- Goldberg, D. and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms, in *Foundations of Genetic Algorithms*, pp. 69-93, Morgan Kaufmann, San Francisco, Calif.
- Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison Wesley.
- Grznár, P., Krajčovič, M., Gola, A., Dulina, E., Furmannová, B., Mozol, Š., Plinta, D., Burganová, N., Danilczuk, W., and Svitek, R. (2021). The Use of a Genetic Algorithm for Sorting Warehouse Optimisation. *Processes*, 9(7), 1197. doi: 10.3390/pr9071197
- Haupt, R. L. and Haupt, S. E. *Practical Genetic Algorithms*, Second Edition. John Wiley & Sons, Inc. (2004).
- Holland, J. *Adaptation in Natural and Artificial Systems*. MI: University of Michigan Press. (1975).
- Howard, D. A., Jørgensen, B. N., and Ma, Z. (2023). Multi-Method Simulation and Multi-Objective Optimization for Energy-Flexibility-Potential Assessment of Food-Production Process Cooling. *Energies*, 16(3), 1514. doi: 10.3390/en16031514.
- Kordos, M., Boryczko, J., Blachnik, M., and Golak, S. (2020). Optimization of Warehouse Operations with Genetic Algorithms. *Applied Sciences*, 10(14), 4817. doi: 10.3390/app10144817

- Larranaga, P., Kuijpers, C. M. H., Murga, R. H., Inza, I., and Dizdarevic, S. (1999). Genetic algorithms for the travelling salesman problem: A review of representations and operators. *Artificial intelligence review*, 13, 129-170.
- Shannon, R. E. Introduction to the art and science of simulation. 1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274), Washington, DC, USA, (1998), pp. 7-14 vol.1, doi: 10.1109/WSC.1998.744892.
- Townsend, A. A. R. (2003). Genetic Algorithm–A Tutorial. Av.: [www-course.cs.york.ac.uk/evo/SupportingDocs/TutorialGAs.pdf](http://www-course.cs.york.ac.uk/evo/SupportingDocs/TutorialGAs.pdf), 8.
- Visual Components. (2024). Visual Components - 3D manufacturing simulation software. <https://www.visualcomponents.com/>
- Xie, X., Zheng, Y., Li, Y. (2015). Genetic Algorithm and Its Performance Analysis for Scheduling a Single Crane. doi: 10.1155/2015/618436

# Evaluating Modelling Performance: Sensitivity Analysis of Data Volume in Industrial Batch Processes

Simon Mählkvist <sup>\*,\*\*</sup> Thomas Helander <sup>\*</sup>  
Konstantinos Kyprianidis <sup>\*\*</sup>

<sup>\*</sup> *Kanthal AB, Hallstahammar, Sweden*  
(*e-mail: simonmkvst@gmail.com*)

<sup>\*\*</sup> *Future Energy Center, Mälardalen University, Västerås, Sweden*

---

**Abstract:** This study conducts a sensitivity analysis to evaluate the influence of varying data volumes on model performance within multi-product batch processes in the iron and steel industry. Nine machine learning models, encompassing both ensemble and parametric methods, were rigorously tested using a data withholding approach. The results demonstrate that ensemble models, particularly Random Forest and Gradient Boosting, consistently outperformed parametric models across different data volumes, showcasing superior generalisation and robustness to outliers. These findings underscore the importance of careful model selection and comprehensive data preprocessing in enhancing model performance and suggest that ensemble methods are particularly well-suited for complex industrial applications where data quality and volume are critical.

*Keywords:* Machine Learning, Model selection, Performance evaluation, Data volume sensitivity, Iron and steel industry, Industrial batch processes,

---

## 1. INTRODUCTION

The iron and steel industry, a cornerstone of global industrial development, is responsible for approximately 7.2% of global Green House Gas (GHG) emissions, highlighting its significant contribution to climate change (Ritchie et al., 2020). With global steel production anticipated to rise by approximately 30% by 2050, the demand for innovation and sustainable practices in this sector has become increasingly urgent (Yoro and Daramola, 2020). Thus, the advancement of more efficient production methods is not only an environmental imperative but also essential for the industry's long-term viability. Moreover, the development of accurate and reliable models can significantly contribute to reducing waste, facilitating process control, and improving overall product quality. By optimising the predictive capabilities of these models, industries can enhance their operational efficiency and sustainability, thereby achieving better outcomes both economically and environmentally.

In recent years, significant interest has been directed towards the application of Machine Learning (ML) techniques in industrial processes, driven by advancements in data acquisition technologies and the increasing complexity and volume of industrial data. These advancements have enabled the development of sophisticated models capable of processing vast amounts of data, thereby improving decision-making and operational efficiency within industrial contexts.

This study aims to systematically evaluate the impact of data volume and complexity on the performance of ML models in multi-product batch processes within the iron and steel industry. A sensitivity analysis is conducted to provide insights that will guide future model development and applications in industrial batch processes.

A rigorous and systematic approach has been adopted in this study, wherein the effect of varying data volumes on model accuracy and complexity is analysed to ensure a comprehensive examination of these critical factors. A diverse range of ML models, with varying degrees of complexity, has been selected to assess their performance across different scenarios. These models include traditional machine learning algorithms, which are recognised for their efficacy in handling tabular data. Neural networks were excluded from this analysis due to the tabular nature of the dataset, which does not inherently suit such models, as evidenced by Shwartz-Ziv and Armon's findings that ensemble models generally outperform deep neural networks on tabular data (Shwartz-Ziv and Armon, 2022).

To explore the relationship between data volume and model performance, a data withholding approach has been implemented, enabling an assessment of how error scores fluctuate with varying data volumes. The data volume in this study ranges from 10 to 10,000 samples, distributed across 10 logarithmically spaced steps. These steps include 10, 22, 46, 100, 215, 464, 1,000, 2,154, 4,642, and 10,000 samples. This logarithmic progression ensures that the analysis covers a broad range of data volumes, providing a nuanced understanding of how data availability impacts model performance.

---

<sup>\*</sup> The authors gratefully acknowledge Kanthal AB, Automation Region Research Academy (ARRAY), and the Swedish Knowledge Foundation (KKS) for their support.

Previous studies have also examined the impact of data volume on model performance, albeit with differing scopes and methodologies. For instance, Bailly et al. (2022) investigated the effect of data volume on model metrics using artificially generated datasets with volumes of 1,000, 10,000, and 100,000 samples. Their findings indicated that, within the specific setup employed, data volume did not significantly influence model metrics, suggesting that the relationship between data volume and model performance may be context-dependent and influenced by factors such as data characteristics and model selection. In another study, Ramezan et al. (2021) explored the effects of training sample size on the performance of six supervised ML algorithms in classifying a large-area high-spatial-resolution remotely sensed dataset. Their work demonstrated that, while larger training sets generally led to better performance, there was considerable variation in how different classifiers responded to changes in sample size. These variations underscore the complexity of the relationship between data volume and model performance, highlighting the need for context-specific analyses.

The dataset employed in this study originates from the production of thermocouple materials at Kanthal Hallstahammar, specifically from the key stages of melting and hot rolling. Thermocouples are vital components in temperature measurement and are among the most common methods used in industrial processes, including those in the iron and steel industry. Their widespread use underscores the practical significance of accurate and reliable temperature measurement in maintaining process efficiency and product quality. The dataset comprises measurements of chemical composition and Electromotive Force (EMF), with the objective of predicting the final properties after hot rolling based on initial measurements taken after melting. A more detailed description of the pre-processed dataset is provided in Section 3.1.

The overarching aim of this work is to enhance the modelling of industrial processes by adhering to the principle of Occam's razor, which advocates for simplicity in model design. While continuous advancements in research contribute to increasingly complex models, it is crucial to balance complexity with practical implementation. This study builds upon previous work, such as Rendall et al. (2019) and Mählkvist et al. (2023), which examines the trade-offs between model complexity and performance. Rendall et al. (2019) succinctly illustrated the relationship between modelling complexity and implementation challenges, providing a framework for assessing the practicality of complex models in real-world applications. Similarly, Mählkvist et al. (2023) evaluated the modelling complexity of different classification models, including Logistic Regression, Random Forest Classifier, and Support Vector Classifier, to determine the most suitable model for the specific data problem at hand. These studies underscore the importance of balancing model sophistication with practical considerations, such as ease of implementation and computational efficiency.

Through a comparative analysis, it is intended to evaluate whether models with specific characteristics offer superior insights into the industrial processes under study. It is hypothesised that some models will perform better with larger data volumes and that a diverse range of model char-

acteristics will yield more comprehensive insights, particularly under conditions of data saturation. Ultimately, this research seeks to determine the optimal balance between data volume and model performance within the context of industrial batch process modelling.

## 2. METHODOLOGY

This section delineates the methodology employed in this study, encompassing details about the development environment, systematic data processing approaches, model training, and evaluation techniques. The approach has been designed to ensure robust and reproducible results through meticulous dataset handling, model selection, and hyperparameter tuning.

### 2.1 Coding and Dependencies

Python is the coding language used for this study. Besides arbitrary dependencies on Pandas, NumPy, and other common libraries, the package `scikit-learn` Pedregosa et al. (2011) is employed for the implementation of the ML models, as well as for hyperparameter tuning.

### 2.2 Systematic Approach for Datasets and Modelling

#### Dataset and Subset sampling

*Dataset and Subset Sampling* This study begins with the product datasets, denoted as  $\mathcal{P}_x$ , where  $x \in \mathcal{L}_P$  and  $\mathcal{L}_P$  represents a list of all product datasets, each identified by a Greek letter, such as  $\mathcal{L}_P = [\alpha, \beta, \dots]$ .

A data withholding approach is employed to generate increasingly larger datasets by sampling from the original product datasets. These smaller datasets are referred to as subsets. Each subset derived from a product dataset  $\mathcal{P}_x$  is denoted by  $\mathcal{S}_{i,j}^x$ , where  $\mathcal{S}_{i,j}^x$  represents the  $i$ -th iteration of sampling from the  $j$ -th subset of the  $x$ -th product dataset  $\mathcal{P}_x$ .

In this notation,  $i$  varies from 1 to  $n$ , where  $n$  is the total number of iterations performed for each subset size. This allows the same volume to be sampled multiple times to capture a more representative dataset. For instance, if the sampling volume is 10 samples,  $n$  subsets of volume 10 are generated by randomly selecting samples.

The index  $j$  corresponds to the position within a list of predefined sampling sizes, denoted as  $\mathcal{L}_V = [v_1, v_2, \dots, v_j]$ . Each element  $v_j$  in  $\mathcal{L}_V$  defines the size of the subset  $\mathcal{S}_{i,j}^x$ , ensuring that  $\mathcal{S}_{i,j}^x \subseteq \mathcal{P}_x$ .

Each iteration  $i$  of a subset  $\mathcal{S}_{i,j}^x$  not only represents a random sampling from  $\mathcal{P}_x$  but also retains all elements from the previous smaller subset  $\mathcal{S}_{i,j-1}^x$ . Consequently, as  $j$  increases (i.e., as the subset size grows within the same iteration  $x$ ), each new subset includes all samples from the preceding smaller subset for the same product dataset. This approach ensures that as the dataset increases in volume, it maintains the same reference samples, thereby preserving consistency across different subset sizes.

*Modelling* A list of machine learning models of arbitrary size is utilised for training. Each element in the list ( $\mathcal{L}_M = [m_1, m_2, \dots, m_k]$ ) denotes a model  $m_k$  indexed by  $k$ .



Table 1. Characteristics of Machine Learning Models Evaluated in the Study

Models (Abbreviations)	Parametric/ Non-Parametric	Regularization (None/L1/L2)	Linearity (Linear/Non-linear)	Sensitivity to Scaling/Outliers
Ordinary Least Squares Linear Regression (OLS)	Parametric	None	Linear	Sensitive
Ridge Regression (Ridge)	Parametric	L2	Linear	Sensitive
Least absolute shrinkage and selection operator (Lasso)	Parametric	L1	Linear	Sensitive
Decision Tree Regression (DTR)	Non-parametric	None	Non-linear	Robust
Random Forest Regression (RFR)	Non-parametric	None	Non-linear	Robust
Gradient Boosting Regression (GBR)	Non-parametric	None	Non-linear	Robust
Linear Support Vector Regression (LIN)	Parametric	L2	Linear	Sensitive
Polynomial Support Vector Regression (POLY)	Parametric	L2	Non-linear	Sensitive
Radial Basis Function Support Vector Regression (RBF)	Parametric	L2	Non-linear	Sensitive

Each model in the list  $\mathcal{L}_M$  is trained individually on each subset  $\mathcal{S}_{i,j}^x$  derived from the product datasets.

The naming convention for a model trained on a specific subset follows the format  $\mathcal{M}_{i,j,k}^x$ . This indicates that the model indexed  $k$  from the list  $\mathcal{L}_M$  has been trained on subset  $\mathcal{S}_{i,j}^x$ , where  $x$  refers to the originating product dataset,  $i$  to the iteration, and  $j$  to the specific subset volume as defined by its position in the list of sampling sizes  $\mathcal{L}_V$ .

### 2.3 Model Description and Parameter Range

This section outlines the models to be implemented, detailing each model in the subsequent subsections. Additionally, it includes the parameters and their respective ranges used for hyperparameter estimation, where applicable.

The model list ( $\mathcal{L}_M$ ) consists of 9 ML models, as shown in the first column of Table 1. Thus, the length of the list of models is  $|\mathcal{L}_M| = 9$ .

A log-uniform distribution is used to define the hyperparameter range for many of the parameters. This distribution is particularly useful for parameters that span several orders of magnitude, as it facilitates the exploration of a wide range of scales effectively. The log-uniform distribution is defined as:

$$\mathcal{U}(x, y) \quad (1)$$

where  $\mathcal{U}$  is the log-uniform distribution, and  $x$  and  $y$  are the lower and upper bounds, respectively.

In addition, a random integer distribution is employed to define the range for integer-valued hyperparameters, such as the number of estimators in ensemble models or the depth of decision trees. This distribution is particularly useful when the hyperparameter must take discrete values within a specified range. The random integer distribution is defined as:

$$\mathcal{I}(a, b) \quad (2)$$

where  $\mathcal{I}$  is the random integer distribution, and  $a$  and  $b$  are the lower and upper bounds, respectively. This distribution uniformly samples integer values between  $a$  and  $b$ , inclusive.

*Ordinary Least Squares Linear Regression (OLS)* The OLS is a widely used approach to linear modelling that fits coefficients for all dimensions in the datasets to minimise

the residual sum of squares between the observed values and the values predicted by the model (James et al., 2013).

*Ridge Regression (Ridge)* The Ridge model, also known as Tikhonov regularisation, extends linear methods such as OLS by incorporating regularisation. This model addresses a regression problem using the l2-norm.

The method was introduced by Hoerl and Kennard (1970a) in their 1970a; 1970b works.

The primary parameter for the Ridge model is the regularisation parameter for the l2-norm. Details of the parameter and the range of values used are provided in Table 2.

*Least absolute shrinkage and selection operator (Lasso)* The Lasso, similar to Ridge, is a linear model trained with regularisation, but it uses the l1-norm instead. The term was introduced by Tibshirani (1996).

The parameters for the Lasso are similar to those of Ridge, focusing on the regularisation parameter. However, in the case of Lasso, the parameter regulates the l1-norm. Details of the parameter and the range of values used are provided in Table 2.

Table 2. Parameters for Ridge and Lasso Regression Models

Model	Parameter	Scope
<b>Ridge</b>	Alpha <sup>a</sup>	$\mathcal{U}(0.01, 100)$
<b>Lasso</b>	Alpha <sup>a</sup>	$\mathcal{U}(0.01, 100)$

<sup>a</sup> The alpha parameter regulates the regularisation strength of the model.

*Decision Tree Regression (DTR)* The DTR is the first of the non-parametric methods and it infers simple decision rules from the data James et al. (2013).

The key parameter for the DTR is the maximum number of features, which determines the number of features to consider when finding the best split. Selecting the appropriate maximum number of features is crucial for controlling the diversity of features considered at each split. Refer to Table 3 for details.

*Random Forest Regression (RFR)* RFR, also known as random decision forests, is an ensemble learning method used for regression tasks. This technique constructs a multitude of decision trees during the training phase. Each tree in the forest relies on the values of a random vector, which is sampled independently and follows the same distribution across all trees (Breiman, 2001).

The parameters for the RFR include the maximum number of features (as with DTR) and the number of estimators, which represents the number of trees in the forest. The choice of the number of trees is critical, as it impacts both the model's performance and the risk of over-fitting.

Refer to Table 3 for details.

**Gradient Boosting Regression (GBR)** GBR is an ensemble learning method used for regression tasks. This technique builds a series of decision trees sequentially, with each tree aiming to correct the errors made by its predecessor. The process involves fitting new models to the residual errors of the previous models, thereby improving accuracy with each iteration. The final model is a weighted sum of all individual models, resulting in a robust predictive model that minimises the overall prediction error (Hastie et al., 2009).

The primary distinction between GBR and RFR lies in their construction strategy: random forests build trees independently and combine their results, whereas gradient boosting builds trees iteratively, with each tree focused on correcting the errors of the previous ones.

The parameters that GBR shares with the previous tree-based models include the maximum number of features, the number of estimators, and the maximum depth, which defines how deep each tree can grow.

Refer to Table 3 for details.

Table 3. Parameters for Decision Tree, Random Forest, and Gradient Boosting Models

Model	Parameter	Scope
<b>Decision Tree</b>	Max Features <sup>a</sup>	$\mathcal{I}(1, 100)$
<b>Random Forest</b>	Max Features <sup>a</sup> # Estimators <sup>b</sup>	$\mathcal{I}(1, 100)$ $\mathcal{I}(100, 1000)$
<b>Gradient Boosting</b>	Max Features <sup>a</sup> # Estimators <sup>b</sup> Max Depth <sup>c</sup>	$\mathcal{I}(1, 100)$ $\mathcal{I}(100, 1000)$ $\mathcal{I}(1, 100)$

<sup>a</sup> Max features determine the number of features considered for each split.

<sup>b</sup> Number of estimators specifies the total number of trees in the ensemble.

<sup>c</sup> Max depth controls the maximum depth of each tree.

**Linear Support Vector Regression (LIN)** SVM enhances the traditional support vector machine regressor by employing kernels to expand the feature space, thereby accommodating non-linear characteristics (Boser et al., 1992; James et al., 2013). Three different kernels are utilised: linear (discussed in this section), polynomial, and Radial Basis Function Support Vector Regression (RBF), which are presented in the subsequent sections.

The linear kernel is the simplest form of kernel function. It maps the input features directly without any transformation, making it suitable for linearly separable data. The decision boundary is a straight line (or hyperplane in higher dimensions), which simplifies the computation and interpretation.

The parameters for the LIN model include the choice of kernel (in this case, linear), the regularisation parameter  $C$ , and epsilon.

For details on the parameters and their ranges, see Table 4.

**Polynomial Support Vector Regression (POLY)** The polynomial kernel maps the input features into a higher-dimensional space using polynomial functions. This allows it to capture non-linear relationships between the features. The degree of the polynomial determines the model's complexity, enabling it to fit more intricate patterns in the data (James et al., 2013).

For details on the parameters and their ranges, see Table 4.

**Radial Basis Function Support Vector Regression (RBF)**

The RBF kernel, also known as the Gaussian kernel, maps the input features into an infinite-dimensional space. It measures the similarity between data points based on their distance, allowing it to capture complex, non-linear relationships. The RBF kernel is particularly powerful for handling data that is not linearly separable (James et al., 2013).

For details on the parameters and their ranges, see Table 4.

Table 4. Parameters for SVM Models with Different Kernels

Model	Parameter	Scope
<b>Support Vector (Linear)</b>	$C^1$ Epsilon <sup>2</sup>	$\mathcal{U}(0.1, 1.1)$ $\mathcal{U}(0.01, 1)$
<b>Support Vector (Polynomial)</b>	$C^1$ Epsilon <sup>2</sup> Gamma <sup>3</sup> Degree <sup>4</sup> Coef <sup>5</sup>	$\mathcal{U}(0.1, 1.1)$ $\mathcal{U}(0.01, 1)$ $\mathcal{U}(0.01, 100)$ $\mathcal{I}(1, 2)$ $\mathcal{U}(0.01, 10)$
<b>Support Vector (RBF*)</b>	$C^1$ Epsilon <sup>2</sup> Gamma <sup>3</sup>	$\mathcal{U}(0.1, 1.1)$ $\mathcal{U}(0.1, 1)$ $\mathcal{U}(0.01, 100)$

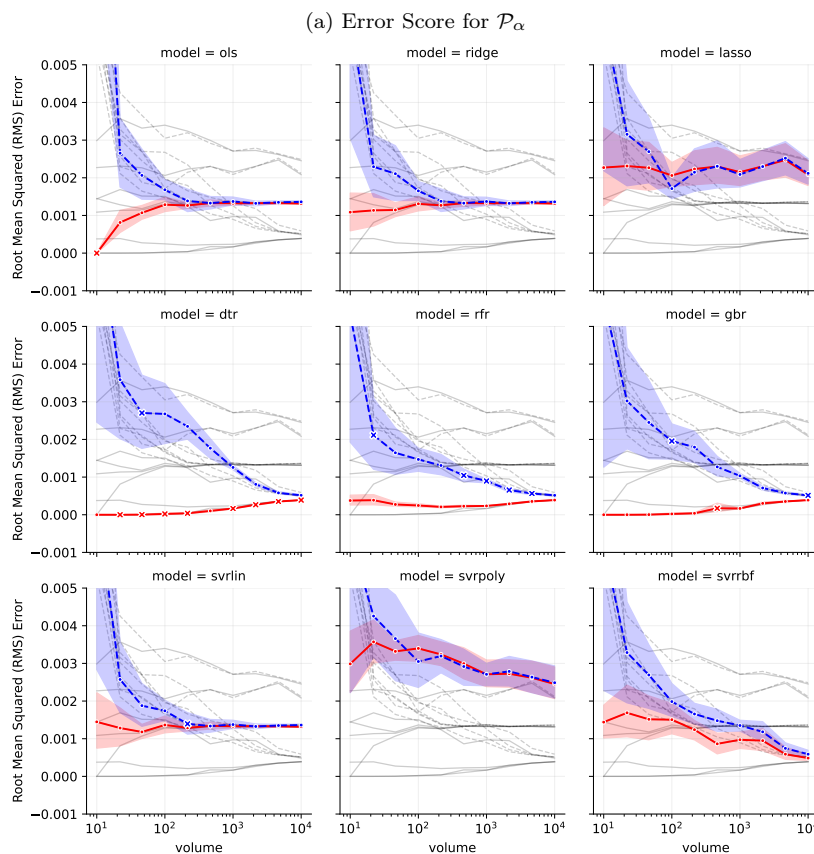
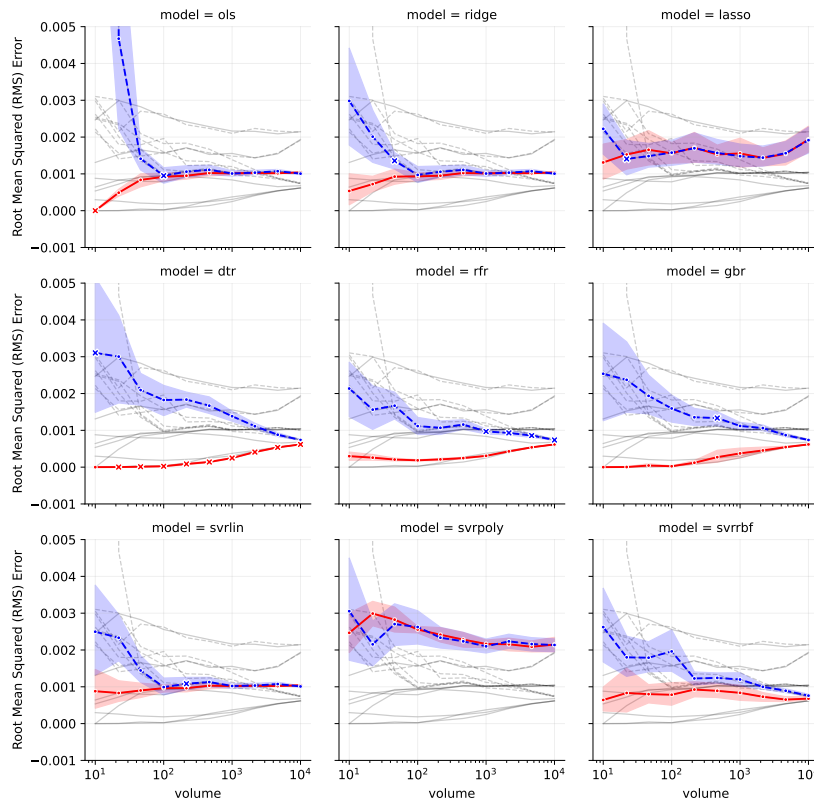
<sup>1</sup> Regularisation Parameter for SVM applies uniformly across other SVM models to ensure consistency in regularisation and sensitivity.

<sup>2</sup> The epsilon parameter defines a margin of tolerance around the regression line within which no penalty is assigned for prediction errors.

<sup>3</sup> The gamma parameter controls the influence of a single training example and determines the spread of the kernel. This affects the smoothness of the decision boundary; lower values imply a broader spread, while higher values imply a narrower spread.

<sup>4</sup> The degree parameter specifies the degree of the polynomial function used to transform the data, determining the flexibility of the decision boundary by defining the highest power of the input features.

<sup>5</sup> The coef0 parameter represents the independent term in the kernel function and adjusts the influence of higher-order versus lower-order terms.



Model Performance  
—●— Training    —●— Test

Fig. 1. Modelling Result for an Element of  $\mathcal{P}_x$  Showing RMSE for All Models Over the  $\mathcal{L}_V$

## 2.4 Data Pre-processing

In this work, two product datasets were compiled,  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$ , each representing different stages in the production process. The goal during pre-processing was to maintain at least 10,000 samples, which influenced the configuration of pre-processing steps. The pre-processing involved two key steps: feature selection and outlier removal.

Feature selection was conducted to identify and retain the most relevant variables by discriminating against those with low variance and high inter-correlation. A variance and correlation threshold was carefully estimated to ensure that the sample volume remained above 10,000, thus preserving the dataset's integrity while enhancing model performance.

Outlier removal was performed using the Interquartile Range (IQR) method. This method involved several steps:

1. Calculating the first quartile ( $Q_1$ ) and the third quartile ( $Q_3$ ).
2. Computing the IQR as  $IQR = Q_3 - Q_1$ .
3. Defining the lower bound as  $Q_1 - 1.5 \times IQR$ .
4. Defining the upper bound as  $Q_3 + 1.5 \times IQR$ .
5. Removing any data points that fell below the lower bound or above the upper bound.

Following the outlier removal, the dataset was scaled to standardise the features, ensuring that all variables contribute equally to the model's performance. Standardisation involved adjusting the features to have a mean of zero and a standard deviation of one, which is particularly important for machine learning models that are sensitive to the scale of the input data.

After these pre-processing steps, the dataset was split into training and testing sets for model evaluation. The separation between the training and test datasets effectively prevents overfitting, as is standard practice. To address underfitting, model parameters were allowed sufficient flexibility, managed through a trial-and-error approach. This approach was supported by a baseline guess informed by experience and conventional practices, ensuring that the models could adequately capture the underlying patterns in the data. The training set was used to fit the models, while the testing set was reserved for assessing the model's predictive performance on unseen data, thereby enhancing the model's ability to generalise and ensuring robust and reliable predictions.

## 2.5 Hyperparameter Estimation

A train-test split is implemented to ensure that the training process is conducted without any data leakage. Hyperparameter estimation is performed using a random grid search approach.

As demonstrated by Bergstra and Bengio (2012), the random search method offers significant advantages over conventional exhaustive grid search, particularly in terms of computational efficiency. It achieves comparable or even superior results while requiring fewer computational resources.

In a random grid search, hyperparameters are randomly sampled from a predefined list or distribution across a set number of iterations. This approach allows for a more

effective exploration of the parameter space, increasing the likelihood of identifying optimal hyperparameters.

## 2.6 Evaluation

The Root Mean Squared Error (RMSE) is employed to evaluate the performance of each subset model  $m_{i,j,k}^x$  on both training and test datasets. The RMSE depends on various factors, including the production database, iteration, volume, and model ( $m(x, i, j, k)$ ). The RMSE is defined by Equation 3:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

where  $n$  is the number of observations,  $y_i$  represents the actual values, and  $\hat{y}_i$  represents the predicted values.

This metric provides a robust measure of the model's predictive accuracy, with lower RMSE values indicating better model performance.

*Grid-plot Evaluation* To effectively interpret the complex system created by multiple layers of iterations, a structured framework for evaluating the results is necessary. The following approach is implemented in this work.

For each product dataset, a grid plot is created, containing one subplot for each model. Given that the number of models is 9, a 3 by 3 grid plot is used.

Each subplot, representing a specific model, displays how the train and test scores vary across the list of volumes. The y-axis shows the training and test error scores, while the x-axis represents the sampling volume. To enhance clarity, the x-axis is displayed on a logarithmic scale.

Additionally, each subplot shows the result scores for all iterations of the random grid search. The results are depicted as an area plot, with the mean indicated by a line (solid red for train data and dashed blue for test data).

To facilitate the comparison of model results within the same product dataset, faint but discernible lines are drawn in the background to represent other models. These background lines correspond to the score type. Consequently, the x- and y-axes of all subplots are synchronised and shared.

*Model Result Ranking* To provide a comprehensive overview of model performance, a heat map ranking plot is created. This heat map shows which models achieve the best test scores (lowest RMSE) for each volume.

Each model in  $\mathcal{L}_M$  is represented by an individual row on the y-axis.

Each volume in  $\mathcal{L}_V$  has a corresponding column on the x-axis, increasing incrementally. Each cell in the heat map displays the rank of the model, ranging from 1 to  $|\mathcal{L}_M|$ .

The top three models are colour-coded individually, while the remaining models share a single colour, as indicated by the colour bar on the right.

3. RESULTS AND DISCUSSION

The results and corresponding discussion are presented in this section. First, the outcomes of the data pre-processing stage are detailed in Section 3.1. This is followed by a description of the subset sampling process in Section 3.2. Next, the details of the hyperparameter search are provided in Section 3.3. The modelling scores are then presented in Section 3.4, followed by the analysis of the model rankings in Section 3.5.

3.1 Preprocessing

The pre-processing stage resulted in two datasets,  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$ , each ultimately containing an equal number of features. However, only half of these features were shared between the two datasets. During the feature selection process, different subsets of features were identified as relevant or superfluous for each dataset, leading to the retention of distinct feature sets in  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$ . Given that both datasets are derived from the same processes, it is expected that they share some underlying characteristics, which is reflected in the final selection of features.

The features retained after pre-processing for both  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$  are summarised in Tables 5 and 6. To ensure consistency in subsequent analyses, the features shared between  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$  were ordered and enumerated in a manner that aligns corresponding features representing the same properties. Each feature was assigned a consistent subscript across both datasets, allowing for direct comparison and facilitating the interpretation of the model results.

Table 5. Features Selection Outcome for  $\alpha$  and  $\beta$

Features	$\alpha$	$\beta$
Initial	22	22
Removed	14	14
Kept	8	8
Missing Value Ratio*	0.96	0.95

\* Constant value what ratio dictating the threshold for feature removal due to missing values.

Table 6. Outlier Removal Results for  $\alpha$  and  $\beta$

Samples	$\alpha$	$\beta$
Initial	12072	13597
Removed	1667	1543
Kept	10405	12054

3.2 Subset Sampling

Subsets  $\mathcal{S}_{i,j}^x$  were extracted for each volume in the list  $\mathcal{L}_V$  and sampled  $A$  times, resulting in  $|\mathcal{L}_V| \times A$  permutations per product dataset. The process begins with 10 samples and progresses to 10,000 samples in 10 steps, with the sampling volumes defined as  $\mathcal{L}_V = [v_1 = 10, v_2 = 22, v_3 = 46, v_4 = 100, v_5 = 215, v_6 = 464, v_7 = 1000, v_8 = 2154, v_9 = 4642, v_{10} = 10000]$ .

Thus,  $|\mathcal{L}_V| = 10$ , states that the length of the list of sampling volumes is 10.

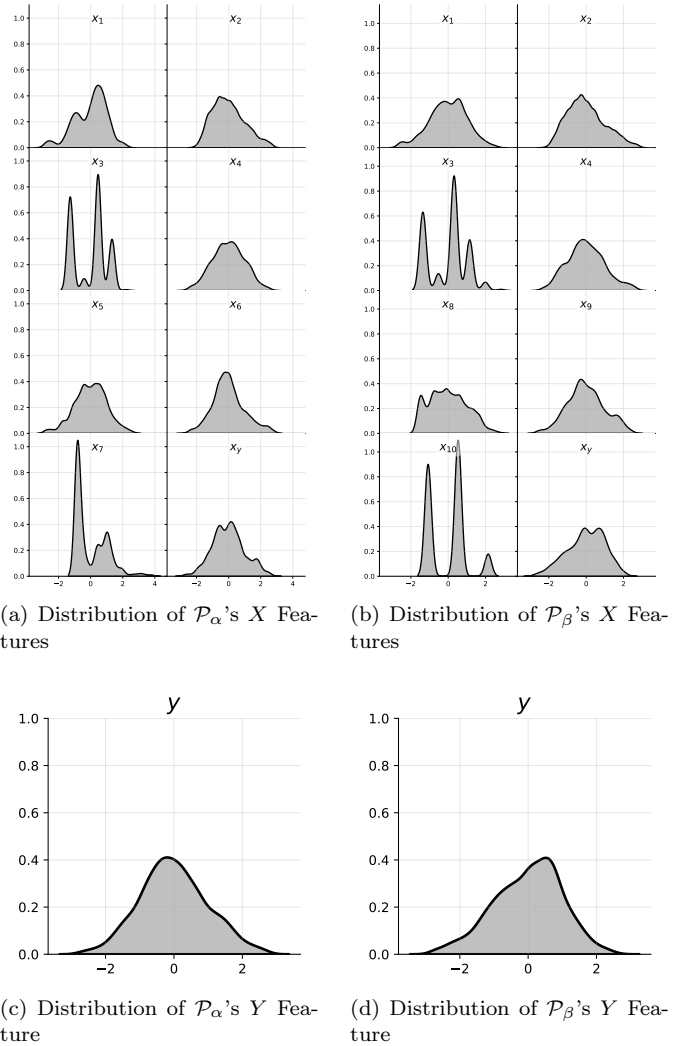


Fig. 2. Feature Distributions

3.3 Hyperparameter Estimation

For each element in the model list  $\mathcal{L}_M$ , the best fit was determined for every dataset in  $\mathcal{S}$ . Table 7 presents the mean values of the hyperparameters selected by the search process for the estimators.

3.4 Modelling Score

This subsection presents and discusses the variation in model scores as the volume of data in the production datasets increases. Figure 1 contains two sub-figures, 1a and 1b, which illustrate how the scores of all models change with increased data volume (see subsection 2.6.1 for details) for  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$ , respectively. In general, the initial data volumes exhibit considerable volatility and are not given significant weight in the overall analysis of results. This volatility is reflected in the variation of the scores. Unless explicitly stated, both  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$  are discussed collectively in the following analysis.

Most models show convergence between training and testing scores as data volume increases, with Lasso being a notable exception. Thus, it can be concluded that generalisation improves with an increase in data volume. How-



Table 7. Hyperparameter Result

Model	Parameter	$\alpha$				$\beta$			
		$10^1$	$10^2$	$10^3$	$10^4$	$10^1$	$10^2$	$10^3$	$10^4$
<b>Ridge</b>	Alpha	5.130	5.190	3.393	7.206	16.470	7.628	2.797	5.435
<b>Lasso</b>	Alpha	0.026	0.017	0.012	0.033	0.171	0.018	0.018	0.019
<b>Decision Tree</b>	Max Features	36.0	35.0	43.0	46.5	47.5	73.5	57.0	66.5
<b>Random Forest</b>	Max Features	45.0	61.0	54.0	27.0	49.5	39.5	59.0	44.5
	# Estimators	327.0	269.0	490.0	636.5	291.5	657.5	604.0	257.0
<b>Gradient Boosting</b>	Max Depth	34.5	37.5	10.5	23.0	56.0	63.5	15.0	25.5
	Max Features	39.5	6.5	25.5	50.5	42.5	4.5	34.0	61.5
	# Estimators	563.5	599.5	513.5	215.0	605.0	579.0	588.0	693.5
<b>Support Vector (Linear)</b>	C	0.047	0.032	0.071	0.154	0.082	0.088	0.057	0.101
	Epsilon	0.024	0.016	0.022	0.028	0.049	0.022	0.029	0.027
<b>Support Vector (Polynomial)</b>	C	0.614	0.281	0.254	0.443	0.145	0.346	0.208	0.289
	Epsilon	0.512	0.119	0.114	0.123	0.636	0.117	0.107	0.130
	Gamma	0.867	0.519	0.263	5.295	0.637	0.852	0.094	0.132
	Degree	2.0	3.0	2.0	2.0	2.0	3.0	2.0	2.0
	Coef0	0.187	0.197	1.117	0.607	0.091	0.644	0.570	0.461
<b>Support Vector (RBF)</b>	C	0.145	0.134	0.168	0.185	0.028	0.103	0.477	0.166
	Epsilon	0.017	0.021	0.024	0.019	0.034	0.031	0.033	0.017
	Gamma	0.300	0.182	0.082	2.177	0.260	0.077	0.136	0.979

ever, a point of diminishing returns in generalisation is discernible at different volumes and to varying degrees.

Lasso exhibits interesting behaviour, where training and testing scores converge quickly but then fluctuate as data volume increases, resulting in subpar overall performance. To explain this atypical behaviour, it is worth considering Table 1, which highlights Lasso's unique use of l1 regularisation. Furthermore, as shown in Table 7, the Lasso hyperparameter (Alpha) remains relatively static as volume increases, indicating that the method may be incompatible with this data or that the parameter ranges need revision.

The training score generally increases monotonically for all models except Lasso and RBF. The previous explanation for Lasso is insufficient when considering RBF, but since RBF shows significant improvement with larger volumes, this is not a major concern.

Examining the score spread, it is clear that ensemble models outperform the other models. A noteworthy runner-up is RBF, which, at larger volumes, approaches the performance of the ensemble models. This suggests that the data is well-suited for a non-parametric approach. Additionally, since all non-linear models, except POLY, show strong performance, it implies that the data has a non-linear nature. Alternatively, it may also indicate that the hyperparameter estimation and parameter ranges are insufficient to fully capture the underlying data patterns.

Referencing Table 1, it is possible that the robustness of ensemble methods to outliers gives them an advantage, suggesting that the pre-processing approach may have been inadequate for models sensitive to outliers.

### 3.5 Model Ranking

To make the overall performance of the models more discernible, their rankings across different data volumes are presented in Figs. 3 and 4 for  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$ , respectively (see subsection 2.6.2 for details).

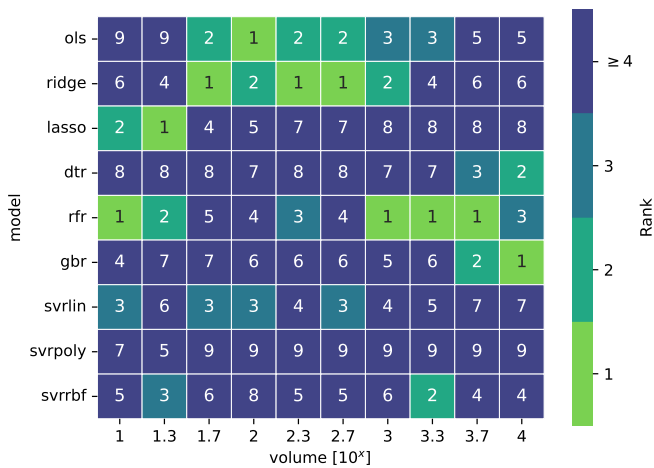
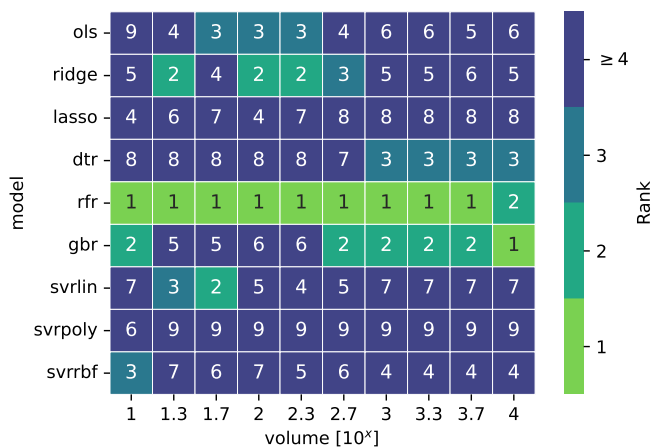
The ranking for  $\mathcal{P}_\beta$  shows convergence earlier than for  $\mathcal{P}_\alpha$ , meaning it stabilizes at a lower data volume. Specifically,  $\mathcal{P}_\beta$  reaches saturation at a volume of approximately  $10^{2.3}$ , while  $\mathcal{P}_\alpha$  does not reach saturation until a volume of around  $10^3$ .

The differences between  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$  can be attributed to the distinct sets of features retained after pre-processing, even though both datasets originate from the same underlying processes. Although each dataset contains approximately 10,000 samples and initially had an equal number of features, the final set of features for  $\mathcal{P}_\alpha$  and  $\mathcal{P}_\beta$  differs. This suggests that the selected features contribute differently to the modelling process, with certain features being more relevant or informative for one dataset than the other.

The divergence in feature selection underscores the varying impact of these features on the predictive models. Some features may offer greater predictive power or relevance depending on the specific context of each dataset, which in turn influences the point at which the model rankings stabilise

## 4. CONCLUSION

This study has demonstrated that most machine learning models show consistent improvement in predictive performance, as evidenced by a reduction in test RMSE,

Fig. 3. Model Rank for  $\mathcal{P}_\alpha$ Fig. 4. Model Rank for  $\mathcal{P}_\beta$ 

with increasing data volume. This finding underscores the importance of larger datasets in enhancing model generalisation, which is crucial in the context of industrial batch process modelling.

Among the models evaluated, ensemble models such as RFR and GBR consistently outperformed other models across various data volumes. Their robustness to outliers and ability to capture complex, non-linear relationships make them particularly effective for the datasets used in this study.

In contrast, non-ensemble models, especially those sensitive to outliers, generally underperformed relative to ensemble methods. Models employing L1 regularisation, such as Lasso, exhibited less stability and improvement in performance, suggesting that the chosen regularisation method may not be optimal for this data.

The disparity in performance between ensemble and non-ensemble models may be attributed to the latter's greater sensitivity to outliers. While stricter outlier removal could potentially enhance the performance of non-ensemble models, it would also reduce the number of available data samples, potentially limiting the study's scope.

The analysis of model rankings revealed a notable difference in the convergence times between models trained on

the  $\mathcal{P}_\beta$  and  $\mathcal{P}_\alpha$  datasets, despite both being derived from the same type of product and processes. This disparity illuminates the effectiveness of the introduced framework in detecting subtle variations in dataset complexity, particularly in terms of the variation of selected features, which significantly impacts the amount of data required for models to achieve saturation.

These findings highlight the importance of careful model selection and robust data pre-processing in industrial applications. Given the superior performance of ensemble models, they should be prioritised in future research within similar contexts. However, non-ensemble models may require more sophisticated pre-processing and parameter tuning to achieve comparable performance. These conclusions provide a foundation for further work aimed at improving model accuracy and robustness in industrial settings, potentially through enhanced data handling techniques and the inclusion of more complex models.

The findings of this study have broader implications beyond the iron and steel industry, extending to other sectors that rely on industrial batch processes, such as the chemical, pharmaceutical, and food processing industries. These industries share common challenges in managing complex, multi-product operations where model performance is heavily influenced by data volume and quality. The demonstrated superiority of ensemble models in handling non-linear relationships and their robustness to outliers suggests that similar approaches could be highly effective in these related industries. Moreover, the insights gained from addressing model sensitivity to outliers and the impact of dataset complexity can inform best practices in these sectors, where optimising process efficiency and product quality is equally critical. By adopting the strategies outlined in this study, industries with comparable batch processing challenges can enhance their predictive modelling capabilities, leading to more sustainable and efficient operations.

## REFERENCES

- Bailly, A., Blanc, C., Francis, É., Guillotin, T., Jamal, F., Wakim, B., and Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, 213, 106504. doi: 10.1016/j.cmpb.2021.106504.
- Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- Boser, B.E., Guyon, I.M., and Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 144–152.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Hastie, T., Tibshirani, R., Friedman, J.H., and Friedman, J.H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, volume 2. Springer.
- Hoerl, A.E. and Kennard, R.W. (1970a). Ridge regression: Applications to nonorthogonal problems. *Technometrics: a journal of statistics for the physical, chemical, and engineering sciences*, 12(1), 69–82.

- Hoerl, A.E. and Kennard, R.W. (1970b). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics : a journal of statistics for the physical, chemical, and engineering sciences*, 12(1), 55–67.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*, volume 112. Springer.
- Mählkvist, S., Ejenstam, J., and Kyprianidis, K. (2023). Cost-sensitive decision support for industrial batch processes. *Sensors*, 23(23), 9464.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Ramezan, C.A., Warner, T.A., Maxwell, A.E., and Price, B.S. (2021). Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data. *Remote Sensing*, 13(3), 368. doi:10.3390/rs13030368.
- Rendall, R., Chiang, L.H., and Reis, M.S. (2019). Data-driven methods for batch data analysis – A critical overview and mapping on the complexity scale. *Computers & Chemical Engineering*, 124, 1–13. doi:10.1016/j.compchemeng.2019.01.014.
- Ritchie, H., Roser, M., and Rosado, P. (2020). CO2 and greenhouse gas emissions.
- Shwartz-Ziv, R. and Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81, 84–90. doi:10.1016/j.inffus.2021.11.011.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288.
- Yoro, K.O. and Daramola, M.O. (2020). Chapter 1 - CO2 emission sources, greenhouse gases, and the global warming effect. In M.R. Rahimpour, M. Farsi, and M.A. Makarem (eds.), *Advances in Carbon Capture*, 3–28. Woodhead Publishing. doi:10.1016/B978-0-12-819657-1.00001-3.

# Machine Learning -based Optimization of Biomass Drying Process: Application of Utilizing Data Center Excess Heat

Henna Tiensuu\* Virpi Leinonen\*\* Jani Isokääntä\*\*  
Jaakko Suutala\*

\* *Biomimetics and Intelligent Systems Group, University of Oulu, P.O. BOX 4500, FI-90014, Oulu, Finland (e-mail: firstname.lastname@oulu.fi).*

\*\* *SFTec Oy, Oulu (e-mail: jani.isokaanta@sftec.fi)*

---

**Abstract:** This research explores the feasibility of using excess heat from data centers for biomass drying, enhancing the biomass energy value. A predictive model was developed to estimate exhaust air humidity from the dryer, indirectly indicating biomass moisture. Machine learning techniques, including linear regression model (LM), gradient boosting machines (GBM), eXtreme gradient boosting (XGBoost), random forest (RF), and multilayer perceptron (MLP), were used. Tree-based models GBM, RF, and XGBoost achieved a coefficient of determination ( $R^2$ ) of 0.88–0.89. Methods were enhanced with transparency through explainable artificial intelligence (XAI) techniques, which facilitated the analysis and visualization of humidity fluctuations. Key factors affecting drying efficiency include weather conditions, supply air humidity, and fan speed. The study provides actionable insights for optimizing the drying process, improving system air tightness, and advancing sustainable energy utilization through AI-driven solutions. The developed model enables future dynamic control of drying processes.

*Keywords:* process monitoring, explainable AI, predictive modeling, decision support, data centers

---

## 1. INTRODUCTION

Excess heat utilization from data center is widely investigated and is high interest of service providers as the climate targets, questions and demands are also increasing around the data center field. The utilization of data center excess heat especially in Nordics is concentrated to utilization in district heating (Wahlroos et al., 2018), as it is a well-established heating method in the area. District heating production in general still relies on fossil fuels and for example in Finland largely also to wood-based bioenergy. The use of wood-based bioenergy is increasing because of climate actions and increased price of CO<sub>2</sub> allowances. For energy production the wood based raw materials like wood chips are highly utilized. One option to utilize the waste heat of data center could be drying of biomass. As very even quality continuous heat flow is available from the datacenter and the drying of biomass as such requires lots of energy, but at the same time the energy value of the dried material increase by drying. The biomass drying has been considered earlier an option for waste utilization as an external process (Wahlroos et al., 2018) and now it was tested in the industrial scale.

In industrial settings, the challenges posed by demanding measurement conditions can significantly impact on quality of the data, and the data can come from several different sources. This multi-source data, often varied in format and structure, adds complexity to the analysis process. Despite these challenges, machine learning (ML)

methods are widely employed to diagnose, optimize, and enhance the quality and efficiency of complex manufacturing processes. However, the volume of data generated in industrial environments can be immense, further complicating the analysis process. ML techniques offer a means to extract valuable insights from these large and diverse datasets, enabling the prediction of process outcomes and the identification of relationships between different process parameters (He et al., 2009).

Previously, wood drying processes have been successfully optimized using ML methods across various applications. Ascher et al. (2022) have shown that ML methods have great potential towards modelling the biomass and waste gasification and pyrolysis processes and predicting the processes' product yields and properties. Chai et al. (2019) utilized feed forward neural network to simulate wood moisture content during the high-frequency drying, while Onsree and Tippayawong (2021) achieved accurate predictions of solid products yields from biomass torrefaction processes using gradient boosting machines (GBM). Studies have explored traditional physical methods in this area as well. For example, Li et al. (2012) investigated integrating a drying process into a power generation plant, using waste energy from process industries. These sources included low-grade heat, such as flue gas or hot cooling water for superheated steam. Additionally, Gebreegziabher et al. (2013) developed a physical model to determine the optimum drying level of wood chips. Furthermore, Li et al. (2022) examined using steel heat carriers for waste heat re-

covery and drying of high-moisture biomass in direct-fired power generation, achieving a 77.4% thermal efficiency in waste heat recovery and reducing fuel moisture content.

To fully leverage data-driven process modeling, the model results must be transparent to humans managing the manufacturing process. Hence, employing explainable machine learning methods is advisable (Goebel et al., 2018). Transparency should be increased, particularly when the model structure fails to explain the root causes behind the outcomes (Hagras, 2018). This transparency not only aids in understanding but also facilitates the optimization of the process, enabling humans to make informed decisions based on the clear explanations.

In this research, first, a simple linear regression model is trained to predict biomass moisture after the drying process using a small dataset. Then, a model for predicting the absolute humidity of the air exhausted from the dryer is developed using LM (James et al., 2013), eXtreme gradient boosting (XGBoost) (Chen and Guestrin, 2016), GBM (Hastie et al., 2001), random forest (RF) (Breiman, 2001), and multilayer perceptron (MLP) (Rosenblatt, 1958; Bishop, 1995) modeling methods, and the root causes behind the undesirable humidity levels are identified. The modeling results have been analysed and visualized using XAI methods.

This study represents a pioneering effort in the field, to our knowledge, being the first to utilize waste heat from data centers for biomass drying and optimize the process through exhaust air humidity prediction. The article is organized as follows: In Section 2, the biomass drying process is described. Section 3 explains the data collection process. The modeling and visualization methods are introduced in Section 4, while Section 5 presents model training and results. Finally, the discussion and conclusion are found in Sections 6 and 7, respectively.

## 2. BIOMASS DRYING PROCESS

The biomass drying using waste heat from a small data center, the Boden Type Data Center (BTDC) located in Boden, Sweden, was tested by installing an industrial scale dryer unit, ModHeat®, at the data center site. The biomass drying test setup is depicted in Fig. 1. The drying air, in this case waste heat from data center was taken straight from the BTDC (from area of 60 m<sup>2</sup>) to the drier. The dryer's fan was used to suck the warm air from the data center to the dryer. The dried test material was a normal energy wood chip.

The wet wood chips were fed to the hopper where the material was then fed by conveyer belt to the dryer's materials feeder. From the material feeder the biomass was fed to dryer. Inside the dryer material was circulated from drying level to another, as the dryer consist of five drying levels. The drying air from the data center was in contact with the material inside the dryer and was circulated in the opposite direction with the material flow. The drying air was circulated by the exhaust fan of the dryer and the moist exhaust air was directed to atmosphere after drying. The dried material discharged from the bottom level of the dryer to the conveyer and unloaded to the skip.

The main feature of the drying test campaign at the data center end was the hot aisle temperature (30/42°C) representing a traditional and high-performance computing data center, which was achieved by the varying the operation modes for the cooling equipment. At the dryer end, the main features were the air flow rate at inlet (50/100%), which was adjusted with the dryer's fan speed and the material feeding rate (50/75/100 %), which was adjusted to control the material flow volume to the dryer. The material flow rates were 2, 2.7, and 3 m<sup>3</sup>/h.



Fig. 1. The layout of the biomass drying test setup.

## 3. DATA COLLECTION

The test campaign lasted about two weeks from September 14<sup>th</sup> to 24<sup>th</sup>, 2020, during which the test setup was instrumented for a data collection. The data collection consists of data from four different systems, of which three were inputs. One of the systems was the internal system of the data center (provided by EcoCooling), which collected environmental data from the computer room evaporation units (CREC). The CREC handles the server cooling and kept the data center aisles at target temperatures. The second data collection system was the ModHeat® dryer data collection system gathering the information about temperature, relative humidity, and air speed going in and out of the dryer. The third system (provided by RISE Research Institutes of Sweden) gathered the data about temperature, relative humidity, air volume flows entering and exhausting the dryer. As the inlet and outlet temperatures and relative humidity of air were the main data points for dryer's efficiency evaluation both data loggers were used to avoid data loss. The fourth system collected the data related to electric power usage data of the servers, the dryer, the CREC units and local weather data at the datacenter site. The sensor placement of data collection is depicted in Fig. 2. In addition to monitored data, the moisture content of the material, initial moisture content as wet and after drying, was measured sample based manually in intervals during the test campaign. However, the dataset was limited, comprising only 15 pairs of values before and after drying.

## 4. METHODS

Several different machine learning methods have been used in this research to find insights from the data. In addition, to interpret these models, methods of XAI are used.

### 4.1 Machine learning models

In this research, first, a linear regression model based on ordinary least square regression was used in biomass moisture prediction and the model's ability to generalize new



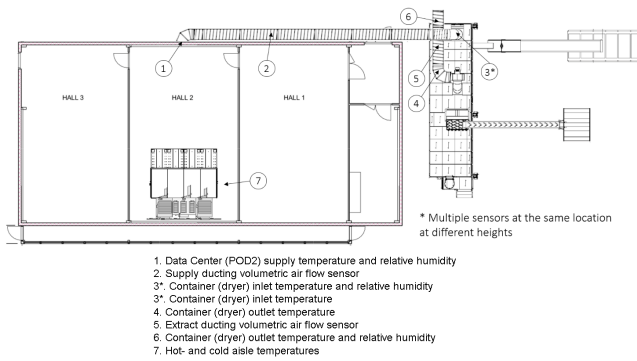


Fig. 2. The data collection points and placement of the sensors at the test site.

data have been inspected based on leave-one-out-cross-validation (LOOCV) (Hastie et al., 2001). Subsequently, five predictive ML methods - LM, GBM, RF, XGBoost and MLP - were employed to predict the absolute humidity of exhaust air, utilizing 10-fold cross-validation during training. For each model except LM, hyperparameter tuning was conducted using a grid-based approach. The tuning process aimed to optimize the models' performance metrics, such as root mean squared error (RMSE), using a cross-validation.

All of these methods are types of supervised machine learning, where the algorithms are trained on labeled data to predict the desired variable. These methods were employed with the goal of identifying the factors that affect biomass end moisture and predicting accurately the exhaust air humidity after one and half hour, which can aid in optimizing the initial settings for the drying process.

GBM, RF, and XGBoost are tree-based ensemble methods (Zhou, 2014). They work by combining the predictions of multiple individual models (trees in the case of RF and GBM, and boosted trees in the case of XGBoost) to make a final prediction. In RF, each tree is built independently using a random subset of features and samples, and the final prediction is made by averaging or voting over the predictions of all trees (Breiman, 2001). In GBM and XGBoost, trees are built sequentially, with each new tree trained to correct the errors made by the previous trees. (Hastie et al., 2001; Chen and Guestrin, 2016) MLP is a type of artificial neural network (ANN) that consists of multiple layers of interconnected neurons, including an input layer, one or more hidden layers, and an output layer. Each neuron in an MLP is connected to every neuron in the adjacent layers, allowing for complex non-linear mappings between input and output features (Bishop, 1995).

#### 4.2 Visualization with XAI

Explainable artificial intelligence (XAI) methods have been used to enhance the understandability of the modeling results (Goebel et al., 2018). With XAI methods, the interpretation of the black box models can be increased (Apley and Zhu, 2019). MLPs are generally considered to be less interpretable compared to tree-based models like RF, GBM, and XGBoost. This is because the internal workings of a neural network are highly complex and opaque, making it difficult to understand how individual features contribute to the model's predictions. In contrast,

decision trees are more transparent and can provide insights into which features are important for making predictions.

Transparency can be increased by providing information about the strength of the importance of each feature of the model regardless of the modeling method used. In this study, model-independent SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017) and Accumulated Local Effects Plots (ALE) (Friedman, 2002) techniques are used. SHAP based on the concept of Shapley values from cooperative game theory to provide a unified approach for explaining the predictions of any machine learning model. They quantify the contribution of each feature by systematically varying its value while keeping others constant and comparing the model's prediction to a baseline. With ALE, the average effect of features on the predictions of ML model can be visualized. The interactions between features are also important, and the strength of interactions can be estimated.

## 5. RESULTS

In this research, there are two study cases, each with its own dataset. The first study case focuses on predicting biomass moisture, while the second study case focuses on predicting the humidity of the exhaust air.

### 5.1 Dataset

In these two cases, about two weeks' dataset were available, although with some limitations. In the biomass moisture study case, the dataset comprises 15 pairs of biomass moisture measurements, recorded manually as spot checks. The pairs were selected with a minimum one-hour delay between start and end moisture readings during the period of time.

In the absolute humidity of exhaust air study case, the data is divided iteratively into training and test sets using 10-fold cross-validation. The independent features used for model training, are 5-minute average values and there are 12 features each containing 140 values. Features are listed in Table 1. The dependent variable, exhaust air absolute humidity, represents the absolute humidity after one and a half hours, relative to the starting state, which corresponds to the drying time of one portion of the biomass.

### 5.2 Study case: Biomass moisture

Rather than predicting end moisture directly, a linear regression model was employed to predict the difference between start and end moisture. Variable selection during modeling resulted in a final model with 9 variables predicting this moisture difference. Figure 3 displays the actual versus predicted moisture differences, while Figure 4 provides detailed model results. Despite an 89% explanation of variance ( $R^2$ ), the adjusted  $R^2$  (0.69) suggests the presence of non-significant variables, notably POD2 supply temperature ( $p = 0.01$ ), supply flow ( $p = 0.01$ ), and outdoor absolute humidity ( $p = 0.02$ ), identified as significant based on p-values. The coefficient plot, with 95% confidence intervals, indicates considerable uncertainty in variables, with some intersecting the reference line at 0,

Table 1. Description of the independent variables

Name	Description
container_supply_hum	incoming air humidity measured in the end of air pipe (%)
container_supply_temp	incoming air temperature measured in the end of air pipe (°C)
dryer_fan_speed	incoming air flow rate to dryer(%)
dryer_feed_rate	material feeding rate to the dryer (%)
humidity_outdoor_absolute	the absolute value of outdoor air humidity ( $g/m^3$ )
humidity_outdoor_relative	the relative value of outdoor air humidity (%)
POD2_supply_hum	incoming air humidity from the data center to air pipe (%)
POD2_supply_temp	incoming air temperature from the data center to air pipe (°C)
precipitation_quantity_absolute	the absolute quantity of half on hour cumulative precipitation (mm)
Supply_air_temp	incoming air temperature from pipe to dryer (°C)
Supply_air_relative_humidity	incoming relative air humidity from pipe to dryer (%)
supply_flow	Air flow from data center to dryer measured in the middle of the pipe ( $m^3/s$ )

signifying insignificance in Fig. 5. Notably, absolute precipitation quantity and relative exhaust air humidity were non-significant, while others influenced biomass moisture differences. The model performance was evaluated using LOOCV, yielding an  $R^2$  value of only 0.57. The generalization of this model appears to be quite poor.

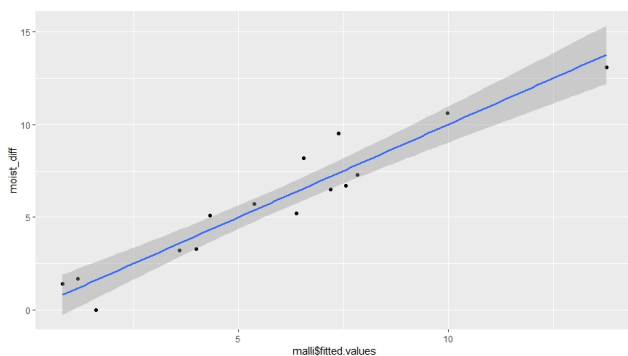


Fig. 3. A comparison between observed and predicted biomass moisture differences with LM.

**MODEL INFO:**  
Observations: 15  
Dependent Variable: moist\_diff  
Type: OLS linear regression

**MODEL FIT:**  
 $F(9, 5) = 4.51, p = 0.06$   
 $R^2 = 0.89$   
Adj.  $R^2 = 0.69$

Standard errors: OLS

	Est.	S. E.	t val.	p
(Intercept)	-26.86	7.22	-3.72	0.01
POD2_supply_temp	31.20	7.53	4.15	0.01
container_supply_hum	44.29	15.62	2.84	0.04
container_exhaust_hum	-10.72	5.28	-2.03	0.10
supply_flow	22.37	5.38	4.15	0.01
supply_air_relative_humidity	22.37	10.53	2.12	0.09
exhaust_air_relative_hum	6.89	5.33	1.29	0.25
exhaust_air_temp	36.89	14.01	2.63	0.05
humidity_outdoor_absolute	-80.51	22.71	-3.55	0.02
precipitation_quantity_absolute	4.59	3.04	1.51	0.19

Fig. 4. LM performance evaluation.

### 5.3 Study case: Absolute humidity of exhaust air

In this study, the absolute humidity of exhaust air is predicted using five ML models: LM, GBM, RF, XGBoost, and MLP. Each model was trained and evaluated using 10-fold cross-validation, and all models except LM underwent hyperparameter tuning via grid search. The LM achieved an average RMSE of 0.468 and average  $R^2$  of 0.799. GBM

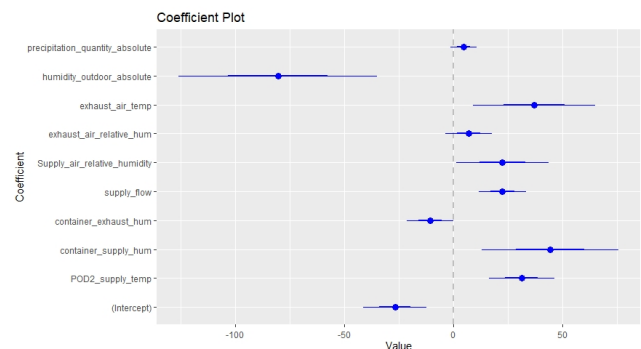


Fig. 5. A coefficient plot of LM with 95% confidence intervals showing the significance of independent features.

Table 2. Comparison of modeling results. Higher  $R^2$  values and smaller RMSE values indicate higher predictive accuracy

Model	$R^2$ Mean $\uparrow$ (sd $\downarrow$ )	RMSE Mean $\downarrow$ (sd $\downarrow$ )
LM	0.799 (0.079)	0.468 (0.113)
GBM	0.880 (0.070)	0.359 (0.107)
RF	0.894 (0.074)	0.345 (0.110)
XGBoost	0.894 (0.075)	0.345 (0.101)
MLP	0.726 (0.005)	0.738 (0.003)

showed the best results with a shrinkage of 0.1, interaction depth of 6, and 200 trees, yielding an average RMSE of 0.359 and an average  $R^2$  of 0.880. RF achieved optimal performance with the number of variables to randomly sample as candidates at each split (mtry) of 4, resulting in an average RMSE of 0.345 and an average  $R^2$  of 0.894. XGBoost performed optimally with a learning rate of 0.3, maximum depth of 3, and 200 rounds, achieving an average RMSE of 0.345 and an average  $R^2$  of 0.894. MLP demonstrated its best performance with a single hidden layer size of 10 and weight decay of 0.01, achieving an average RMSE of 0.738 and an average  $R^2$  of 0.726. All the modeling results, including mean and standard deviation values of  $R^2$  and RMSE for each of five models, are shown in Table 2. As can be seen, the most accurate models are RF and XGBoost based on the RMSE and  $R^2$  values.

The SHAP feature importance plots for both models are presented in Fig. 6 and Fig. 7. In the RF plot, the X-axis values indicate the average absolute SHAP value across all samples, representing the average impact of each feature has on the model's output across different data points. A higher absolute SHAP value suggests that the feature has a stronger influence on the model's prediction. Conversely,

In XGBoost, the X-axis labeled "Gain" denotes the gain in model performance achieved by splitting on each feature during the training process. This gain is calculated based on various metrics such as information gain or reduction in impurity. Features with higher gain values indicate that splitting on those features leads to greater improvements in the model's predictive accuracy. In both plots, the top four features are the same but in a different order. The most important features are the amount of humidity expelled from the data center (POD2\_supply\_hum), absolute humidity of outdoor (humidity\_outdoor\_absolute), absolute quantity of precipitation (precipitation\_quantity\_absolute), and the speed of the dryer fan (dryer\_fan\_speed). In other words, these features exhibit the most significant influence on the target variable, the exhaust air humidity from the dryer.

The correlation between features has been calculated and is shown as a heatmap in Fig. 8. It can be seen that the most correlated feature pairs are those measuring the supply air humidity and temperature. Additionally, the speed of the fan correlates quite strongly (0.80) with the amount of humidity expelled from the data center (POD2\_supply\_hum), the relative outdoor humidity (0.89), and the absolute precipitation quantity (0.87). The RF model will be interpreted in more detail next.

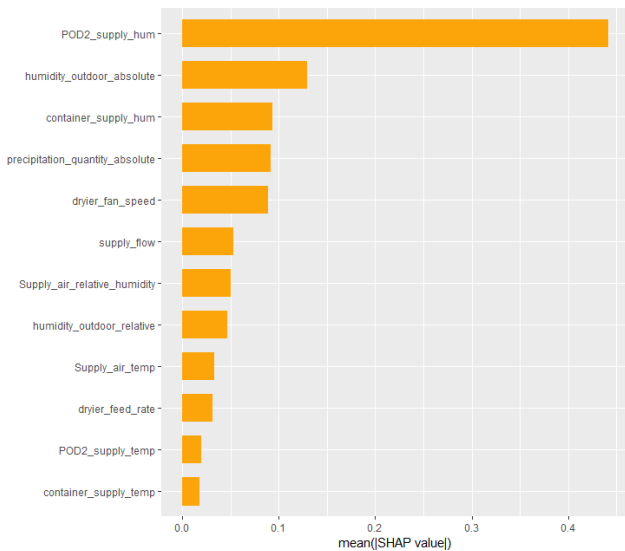


Fig. 6. The SHAP feature importance for RF.

The effect of each feature on the dependent variable is visualized with Accumulated Local Effects (ALE) plots. Figure 9 presents the ALE plot illustrating the influence of the input feature 'POD2\_supply\_hum' on the exhaust air absolute humidity. ALE shows the main effect of the feature at a certain value compared to the average prediction of the exhaust air absolute humidity of RF and also the distribution of data points. As can be seen, the predicted exhaust air absolute humidity is higher when the POD2 supply humidity is below 12.8%. Similarly, a clear boundary can be observed where the feature begins to have a negative effect on the exhaust air absolute humidity in Figs. 10 to 13. These plots show that the exhaust air absolute humidity is higher when the absolute humidity of outdoor (humidity\_outdoor\_absolute) is below  $6.7g/m^3$ , the container supply humidity is below

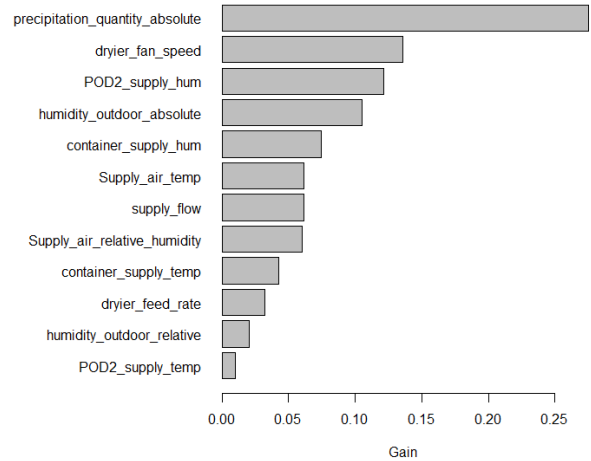


Fig. 7. The SHAP feature importance for XGBoost.

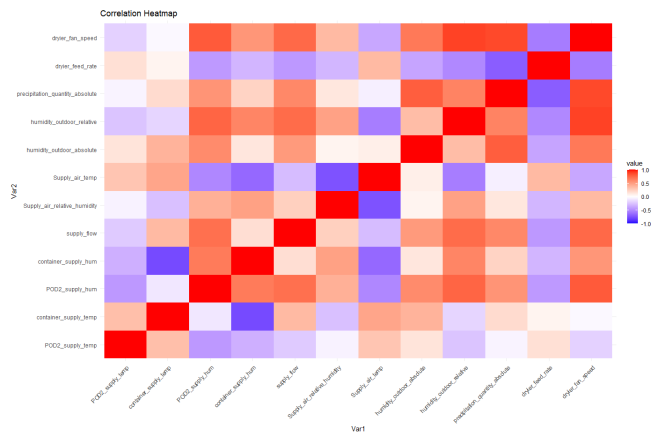


Fig. 8. Correlation heatmap of the independent features.

15.3%, the absolute quantity of precipitation is below  $139mm$  (precipitation\_quantity\_absolute), and the speed of the dryer fan (dryer\_fan\_speed) should be 50% instead of 100%. The modeling results are influenced not only

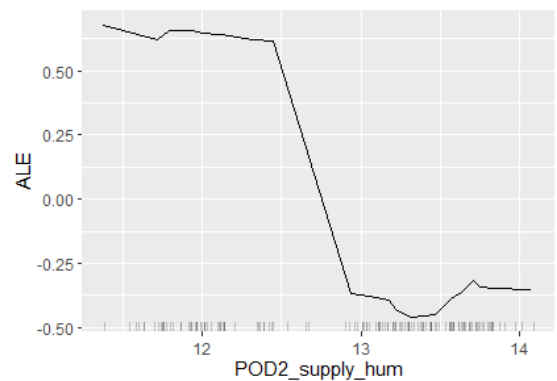


Fig. 9. ALE of the POD2 supply humidity (%) on the predicted humidity of RF (black solid line) are shown on the Y-axis, with the distribution of data points represented by black bars along the X-axis.

by individual features but also by their interactions. In addition to analyzing features independently, it's essential to study their interactions. These interactions are depicted

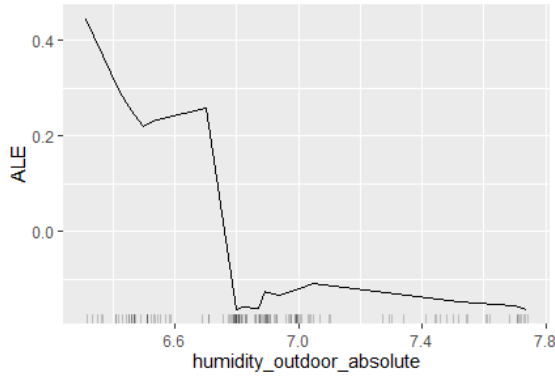


Fig. 10. ALE of the absolute humidity of outdoor ( $g/m^3$ ) on the predicted humidity of RF.

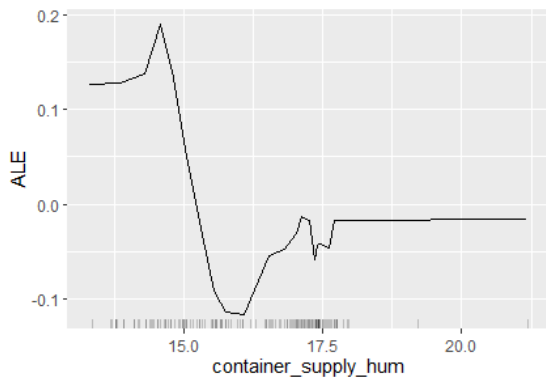


Fig. 11. ALE of the container supply humidity (%) on the predicted humidity of RF.

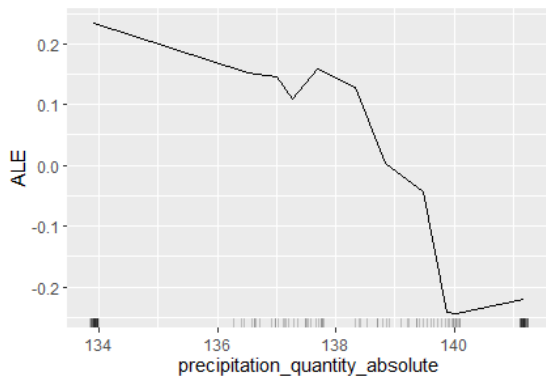


Fig. 12. ALE of the absolute quantity of precipitation (mm) on the predicted humidity of RF.

in Fig. 14, illustrating the strength of interplay between features. Each feature is assigned an interaction strength value, ranging from 0 to 1, indicating the proportion of explained variance of  $f(x)$ . A value of zero denotes no interaction, while one indicates complete dependence on the interaction of the given feature. Notably, POD2 supply humidity, container supply humidity, relative and absolute humidity of outdoor, and absolute precipitation quantity exhibit the highest interaction strengths. Moreover, Figure 15 visually identifies the strongest interaction partners for each feature. For instance, the POD2 supply humidity's primary interaction partner is the absolute humidity of outdoor. Furthermore, interactions between absolute pre-

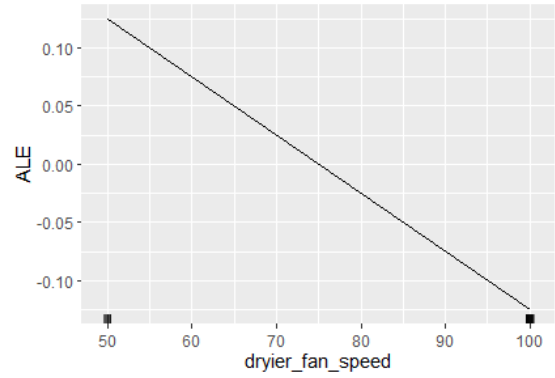


Fig. 13. ALE of the speed of the fan (%) on the predicted humidity of RF.

cipitation quantity, relative humidity of outdoor, container supply air humidity, and the speed of fan are also observed. But overall, the interaction rates are quite low, indicating a relatively low interaction between the features.

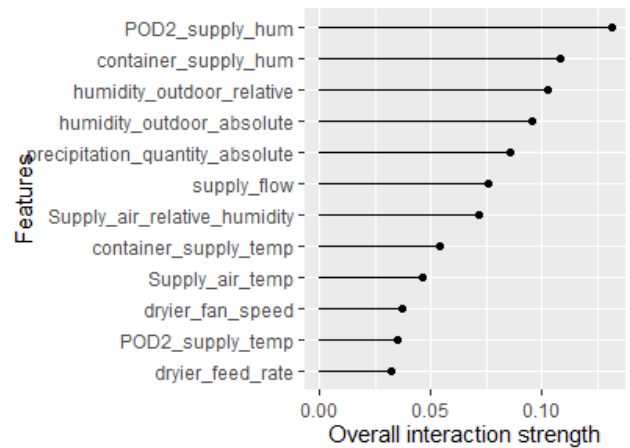


Fig. 14. Overall interaction strengths for each feature independently analyzed, derived from the RF model.

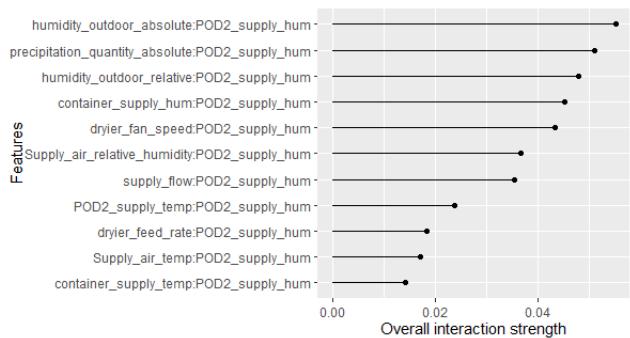


Fig. 15. The 2-way interaction strengths between the most important feature, POD2 supply humidity, and the other features, as determined by the RF model.

For individual predictions, SHAP summary plots provide a visualization of the SHAP values for each feature. A summary of SHAP values illustrates the impact of each feature on the predicted exhaust air absolute humidity in Fig. 16. A positive SHAP value for a feature indicates



that its presence increases the prediction, while a negative SHAP value indicates that its presence decreases the prediction. The length of the bar shows the magnitude of the impact. Longer bars mean a greater impact on the prediction. It's evident that lower POD2 supply humidity, absolute humidity outdoors, and container supply humidity values correspond to higher predicted exhaust air absolute humidity. Conversely, higher values result in lower predicted humidity. As can be seen, with the low POD2 supply humidity, absolute humidity of outdoors and container supply humidity values, the predicted exhaust air absolute humidity is higher. Similarly, with higher values, the predicted humidity is lower.

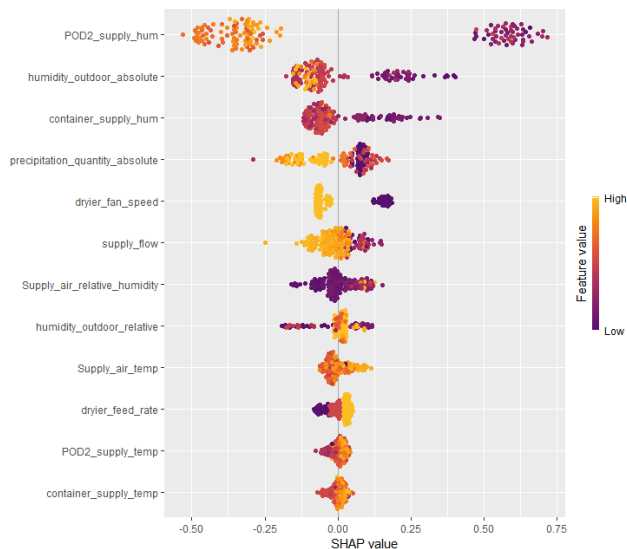


Fig. 16. Summary of SHAP values illustrating the impact of each feature on RF model predictions.

## 6. DISCUSSION

The prediction of biomass end moisture poses a significant challenge due to the inherent difficulty indirectly measuring this target variable. In this study, a small set of initial measurements was measured both before and after the drying process. However, the small size of the dataset and the quality issues associated with the measurements contribute to the overall low modeling accuracy. Consequently, an alternative approach focusing on the exhaust absolute humidity was adopted, as it proves to be more easily measurable. To mitigate the challenges posed by the limited datasets, cross-validation techniques were utilized to enhance the robustness of model evaluation and minimize the potential for overfitting.

The relationship between exhaust air humidity and moisture captured from the material reveals valuable insights into the drying process. Specifically, in the absolute humidity of exhaust air study case, it was assumed that higher humidity levels in the exhaust air corresponded to greater moisture extraction from the biomass, resulting in a drier biomass output. Root cause analysis of exhaust air humidity reveals underlying root causes affecting biomass moisture, as well.

The most important factor behind the humidity expelled from the dryer is the amount of humidity expelled from the

data center (referred to as the POD2 supply humidity), as indicated by the RF feature importance analysis. Additionally, the absolute quantity of precipitation (referred to as precipitation quantity absolute) emerges as another crucial factor, as identified by the XGBoost analysis. Both features rank in the top four in both analyses. Variations in air humidity along the path from the data center to the dryer, as well as the initial supply humidity (referred to as the container supply humidity) at the onset of drying, play significant roles in both feature analyses. In both analyses, the top two most important features include the absolute outdoor humidity and the absolute quantity of precipitation, which are quite strongly correlated based on the heatmap. Therefore, weather conditions, particularly rainy weather, are found to reduce the effectiveness of the biomass drying process due to alterations in the quality of leaked air.

The SHAP summary plots provided valuable insights into the impact of each feature on predicted exhaust air absolute humidity. The analysis revealed that lower POD2 supply humidity, absolute humidity outdoors, and container supply humidity were associated with higher predicted exhaust air absolute humidity, as indicated by positive SHAP values. Conversely, higher values of these features were linked to lower predicted humidity, as evidenced by negative SHAP values. These findings highlight the significance of controlling and optimizing factors such as POD2 supply humidity and outdoor humidity to manage and regulate exhaust air absolute humidity levels during the drying process.

Through data analysis, three threshold values were identified to increase the exhaust absolute humidity: POD2 supply humidity below 12.8%, container supply humidity below 15.3%, and the speed of the dryer fan set at 50% instead of 100%. Additionally, favorable weather conditions related to air humidity were identified: the absolute humidity outdoors should be below  $6.7 \text{ g/m}^3$ , and the absolute quantity of precipitation should be below 139 mm. If these thresholds are exceeded, process settings can be adjusted to optimize biomass drying, resulting in a drier biomass output.

Moreover, weather conditions interact with other independent features, highlighting the importance of improving system air tightness to enhance drying efficiency. Shorter and more airtight pipes between the data center and dryer, along with airtight design of air entry and output for the dryer, can mitigate the adverse effects of humid weather conditions, further optimizing biomass drying effectiveness.

The developed model accurately predicts exhaust air absolute humidity for one hour and a half into future with present settings. Additionally, both the RF and XGBoost models demonstrate strong generalization capabilities, as evidenced by the metrics calculated on an independent test dataset obtained through 10-fold cross-validation. These models explain 89% of the variation in exhaust air absolute humidity using the selected independent features. Nonetheless, further research is needed to refine the model's ability to indirectly predict biomass end moisture. With access to more data for model training, testing, and validation, dynamic modeling could enable long-term



predictions of end moisture and adaptive system control, further optimizing the entire biomass drying process.

## 7. CONCLUSION

It has been shown that the developed models, even with a small dataset, showcased the potential of ML and its capabilities in developing process control for biomass drying, particularly through tree-based ML methods. As the test campaign showed the ambient conditions effect strongly to the dryer's efficiency especially when low temperature drying is done. And when operating a multivariant system like drying in a changing environment there needs to be possibilities to measure and react to the changes. The machine learning could provide a tool for handling these changes by predicting the upcoming changes based on the history data or forecast data e.g. of weather conditions or material initial moisture content changes. More exact evaluation of process parameters variations enables more accurate prediction of e.g., production capacities.

## ACKNOWLEDGEMENTS

The work for this paper was done as part of ArctiqDC project, with financial support from the European Regional Development Fund via the Interreg Nord program. RISE Research Institutes of Sweden is acknowledged for their contribution to data collection.

## REFERENCES

- Apley, D.W. and Zhu, J. (2019). Visualizing the effects of predictor variables in black box supervised learning models.
- Ascher, S., Watson, I., and You, S. (2022). Machine learning methods for modelling the gasification and pyrolysis of biomass and waste. *Renewable and Sustainable Energy Reviews*, 155, 111902.
- Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press.
- Breiman, L. (2001). Random forests. *Mach. Learn.*, 45(1), 5–32. doi:10.1023/A:1010933404324.
- Chai, H., Chen, X., Cai, Y., and Zhao, J. (2019). Artificial neural network modeling for predicting wood moisture content in high frequency vacuum drying process. *Forests*, 10(1). doi:10.3390/f10010016.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, 785–794. Association for Computing Machinery, New York, NY, USA. doi:10.1145/2939672.2939785.
- Friedman, J.H. (2002). Stochastic gradient boosting. *Comput. Stat. Data Anal.*, 38(4), 367–378. doi:10.1016/S0167-9473(01)00065-2.
- Gebreegziabher, T., Oyedun, A.O., and Hui, C.W. (2013). Optimum biomass drying for combustion – a modeling approach. *Energy*, 53, 67–73. doi:10.1016/j.energy.2013.03.004.
- Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., and Holzinger, A. (2018). Explainable ai: The new 42? In A. Holzinger, P. Kieseberg, A.M. Tjoa, and E. Weippl (eds.), *Machine Learning and Knowledge Extraction*, 295–303. Springer International Publishing, Cham.
- Hagras, H. (2018). Toward human-understandable, explainable ai. *Computer*, 51(9), 28–36. doi:10.1109/MC.2018.3620965.
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York.
- He, S., Wang, G.A., and Li, L. (2009). Quality improvement using data mining in manufacturing processes. In J. Ponce and A. Karahoca (eds.), *Data Mining and Knowledge Discovery in Real Life applications*, 357–372. I-Tech.
- James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning*, volume 112. Springer.
- Li, G., Li, Z., Yin, T., Ren, J., Wang, Y., Jiao, Y., and He, C. (2022). Drying biomass using waste heat from biomass ash by means of heat carrier. *BioResources*, 17(3), 5243–5254.
- Li, H., Chen, Q., Zhang, X., Finney, K.N., Sharifi, V.N., and Swithenbank, J. (2012). Evaluation of a biomass drying process using waste heat from process industries: A case study. *Applied Thermal Engineering*, 35, 71–80. doi:10.1016/j.applthermaleng.2011.10.009.
- Lundberg, S.M. and Lee, S.I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, 30, 4765–4774. Curran Associates, Inc.
- Onsree, T. and Tippayawong, N. (2021). Machine learning application to predict yields of solid products from biomass torrefaction. *Renewable Energy*, 167(C), 425–432.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386–408.
- Wahlroos, M., Pärssinen, M., Rinne, S., Syri, S., and Manner, J. (2018). Future views on waste heat utilization – case of data centers in northern europe. *Renewable and Sustainable Energy Reviews*, 82, 1749–1764. doi:10.1016/j.rser.2017.10.058.
- Zhou, Z.H. (2014). Ensemble methods. *Combining pattern classifiers*. Wiley, Hoboken, 186–229.

# Enhanced Anomaly Detection in Aero-Engines using Convolutional Transformers

M.R. Babaei, A.D. Fentaye, K. Kyprianidis

*School of Business, Society and Engineering (EST), Mälardalen University, 72220 Västerås, Sweden*

*(e-mail: [mohammad.reza.babaei@mdu.se](mailto:mohammad.reza.babaei@mdu.se), [amare.desalegn.fentaye@mdu.se](mailto:amare.desalegn.fentaye@mdu.se), [konstantinos.kyprianidis@mdu.se](mailto:konstantinos.kyprianidis@mdu.se))*

---

**Abstract:** Gas turbines are vital in power generation and propulsion systems. However, these engines are exposed to complex and variable operating conditions, which makes early and accurate fault detection essential for predictive maintenance and minimizing unplanned downtime. This paper proposes a novel approach that combines convolutional neural networks (CNNs) with transformer architectures to address these challenges. The proposed Convolutional transformer model aims to enhance the accuracy and robustness of turbofan fault classification by integrating the feature extraction capabilities of CNNs with the contextual learning strengths of transformers. Through rigorous experiments, we seek to demonstrate our approach's performance in classification accuracy and generalization across different operating conditions. We utilize a comprehensive synthetic dataset, C- MAPSS, derived from multiple aircraft engine units as the benchmark for this study. The results for the proposed model show an accuracy of 99.6% on the test dataset. The outcome has the potential to be extended and fine-tuned for different types of gas turbines for diverse applications.

*Keywords:* fault classification, gas turbine engines, attention mechanism, convolutional neural network, machine learning

---

## 1. INTRODUCTION

The reliable operation of turbofan gas turbines is crucial to ensuring the safety and efficiency of modern aviation. As these engines are exposed to complex and variable operating conditions, early and accurate fault detection is essential for predictive maintenance and minimizing unplanned downtime. Fault classification plays a pivotal role in identifying and rectifying potential issues before they become significant failures.

The classification of faults in turbofan gas turbines is confronted with multiple significant challenges (Fentaye et al. (2019); X. Yang et al. (2023); Z. Yang et al. (2013)). Firstly, these engines are fitted with numerous sensors that monitor a range of parameters such as temperature, pressure, and vibration, leading to the generation of high-dimensional data streams that are challenging to process and analyze effectively. Secondly, the operational behavior of turbofan engines is characterized by complex temporal patterns, as faults often develop progressively over time, necessitating the capture of these temporal dependencies for accurate classification. Additionally, faults can progress at varying rates depending on operating conditions, complicating the development of models that can generalize across different scenarios. Moreover, many datasets exhibit an imbalance in fault classes, where some fault types are significantly more prevalent than others, potentially leading to biased classification models. Finally, for practical implementation, fault classification models must operate in real-time or near-real-time, necessitating the use of efficient algorithms capable of managing large volumes of streaming data.

In recent years, various machine learning (ML) and deep learning approaches have been explored for turbofan gas turbine data exploration and fault classification. A study by Xie et al. (2023) shows that feature extraction is a vital step in the ML pipeline, involving the transformation of raw sensor readings into informative and discriminative features that can improve the performance of classification models. Barrera et al. (2022) combines gained information from clustering with an auto-encoder for anomaly detection. Traditional techniques like support vector machines (SVM) and random forests have been used with reasonable success. For instance, Zhou et al. (2015) applied SVMs to classify different fault conditions with reasonable accuracy. Losi et al. (2022) were utilized random forest (RF) to predict gas turbine trips based on the snapshot matrix of records. However, these methods often struggle to oversee the high dimensionality and complex temporal dependencies inherent in turbine data.

The advent of deep learning has brought significant advancements in fault classification. CNNs have been highly effective due to their ability to extract hierarchical features from raw sensor data automatically Sateesh Babu et al. (2016) for remaining useful life (RUL) estimation. Long short-term memory (LSTM) networks have also been employed in research by Cao et al. (2021) to capture temporal dependencies, further improving classification performance. A study by Arpit et al. (2019) suggests a variant of LSTM to overcome well-known limitations like the vanishing gradient discussed by Pascanu et al. (2012). Fentaye et al. (2021) implements a Deep CNN to detect and isolate multiple gas path faults with over 96% accuracy. Despite these advances,

existing models still face challenges balancing local feature extraction and capturing global dependencies within the data.

This paper proposes a novel approach that combines CNNs with transformer architectures to address these challenges. Transformers, originally introduced by Vaswani et al. (2017) for natural language processing, have shown exceptional performance in capturing long-range dependencies and contextual information. By integrating the feature extraction capabilities of CNNs with the contextual learning strengths of transformers, our proposed convolutional transformer model aims to enhance the accuracy and robustness of turbofan gas turbine fault classification.

We will conduct a comparative analysis of the Convolutional transformer model against other state-of-the-art models, including a feed-forward transformer and standalone CNN models. Through rigorous experiments, we seek to demonstrate our approach's superiority in classification accuracy and generalization across different operating conditions. The proposed model should be able to explore and learn imbalanced dataset with complex dynamics effectively. In addition, address common issues with recurrent neural networks and provide the possibility to extend into deep neural networks (DNNs). This research contributes to advancing fault classification methodologies and has significant implications for implementing predictive maintenance in real-world aviation scenarios.

## 2. METHODOLOGY

### 2.1 Data collection

This study uses a comprehensive synthetic dataset published by Arias Chao et al. (2021) from multiple aircraft engine units, the NASA Commercial Modular Aero-Propulsion System Simulation (C-MAPSS). C-MAPSS is a turbofan engine degradation dataset which is a widely used dataset in the field of predictive maintenance and prognostics.

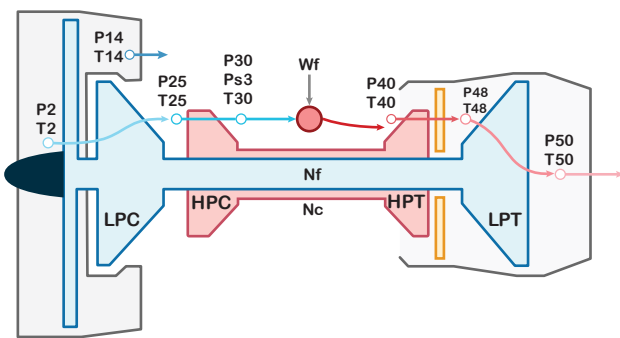


Fig. 1. A schematic of a turbofan engine gas turbine

It was developed to simulate realistic engine degradation scenarios under various operational conditions, providing researchers and practitioners with a valuable resource for developing and benchmarking fault detection and remaining useful life (RUL) prediction algorithms. The dataset comprises multiple sets of time-series data, capturing different degradation patterns across several turbofan engines. The C-MAPSS dataset is instrumental for advancing ML and statistical methods in predictive maintenance, allowing for the

creation and validation of models that can predict equipment failures, optimize maintenance schedules, and ultimately enhance the reliability and safety of aerospace systems. Nevertheless, there has been relatively limited research focused on fault diagnostics and classification in gas turbines, indicating a significant opportunity for further exploration of advanced ML methods in anomaly detection within this context. The variables in the dataset are categorized into scenario descriptors  $\omega$ , health parameters  $\theta$ , measurements  $\hat{x}_s$ , and virtual sensors  $\hat{x}_v$ . Figure 1 shows a schematic of turbofan engine gas turbine.

This dataset consists of 9 units (2, 5, 10, 11, 14, 15, 16, 18, and 20). The units 2, 5, and 10 are subject to degradation of the efficiency of the high-pressure turbine (HPT) (Fault 1). The units 11, 14, 15, 16, 18, and 20 are subject to the flow and efficiency of low-pressure turbine (LPT) in addition to HPT efficiency degradation (Fault 2). Figure 2 shows the health parameters corresponding to the faults occurring within the dataset.

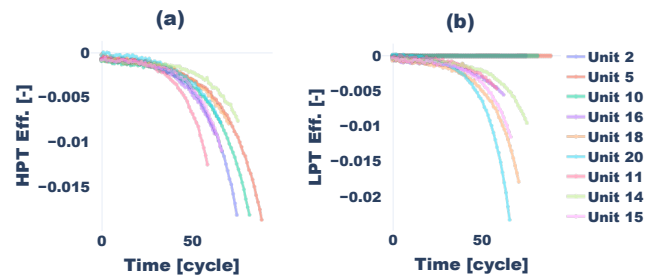


Fig. 2. Health parameters for LPT and HPT during the cycles

To eliminate the effects of ambient variables on the measure, the data goes through a correction process based on Volponi (2020). For example, the ambient temperature affects the system temperatures, pressure, and, consequently, shaft speed and fuel flow. Although the effect of the inlet temperature on the pressure is negligible, according to (1-3), the corrected values of temperatures, fuel flow, and shaft speed replaced the raw values to analyze the system's state during the operation.

$$\theta = \frac{T_{amb}}{T_{ref}}, \quad T_{ref} = 288.15 \text{ K} \quad (1)$$

$$T_i^{corr} = T_i / \theta \quad (2)$$

$$W_f^{corr} = \frac{W_f}{\sqrt{\theta}}, \quad N_i^{corr} = \frac{N_i}{\sqrt{\theta}} \quad (3)$$

Figure 3 demonstrates the distribution of total pressure at LPT outlet pressure (P50) during cycles for different units in the dataset.

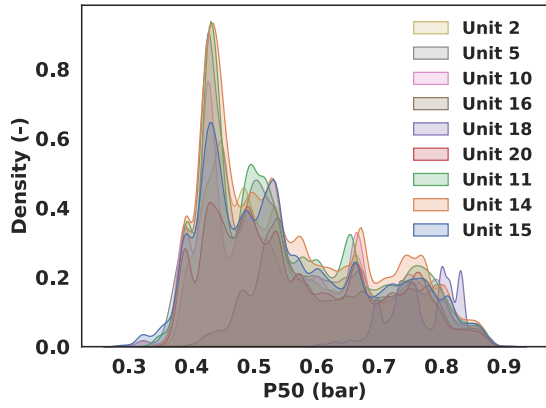


Fig. 3. LPT outlet pressure distribution for the engine units in the dataset

2.2 Attention mechanism

The transformer model, initially introduced by Vaswani et al. (2017), has become a cornerstone in natural language processing and sequence modeling tasks. The standard transformer consists of an encoder and a decoder, comprising stacked self-attention and feedforward (FF) neural network layers. Self-attention layer takes key, query, and value as the input. On the other hand, convolutional neural networks (CNNs) are well-known for their ability to extract essential features effectively from local spatial hierarchies. This study proposes a combination of a multi-head attention (MHA) layer and a one-dimensional convolutional layer instead of FF for time-series data diagnostic. As the convolutional layer extracts essential features, the multi-head attention layer enhances the model's performance by directing each head to focus on distinct aspects of the data, thereby providing complementary information crucial for accurate fault classification.

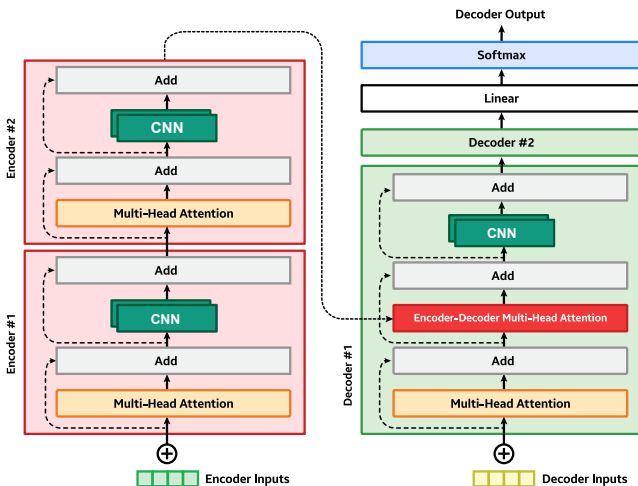


Fig. 4. A schematic of the transformer neural network structure including convolutional layers

Figure 4 represents a schematic of the proposed transformer neural network. A SoftMax activation layer is implemented at the output layer to yield a probability that each fault occurs within the system.

2.3 Training methodology

CNNs are more likely to experience overfitting than traditional ML models or simpler neural network architectures. One way to tackle this issue is to include dropout layers in the CNN structure. This will disable some of the connections between layers in the training process to achieve a generalized model. Another way is to define loss value criteria to terminate the training if the model has not improved in multiple consequence epochs. This research applies a dropout rate of 10-20% to the neural network structure. The loss was calculated using the categorical cross-entropy function, as shown in (4).

$$CCE(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log(\hat{y}_{i,j}) \quad (4)$$

Where N is the number of samples, C is the number of classes,  $y$  is a Boolean to show if the sample  $i$  belongs to class  $j$ , and  $\hat{y}$  is the predicted probability that the sample  $i$  belongs to class  $j$ .

Classification accuracy, a fundamental performance metric in supervised learning, is calculated as the ratio of correctly predicted instances to the total number of instances evaluated. Formally, accuracy ( $A$ ) is defined as:

$$A = \frac{1}{N} \sum_{i=1}^N l(y_i = \hat{y}_i) \quad (5)$$

Where in (5),  $N$  denotes the total number of predicted class labels, and  $l(\cdot)$  is an indicator function that returns 1 if the argument is true and 0 otherwise. This metric effectively quantifies the proportion of correct predictions made by the model, providing an intuitive measure of its performance.

A batch size of 256 was selected to gain consistent training, testing, and validation results. Considering the batch size, a learning rate of  $\eta = 5e-5$  is assumed for the training. The input data of the neural network is a window of input variables with the size  $w = 10$  and created using past and present records of measured values.

Starting from epoch 70, an early stopping condition on the validation loss was applied to terminate the training process after 5-20 consecutive epochs without any improvement. The models were optimized by Adam optimizer. The Adam optimizer offers adaptive learning rates and efficient computation, which makes it suitable for large datasets and models with sparse gradients. Its bias correction, robustness to noisy gradients, and user-friendly hyperparameters ensure fast convergence and broad applicability in deep learning.

A summary of the model structures studied in this research is shown in Table 1.

Table 1. Summary of the model structures of Vanilla CNN, CNN, FF Transformer, and Convolutional Transformer

1. Vanilla CNN				2. CNN				3. FF Transformer				4. Convolutional Transformer			
#	layer type	unit/ filter	activation function	#	layer type	unit/ filter	activation function	#	layer type	unit/ filter	activation function	#	layer type	unit/ filter	activation function
1	Conv1D	100	ReLU	1	Conv1D	100	ReLU	1	MHA	4	-	1	MHA	4	-
2	Avg Pooling	-	-	2	Conv1D	100	ReLU	2	Dense**	100	ReLU	2	Conv1D	100	ReLU
3	Dropout*	10%	-	3	Dropout	10%	-	3	Dropout	10%	-	3	Dropout	20%	-
4	Conv1D	100	ReLU	4	Dense**	2	SoftMax	4	MHA	4	-	4	MHN	4	-
5	Avg Pooling	-	-					5	Dense**	100	ReLU	5	Conv1D	100	ReLU
6	Dense**	2	SoftMax					6	Dense	100	ReLU	6	Dense**	2	SoftMax
								7	Dense	2	SoftMax				

\* Dropout rate  
 \*\* A flattened layer is applied before the dense layer

Table 2 provides an overview of the data split into train, validation, and test sets. The test data set uses units 10 and 16 to ensure that the model evaluates with a balanced data set and that there is no data leakage between train/validation and test splits. Of the remaining units, 80% and 20% were utilized for training and validation, respectively.

Table 2. An overview of the training, validation, and test datasets

Dataset	Unit	Fault Mode
training and validation	2	HPT
	5	
	11	HPT + LPT
	14	
	15	
test	18	HPT
	20	
	10	HPT + LPT
	16	HPT + LPT

Data were normalized from 0 to 1 using a Min-Max scaler to avoid saturation in the activation function and achieve faster convergence. We used TensorFlow 2.13 for the implementation of our neural network models, running on a system with an AMD Ryzen 9 5950X CPU (16 cores, 32 threads), an NVIDIA GeForce RTX 3090 GPU (24 GB of GDDR6X memory), and 128 GB of DDR4 RAM at 3200 MHz.

### 3. RESULTS AND DISCUSSION

The K-Best feature selection was implemented to analyze the effect of each variable on fault classification. The K-Best score for each variable and the cumulative score are presented in Fig. 5. The results indicate that P50 contributes the most to capturing the variance of the output, and the LPT outlet temperature (T50) has the most negligible effect on capturing the output. In other words, the impact of T50 on fault classification is insignificant, and 13 variables are enough to predict the output. Notably, 14 out of 45 variables were selected as the model's input. In this case, the effect of degradation on T50 was not considerable and resulted in a lower K-Best score. However, based on the physics of the problem, we decided to keep T50 as the model's input.

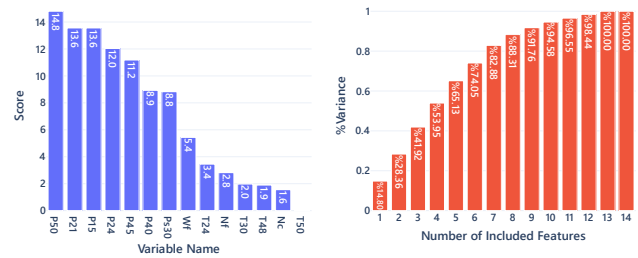


Fig. 5. K-Best feature selection results: variable scores (left) and cumulative score (right)

Figure 6 shows the result of the K-Means clustering. The optimal number of clusters was identified by the knee method. As shown, the input variables can be classified into 4 clusters. The variables in each cluster are expected to have a similar dynamic behavior. For example, the HPT outlet temperature (T48) and the T50 have similar dynamic behavior and can be assumed to be a cluster. Table 3 shows the selected features for the classification problem.

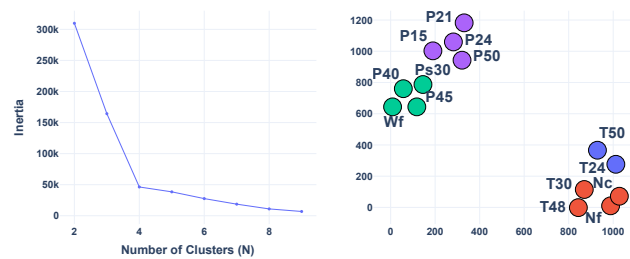


Fig. 6. K-Means inertia value based on the number of clusters (left) and K-Means for  $n_c = 4$  (right)

Figure 7 presents training and validation loss curves. Reducing the learning rate resulted in smoother training curves and stable results. The model reaches the best accuracy around epoch 120. Starting from epoch 90, the training loss starts to fluctuate around the local minimum. No significant improvements were observed after 120 epochs, and the training was terminated to avoid overfitting. In the presence of a dropout layer, a lower validation loss is expected compared to the training loss.



Table 3. Description and notation of variables within the dataset based on Arias Chao et al. (2021)

#	Symbol	Description	Units
1	Wf	Fuel flow	kg/s
2	Nf	Physical fan speed	rpm
3	Nc	Physical core speed	rpm
4	T24	Total temperature at LPC outlet	K
5	T30	Total temperature at HPC outlet	K
6	T48	Total temperature at HPT outlet	K
7	T50	Total temperature at LPT outlet	K
8	P15	Total pressure in bypass-duct	kPa
9	P21	Total pressure at fan outlet	kPa
10	P24	Total pressure at LPC outlet	kPa
11	Ps30	Statics pressure at HPC outlet	kPa
12	P40	Total pressure at burner outlet	kPa
13	P50	Total pressure at LPT outlet	kPa
14	P45	Total pressure at HPT outlet	kPa

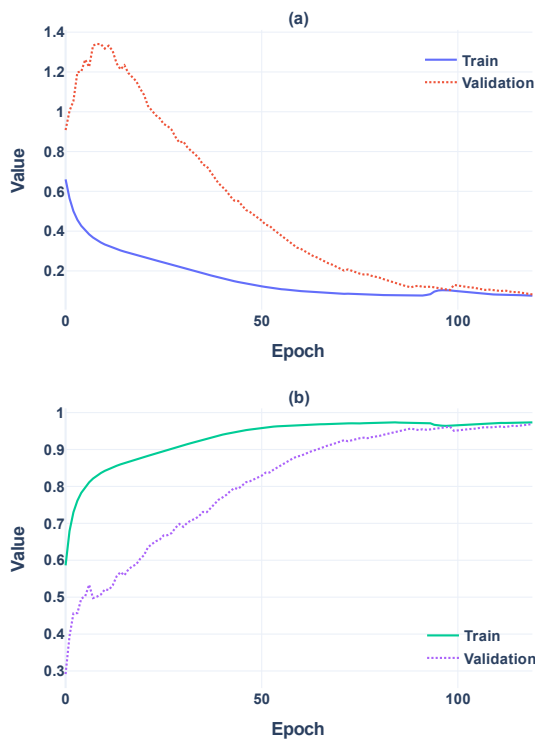


Fig. 7. Convolutional Transformer loss (a) and accuracy (b) curves for the training and validation datasets

Imbalanced classes in the training dataset cause a temporary increase in the validation loss at the preliminary stages of the training. The prediction results at the beginning are biased toward the class with a higher population. On the other hand, the categorical cross-entropy loss is sensitive to the probabilities assigned to the correct classes. If the model's confidence in its predictions fluctuates (even if predictions are mostly correct), the loss can increase while accuracy improves. This behavior was observed across all the models studied during this research. Training and validation loss/accuracy

curves for Vanilla CNN, CNN, and FF Transformer are presented in Appendix A.

The validation and test accuracies prove these claims since both show remarkable results in fault classification. Figure 8 shows the confusion matrix for the transformer model with convolutional layers. As shown, imbalanced data for fault classes resulted in a considerable gap between Fault 1 and 2 accuracies for models 1-3. These models can identify Fault 1 with higher accuracy despite a higher population of classes with Fault 2. As mentioned, it has simple features and only corresponds to HPT fault, while Fault 2 is a combination of HPT and LPT failures.

Hence, a convolutional transformer (model 4) shows significant improvement in finding the complex correlation in the data considering the imbalance training/validation dataset. It is good to note that Fault 2 is a combination of Fault 1 and other faults and distinguishing one from another is challenging.

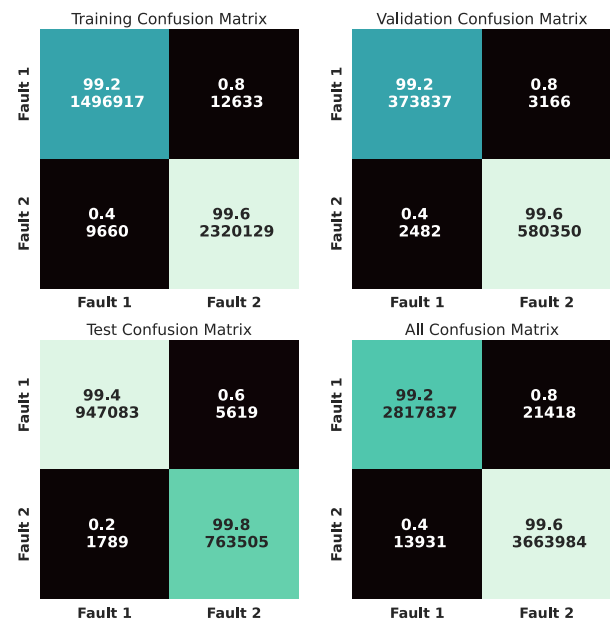


Fig. 8. Convolutional transformer confusion matrix for training, validation, testing, and total dataset

Interestingly, the accuracy gained from the test dataset is promising and higher than both training and validation in some cases. Considering that the test dataset is from separate turbofan engine units and was never used during the training season, it is another clue for the quality of training and evaluation. Implementing the dropout layer avoids overfitting and makes the model learn the pattern instead of overfitting the current dataset. In this case, the final model performs better in predicting the test dataset.

Figure 9 summarizes the overall accuracy of the evaluated modes in this study. CNN model lacks MaxPooling layers and has higher parameters compared to Vanilla CNN. In this case, the CNN model performs slightly better in accuracy. The transformer model with feed-forward layers improved overall accuracy by 0.8%.



Fig. 9. Overall fault classification accuracy of Vanilla CNN, CNN, FF Transformer, and Convolutional Transformer neural networks

What is interesting here is the accuracy of 98.89% achieved by substituting convolutional layers by feed-forward and adding the attention mechanism. A combination of multi-head attention and convolutional layers improves spatial pattern recognition and results in around 99.5% overall accuracy in fault classification. Despite the highest accuracy in the Convolutional transformer model, the FF transformer offers a simple model structure with fewer parameters and respectful accuracy. However, compared to the Convolutional transformer, this structure fails to handle an imbalanced dataset.

#### 4. CONCLUSIONS

Gas turbines play a vital role in transportation and energy systems. Modern engineering considers optimal design, efficient operation, sustainability, and safety simultaneously when manufacturing turbofan engines, with the progress of ML and artificial intelligence process optimization and maintenance going toward automation to balance safety and diagnostics costs. This research investigated the application of recent developments in NLP and adopted a minimal structure to provide a fault classifier for the gas turbine. The final platform fulfills the needs of AI assistants to identify the faults within the turbofan engines based on a snapshot of measured data in real-time. The transformer takes advantage of the multi-head attention mechanism and convolutional layers to investigate a horizon of information and capture the essential information. The results for the proposed model show an accuracy of 99.6% on the test dataset. In future work, a complete dataset of faults will be investigated to analyze the effect of each measurement on anomaly detection. While proposing a general model for the gas turbine diagnostic is challenging, providing more information will help the model recognize more patterns and distinguish the system behavior under different conditions. The outcome can be fine-tuned for distinct types of gas turbines for various kinds of applications.

#### REFERENCES

- Arias Chao, M., Kulkarni, C., Goebel, K., and Fink, O. (2021). Aircraft Engine Run-to-Failure Dataset under Real Flight Conditions for Prognostics and Diagnostics. *Data*, 6(1), Article 1. doi:10.3390/data6010005
- Arpit, D., Kanuparthi, B., Kerg, G., Ke, N. R., Mitliagkas, I., and Bengio, Y. (2019). *h-detach: Modifying the LSTM Gradient Towards Better Optimization* (arXiv:1810.03023). arXiv. doi:10.48550/arXiv.1810.03023
- Barrera, J. M., Reina, A., Mate, A., and Trujillo, J. C. (2022). Fault detection and diagnosis for industrial processes based on clustering and autoencoders: A case of gas turbines. *International Journal of Machine Learning and Cybernetics*, 13(10), 3113–3129. doi:10.1007/s13042-022-01583-x
- Cao, Q., Chen, S., Zheng, Y., Ding, Y., Tang, Y., Huang, Q., Wang, K., and Xiang, W. (2021). Classification and prediction of gas turbine gas path degradation based on deep neural networks. *International Journal of Energy Research*, 45(7), 10513–10526. doi:10.1002/er.6539
- Fentaye, A. D., Baheta, A. T., Gilani, S. I., and Kyprianidis, K. G. (2019). A Review on Gas Turbine Gas-Path Diagnostics: State-of-the-Art Methods, Challenges and Opportunities. *Aerospace*, 6(7), Article 7. doi:10.3390/aerospace6070083
- Fentaye, A. D., Zaccaria, V., and Kyprianidis, K. (2021). Aircraft Engine Performance Monitoring and Diagnostics Based on Deep Convolutional Neural Networks. *Machines*, 9(12), Article 12. doi:10.3390/machines9120337
- Losi, E., Venturini, M., Manservigi, L., Ceschini, G., Bechini, G., Cota, G., and Riguzzi, F. (2022). Prediction of Gas Turbine Trip: A Novel Methodology Based on Random Forest Models. *Journal of Engineering for Gas Turbines and Power*, 144(031025). doi:10.1115/1.4053194
- Pascanu, R., Mikolov, T., and Bengio, Y. (2012). On the difficulty of training Recurrent Neural Networks. *30th International Conference on Machine Learning, ICML 2013*.
- Sateesh Babu, G., Zhao, P., and Li, X.-L. (2016) *Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life* (S. B. Navathe, W. Wu, S. Shekhar, X. Du, X. S. Wang, and H. Xiong, Eds.; Vol. 9642, pp. 214–228). Springer International Publishing. doi:10.1007/978-3-319-32025-0\_14
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). *Attention Is All You Need* (arXiv:1706.03762). arXiv. doi:10.48550/arXiv.1706.03762
- Volponi, A. J. (2020). *Gas Turbine Parameter Corrections*. Springer International Publishing. doi:10.1007/978-3-030-41076-6
- Xie, J., Sage, M., and Zhao, Y. F. (2023). Feature selection and feature learning in machine learning applications for gas turbines: A review. *Engineering Applications of*

*Artificial Intelligence*, 117, 105591.  
doi:10.1016/j.engappai.2022.105591

Yang, X., Cheng, K., Zhao, Q., and Wang, Y. (2023). Unknown fault detection for EGT multi-temperaturesignals based on self-supervised feature learning and unary classification. *Frontiers in Energy*, 17(4), 527–544. doi:10.1007/s11708-023-0880-x

Yang, Z., Wong, P. K., Vong, C. M., Zhong, J., and Liang, J. (2013). Simultaneous-Fault Diagnosis of Gas Turbine Generator Systems Using a Pairwise-Coupled Probabilistic Classifier. *Mathematical Problems in Engineering*, 2013(1), 827128. doi:10.1155/2013/827128

Zhou, D., Zhang, H., and Weng, S. (2015). A New Gas Path Fault Diagnostic Method of Gas Turbine Based on Support Vector Machine. *Journal of Engineering for Gas Turbines and Power*, 137(102605). doi:10.1115/1.4030277



Fig. A2. CNN confusion matrix for training, validation, testing, and total dataset

Appendix A: Confusion matrix and training/validation loss curves for Vanilla CNN, CNN, and FF transformer



Fig. A1. Vanilla CNN confusion matrix for training, validation, testing, and total dataset

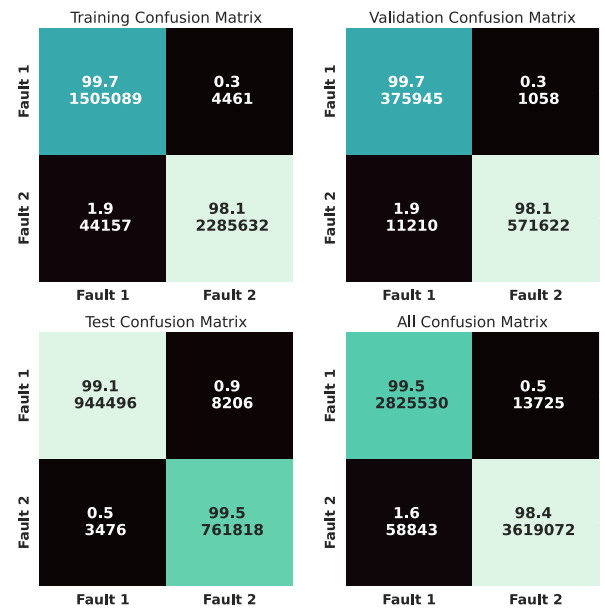


Fig. A3. FF transformer confusion matrix for training, validation, testing, and total dataset

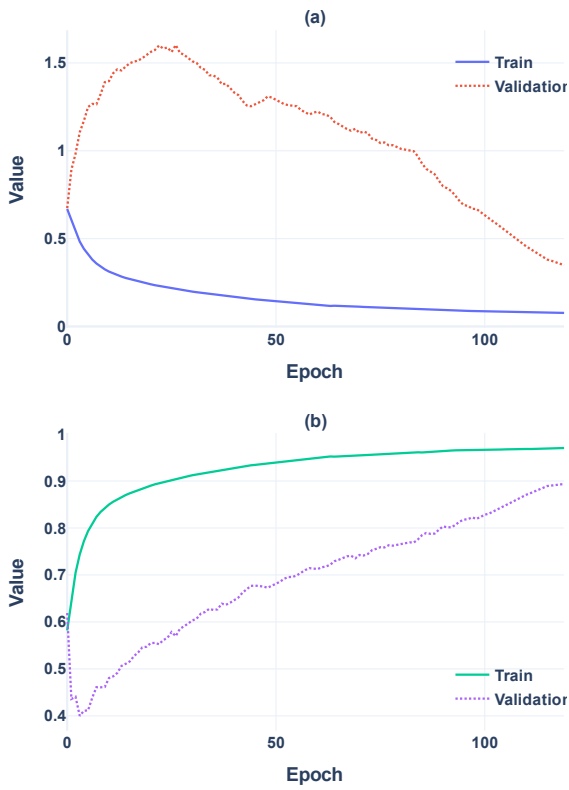


Fig. A4. Vanilla CNN loss (a) and accuracy (b) curves for the training and validation datasets

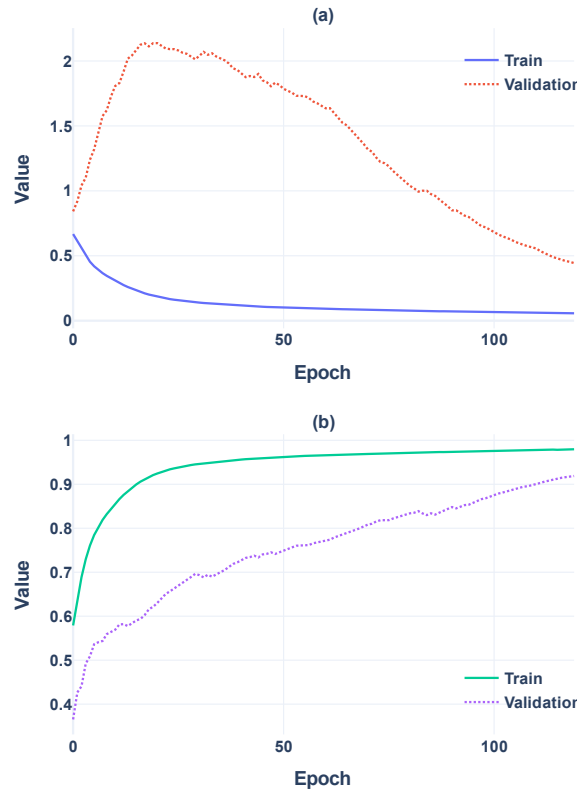


Fig. A6. FF Transformer loss (a) and accuracy (b) curves for the training and validation datasets

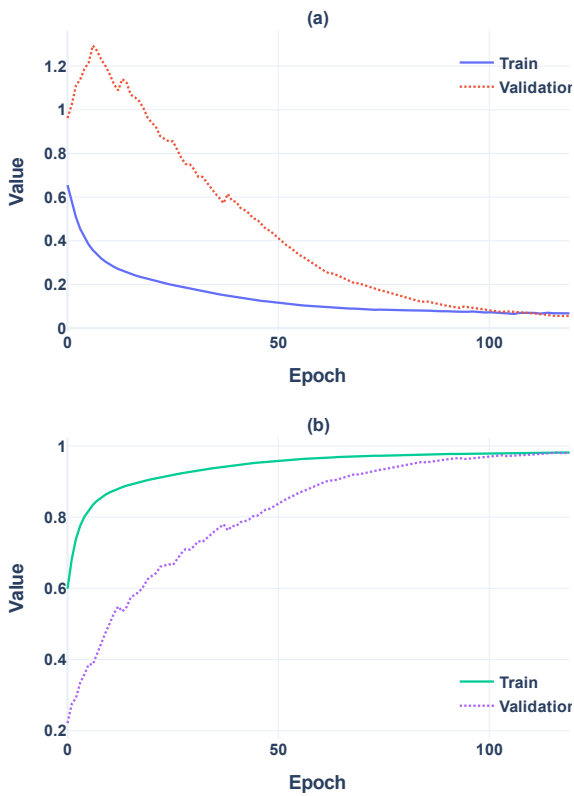


Fig. A5. CNN loss (a) and accuracy (b) curves for the training and validation datasets

# A Deep-Unfolding Approach to RIS Phase Shift Optimization Via Transformer-Based Channel Prediction

Ishan Koralege , Arthur S. de Sena, Nurul H. Mahmood,  
Farjam Karim, Dimuthu Lesthuruge, Samitha Gunarathne

*Centre for Wireless Communications, University of Oulu,  
Oulu, Finland*

{ishan.koralege, arthur.sena, nurulhuda.mahmood, farjam.karim,  
lesthuruge.silva, samitha.gunarathne}@oulu.fi

---

**Abstract:** Reconfigurable intelligent surfaces (RISs) have emerged as a promising solution that can provide dynamic control over the propagation of electromagnetic waves. The RIS technology is envisioned as a key enabler of sixth-generation networks by offering the ability to adaptively manipulate signal propagation through the smart configuration of its phase shift coefficients, thereby optimizing signal strength, coverage, and capacity. However, the realization of this technology's full potential hinges on the accurate acquisition of channel state information (CSI). In this paper, we propose an efficient CSI prediction framework for a RIS-assisted communication system based on the machine learning (ML) transformer architecture. Architectural modifications are introduced to the vanilla transformer for multivariate time series forecasting to achieve high prediction accuracy. The predicted channel coefficients are then used to optimize the RIS phase shifts. Simulation results present a comprehensive analysis of key performance metrics, including data rate and outage probability. Our results confirm the effectiveness of the proposed ML approach and demonstrate its superiority over other baseline ML-based CSI prediction schemes such as conventional deep neural networks and long short-term memory architectures, albeit at the cost of slightly increased complexity.

*Keywords:* Channel prediction, deep neural network, machine learning, reconfigurable intelligent surface, transformer.

---

## 1. INTRODUCTION

Sixth-generation (6G) wireless networks will have the ability to push beyond the limits of today's wireless systems with their groundbreaking expected features such as very low latency, ultra-high reliable connectivity, enhanced data security, and integrated intelligence that leverages machine learning (ML) capabilities (Chowdhury et al., 2020). Within this framework, reconfigurable intelligent surface (RIS) has emerged as a potential game-changer. RIS introduces a new layer of intelligence into the physical layer, allowing the stochastic wireless propagation environment to be somewhat controlled (Mahmood et al., 2023). This technology holds the promise of substantial improvements in signal strength, signal coverage, and network capacity by offering advanced control over the properties of electromagnetic waves (Basar et al., 2019). RISs are composed of passive reflecting elements that can be dynamically configured to manipulate the way wireless signals propagate toward receivers, enabling the achievement of diverse goals.

---

\* The research leading to this paper was supported by the Research Council of Finland (former Academy of Finland) 6G Flagship program (Grant Number: 346208), and Business Finland's 6GBridge program through the projects Local 6G (Grant Number 8002/31/2022) and 6CORE (Grant Number 8410/31/2022).

Since RISs typically comprise only nearly passive components with no data processing capabilities, acquiring RIS-associated channel state information (CSI) is a fundamental challenge (Yuan et al., 2021). Classical methods of CSI estimation such as least squares and minimum mean-squared error, as in the work in (Ardah et al., 2021), rely on pilot signals, which incurs significant signaling overhead and channel acquisition delay, especially when the number of RIS elements is large. It is further exacerbated under dynamic wireless environments when the channel coherence time is short, thus requiring frequent CSI acquisition. These issues can be mitigated via learning-based approaches as they harness the power of ML to learn the optimal RIS phase shifts, thereby eliminating the requirement of complex mathematical modeling or overwhelming pilot training (Hashemi et al., 2023). However, such an approach suffers from slow convergence, large training overhead, and poor generalizability. A deep unfolding approach, where ML-based techniques are used to learn partial system blocks while adhering to conventional optimization approaches for the overall system design (Balatsoukas-Stimming and Studer, 2019) can be adopted to address these limitations. This results in improved performance at a much lower complexity. More specifically, we propose to replace the pilot-based CSI estimation procedure with an



ML-based CSI prediction method, which is then used to mathematically optimize the RIS phase shifts.

ML-assisted solutions have the potential to adapt and learn the dynamics of the CSI autonomously. By training ML models with large data sets, unforeseen channel characteristics can be captured, even in complex communications environments assisted by RISs. Conventional deep neural network (DNN) and recurrent neural network (RNN) architectures can provide, to some extent, satisfactory prediction results (Gao et al., 2021). However, both DNNs and RNNs have limitations such as the vanishing gradient phenomenon. This occurs when training DNNs with a large number of layers, including activation functions. In such cases, the gradients used to update the network become extremely small or even vanish as they are backpropagated. As a result, the convergence of the algorithm is substantially slowed. (Glorot and Bengio, 2010). Long short-term memory (LSTM) networks are an evolution of RNNs that have been proposed to prevent the vanishing gradient problem (Hochreiter and Schmidhuber, 1997), which enables the processing of longer data sequences. A recent architecture named *transformers* pushes the boundaries further. Transformers have excelled across various domains due to their superior characteristic performance (Tay et al., 2022), significantly outperforming most of the previous deep learning approaches.

Diverse ML-based strategies have been employed in the literature to predict RIS-associated CSI. In (Elbir and Coleri, 2022), a federated learning strategy with distributed convolution neural networks (CNNs) was used for channel prediction in a RIS-assisted multi-antenna system employing orthogonal frequency-division multiplexing. A distributed CNN framework for downlink channel prediction in a narrow-band multi-user system was proposed in (Dai and Wei, 2022). A real-time reinforcement learning-aided CSI measurement scheme for a RIS-assisted millimeter wave (mmWave) system was considered in (Zhang et al., 2022). The authors in (Zhang et al., 2021) developed schemes based on CNNs and fully connected DNNs for RIS channel extrapolation and beam searching in a single-antenna system. A deep denoising CNN was used for aiding RIS channel estimation in a mmWave system in (Liu et al., 2020). In (Nguyen et al., 2023), a strategy based on CNNs and LSTMs was proposed to predict the CSI of a RIS-assisted system employing non-orthogonal multiple access, and in (Xia et al., 2024), a transformer-aided scheme was proposed for predicting the CSI of an uplink RIS-assisted mmWave system.

To the best of our knowledge, only the work in (Xia et al., 2024) has exploited the transformer architecture for channel prediction in RIS-assisted networks. In this paper, we propose a novel transformer-based approach to predict the unknown CSI in a RIS-assisted communication system. The predicted CSI is then used to optimize the RIS phase shifts for downlink data transmission. We compare our proposed prediction scheme with other state-of-the-art learning-based approaches, evaluating the performance using metrics such as data rate and outage probability. Based on the ML architectures we developed, it is shown that the transformer outperforms LSTM and DNN architectures, though the architecture is slightly more complex. Furthermore, the DNN architecture shows the lowest perfor-

mance with the least complexity, while the LSTM method demonstrates higher performance and greater complexity than the DNN approach, yet lower performance and less complexity compared to the transformer architecture.

The remainder of the article is organized as follows. In Section 2, the system model is provided. Section 3 gives insights into the proposed ML-based CSI predictor. Section 4 provides the results and discussion. Finally, the conclusions are presented in Section 5.

## 2. SYSTEM MODEL

Consider a downlink RIS-aided single-input single-output (SISO) communication system as in Fig. 1, where the base station (BS) and the user equipment (UE) are each equipped with a single antenna, and the RIS comprises  $N$  reflecting elements. Let  $f \in \mathbb{C}$ ,  $\mathbf{h} \in \mathbb{C}^{N \times 1}$ ,  $\mathbf{g} \in \mathbb{C}^{N \times 1}$  be the fading channels between the BS and the UE, the BS and the RIS, and the RIS and the UE, respectively. The channel vectors  $\mathbf{h} = [h_1, \dots, h_i, \dots, h_N]^T$  and  $\mathbf{g} = [g_1, \dots, g_i, \dots, g_N]^T$ , where  $h_i$  ( $g_i$ ) is the channel between the BS (UE) and  $i^{\text{th}}$  RIS element. We assume that the channel coefficients are Rayleigh distributed, i.e.,  $f \sim \mathcal{CN}(0, L_{\text{BS-UE}})$ ,  $h_i \sim \mathcal{CN}(0, L_{\text{BS-RIS}})$ , and  $g_i \sim \mathcal{CN}(0, L_{\text{RIS-UE}})$ , for  $i = 1 \dots, N$ , where  $L_a$  denotes the path loss coefficients in linear scale for the  $a^{\text{th}}$  link with  $a \in \{\text{BS-UE}, \text{BS-RIS}, \text{RIS-UE}\}$ . In this work, we consider the log-distance path loss model, which can be given in dB by

$$L_{a,\text{dB}} = L_{0,\text{dB}} + 10\eta_a \log_{10} \left( \frac{d_a}{d_0} \right), \quad (1)$$

where  $\eta_a$  and  $d_a$  represent the path loss exponent and the distance for the link  $a$  respectively, and  $L_{0,\text{dB}}$  is the reference path loss in dB at a reference distance  $d_0$ . Under this model, the received signal at the UE can be expressed as

$$y = \sqrt{P} (f + \mathbf{g}^T \Phi \mathbf{h}) x + n, \quad (2)$$

where  $n \sim \mathcal{CN}(0, N_0)$  is the additive white Gaussian noise (AWGN),  $x$  is the transmitted symbol,  $P$  represents the transmit power and  $\Phi$  is the matrix comprising the reflection coefficients of the RIS, which is a diagonal matrix defined as (Basar et al., 2019)

$$\Phi = \text{diag}(\phi), \quad (3)$$

where  $\phi = [a_1 e^{j\phi_1}, a_2 e^{j\phi_2}, \dots, a_N e^{j\phi_N}]^T$  is the RIS phase shift vector, in which  $a_i$  is the amplitude and,  $\phi_i$  is the adjustable phase induced by the  $i^{\text{th}}$  reflecting element. Due to the passive nature of the RIS, we can assume  $a_i = 1, \forall i = 1, \dots, N$ .

### 2.1 RIS Phase Shift Optimization

We wish to compute a phase shift matrix  $\Phi$  that can maximize the signal power propagating through the RIS at the UE during downlink data transmission. To this end, we assume that the direct link is non-dominant. This can be the case when it is non-line of sight, while the BS-RIS and RIS-UE links are in line of sight. Under this assumption, we can recall the Cauchy-Schwartz inequality and achieved (Björnson et al., 2024)

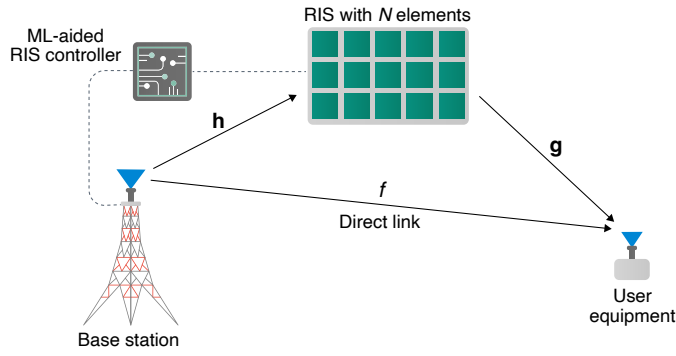


Fig. 1. System model. An RIS comprising  $N$  elements assists the communication between one single-antenna BS and one single-antenna UE.

$$|\mathbf{g}^T \Phi \mathbf{h}|^2 = |(\mathbf{h} \odot \mathbf{g})^T \boldsymbol{\phi}|^2 \leq \left( \sum_{i=1}^N |h_i g_i e^{j\phi_i}| \right)^2, \quad (4)$$

where  $\odot$  denotes the Hadamard product. The inequality in (4) provides an upper bound to the reflected power i.e., the maximum power that can be reflected by the RIS. The equality can be achieved if and only if  $\phi_i = -\arg(h_i g_i)$ . As a result, the desired optimal RIS phase shift vector can be given by

$$\boldsymbol{\phi}^* = e^{-j\arg(\mathbf{h} \odot \mathbf{g})}. \quad (5)$$

The optimum phase shift coefficients in (5) can be readily computed after predicting the CSI through ML, as discussed in the subsequent sections.

## 2.2 Performance Metrics

In this subsection, we present the performance metrics used to assess the effectiveness of the proposed ML-based CSI prediction framework. To this end, we first provide an expression for the signal-to-noise ratio (SNR). More specifically, by using (2) and (4), the instantaneous SNR observed at the UE, denoted as  $\gamma$ , can be computed by

$$\gamma = \frac{|(\mathbf{h} \odot \mathbf{g})^T \boldsymbol{\phi} + f|^2 P}{N_0}. \quad (6)$$

The outage probability,  $P_{\text{out}}$ , is the first important performance metric used to assess the performance of the UE. It is the probability of the instantaneous SNR falling below a defined threshold  $\gamma_{\text{th}}$ , i.e.,

$$P_{\text{out}} = P(\gamma < \gamma_{\text{th}}), \quad (7)$$

where  $P$  denotes probability.

Last, we consider the maximum data rate that can be achieved when the BS is transmitting at a fixed rate  $R_{\text{th}}$ , subject to a given outage probability  $P_{\text{out}}$ , which can be obtained as

$$R = (1 - P_{\text{out}})R_{\text{th}}, \quad (8)$$

where  $R_{\text{th}}$  represents the data rate threshold given by

$$R_{\text{th}} = \log_2(1 + \gamma_{\text{th}}). \quad (9)$$

## 3. CSI PREDICTION AND DATA PREPARATION

This section presents the proposed architecture of our transformer model which is used to make multivariate time

series predictions of the channel coefficients. Details on the generation of the data sets used to train, validate, and test the implemented ML models are also presented.

### 3.1 Transformer-Based CSI Prediction

In this subsection, we describe our proposed CSI prediction approach based on the ML transformer architecture, which we later exploit to optimize the RIS phase shifts. Transformers are a deep learning architecture initially introduced in (Vaswani et al., 2017) by a research group in Google. The key feature of this architecture is its state-of-the-art attention mechanism. The attention mechanism is a computational method that focuses on specific parts of the input data sequence, assigning varying degrees of importance to the different parts. This mechanism helps to inform the model where to pay attention when processing data. The vanilla transformer architecture consists of two parts, an encoder and a decoder, which consists of a sequence-to-sequence architecture. In this method, the model generates an output data sequence according to the given input sequence.

In this work, we aim to predict the CSI at the next time instance based on a given sequence of previous CSI samples. Therefore, the original sequence-to-sequence architecture is modified into a sequence-to-one architecture using only one encoder. More specifically, the encoder of our transformer model comprises an input embedding module, a positional encoding module, and a transformer encoder module, as shown in Fig. 2. The input embedding module transforms the dimension of the input data to the model dimension of the subsequent inner layer of the transformer. Since the transformer lacks a recurrent structure as in recurrent neural networks, it feeds the positional information to the output of the embedding layer separately. After that, using the acquired knowledge from the input sequence, an abstract representation is generated by the transformer encoder. The encoder consists of a multi-head attention block, layer normalization blocks, and a feedforward layer, as illustrated in Fig. 2. Then, the encoder output is sent through a fully connected (FC) linear layer and an activation function. Figure 2 shows the modified sequence-to-one transformer architecture used for CSI prediction, where we feed samples of real and imaginary channel data separately to forecast the next time instance of the channel coefficients as output.

The vanilla transformer architecture, originally introduced for natural language processing applications, contains two FC layers in the feedforward block and uses the rectified linear unit (ReLU) as the activation function (Vaswani et al., 2017). This setup is effective for capturing long-term variations in the multi-head attention output. However, given the relatively small window size, our problem requires observing and capturing short-term variations. Therefore, as shown in Fig. 3, we modified the original architecture by implementing two one-dimensional (1D) CNN layers instead of the two FC layers, as CNNs effectively capture local temporal patterns crucial for accurate forecasting. Additionally, we replaced the ReLU activation function with a hyperbolic tangent (Tanh) function. These modifications result in better performance in our time series forecasting problem.

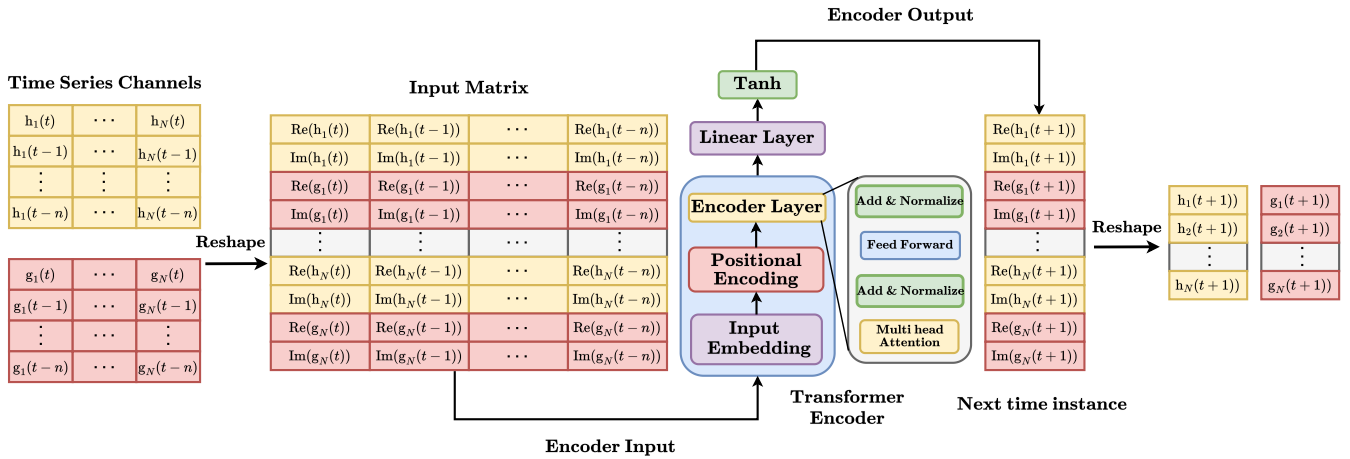


Fig. 2. Sequence-to-one transformer architecture with input and output.

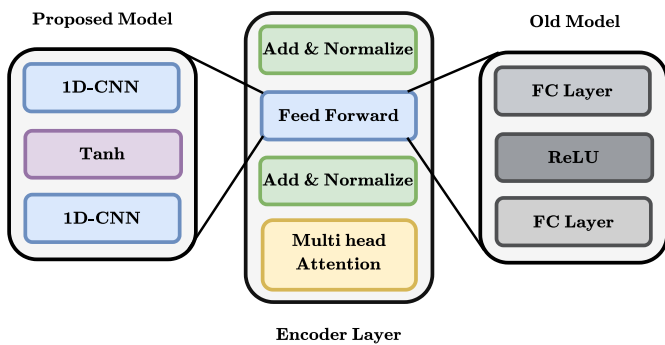


Fig. 3. Feedforward layer in the vanilla transformer architecture (right) and the proposed model (left).

3.2 Data Preparation

Most of the cited works on CSI prediction assume independent/uncorrelated fading, whereas real-world scenarios often exhibit correlation. We incorporate this important property in our work by considering time-correlated Rayleigh-distributed fading channel coefficients. We first generate uncorrelated Rayleigh-fading samples through simulation with zero mean and unit variance, which are then convoluted with a finite impulse response filter representing the correlation function. Before feeding time-correlated channel data to the ML models, the data is normalized for efficient convergence. Out of the 2550 time-correlated data samples, 80% is allocated for training, while the remaining 20% are allocated for validation. Another 50 samples are taken for testing per Monte Carlo iteration. In all the architectures, a moving window is deployed for iterative training, validation, and testing processes. To learn the temporal patterns, consecutive 10 samples are taken. Using this learned knowledge, 11<sup>th</sup> sample is predicted. A perfect CSI assumption is made for the training data. Though this is a strong assumption, it helps us to obtain useful information about the performance upper bound.

4. RESULTS AND DISCUSSIONS

This section presents comprehensive simulation results to verify the effectiveness of our proposed transformer-

based CSI prediction framework. The number of parameters and other architectural details used by different ML strategies are also put into perspective. To demonstrate the performance advantages of our approach, we consider two baseline ML architectures, namely DNN and LSTM, where metrics such as data rate and outage probability are compared.

We consider the path loss exponent for the BS-RIS and the RIS-UE links to be 2.2, whereas the exponent for the direct BS-UE link is assumed to be 4.2. The reference path loss value and noise power are taken as -30 dB and -100 dB, respectively. The distances from the transmitter to the RIS, the RIS to the receiver, and the transmitter to the receiver are set to 38 m, 5 m, and 40 m. Here, the SNR threshold  $\gamma_{th} = 1$ , correspondingly  $R_{th} = 1$  too. Unless stated otherwise, an RIS with eight elements is considered and the transmit power is set to 0dB.

4.1 Comparison with ML Baseline Architectures

In this section, the performance of the proposed transformer approach is compared with the conventional DNN and the LSTM baseline ML architectures. All ML models are trained for 100 epochs employed with the Adam optimizer alongside the root mean square error (RMSE) as the loss function. ReLU is used as the activation function for DNN and LSTM architectures while Tanh is used for the transformers. Architectures in all the models used a window size of 10. In each encoder layer, four attention heads and 20 feedforward layer dimensions are used across all transformer models.

Variation of Transmit Power with Fixed RIS Elements

Let us consider the system model as mentioned in Fig. 1 with the number of RIS elements set to  $N = 8$ . There are 16 channels associated with the RIS model, eight each for the BS-RIS and RIS-UE link, respectively. Since we are predicting real and imaginary channel values separately, 32 features are required to predict at once. Since the number of RIS elements remains constant in this scenario, we can employ a single model for each architecture to make predictions, as the input to each model remains unchanged.

Table 1. Optimized hyperparameters of DNN, LSTM, and Transformer models.

Hyperparameters	Values														
	DNN					LSTM					Transformer				
Number of RIS elements	4	8	12	16	20	4	8	12	16	20	4	8	12	16	20
Input and Output features	16	32	48	64	80	16	32	48	64	80	16	32	48	64	80
Neurons in LSTM layer	-	-	-	-	-	22	44	60	76	100	-	-	-	-	-
Neurons in FC layer 1	20	36	52	68	88	18	36	56	68	90	16	32	48	64	80
Neurons in FC layer 2	24	40	56	72	92	-	-	-	-	-	-	-	-	-	-
Neurons in FC layer 3	24	44	56	68	94	-	-	-	-	-	-	-	-	-	-
Neurons in FC layer 4	20	36	52	68	88	-	-	-	-	-	-	-	-	-	-
Transformer model dimensions	-	-	-	-	-	-	-	-	-	-	24	48	60	80	120

Table 2. Summary of metrics associated with DNN, LSTM, and Transformer models.

RIS Elements	Architecture	Train RMSE	Test RMSE	$P$
4	DNN	0.0335	0.0271	2, 280
	LSTM	0.0149	0.0172	3, 796
	Transformer	0.0134	0.0158	6, 838
8	DNN	0.0364	0.023	6, 968
	LSTM	0.0282	0.0191	16, 532
	Transformer	0.0131	0.0142	26, 702
12	DNN	0.0413	0.0259	14, 216
	LSTM	0.0329	0.0219	32, 552
	Transformer	0.0159	0.0157	46, 818
16	DNN	0.0535	0.0317	24, 024
	LSTM	0.0425	0.0266	52, 820
	Transformer	0.0177	0.0165	82, 894
20	DNN	0.0555	0.0322	39, 538
	LSTM	0.036	0.0228	89, 170
	Transformer	0.0182	0.0166	164, 630

The transmit power is varied from 0 dBm to 50 dBm by keeping other initial model parameters of the RIS the same. Optimized hyperparameters of the ML models only used for this scenario are tabulated under the columns where the number of RIS elements is eight for each approach as in Table 1. To measure the effectiveness of the architectures, a performance evaluation should be carried out. In this study, we use RMSE as the prediction evaluation criterion. It measures the RMSE between the predicted feature sequence and the actual sequence values in the test and training data sets separately. Apart from that, we have obtained the number of model parameters  $P$  associated with models which are calculated using weights and biases. Metrics obtained through the above model simulations are organized in Table 2. Moreover, when  $N$  is fixed at 8 it is evident that the proposed transformer-based approach significantly outperforms LSTM and DNN architectures, thereby reducing the training RMSE approximately by 115% and 177% and test RMSE approximately by 34% and 62%, respectively.

#### Variation of RIS Elements with Fixed Transmit Power

In this scenario, number of RIS elements is the variable. Let us consider instances where the number of RIS elements is 4, 8, 12, 16, and 20 by keeping other initial model parameters of the RIS the same. Therefore, the number of channels associated with the RIS are 8, 16, 24, 32, and 40, respectively. Then, according to our model configuration (considering real and imaginary values separately as fea-

tures), the number of features we need to handle would be 16, 32, 48, 64, and 80, respectively. For each case, we need to optimize the ML models separately since the number of input features is changing. The optimized hyperparameters for each architecture are summarized in Table 1. It shows the number of neurons available in each layer of the respective architecture, the number of neurons in the LSTM layer, and the transformer model dimensions.

Table 2 presents the summary of metrics associated with all three ML architectures such as the train RMSE, the test RMSE, and the number of model parameters. According to the observations, it is clear that the transformer architecture has outperformed the DNN and LSTM architectures in terms of performance when compared with both the train and test RMSE values. This shows that the transformer architecture can significantly outperform the state-of-the-art ML architectures, albeit at the cost of higher complexity. Hence, the transformer architecture is preferable in scenarios where the prediction accuracy is to be prioritized over computational complexity. However, the increased computational complexity can be easily handled by using optimized hardware such as tensor processing units.

#### 4.2 RIS Phase Optimization with Predicted CSI

Using the prediction values obtained from the ML models, we can calculate the optimal phase shift vector of the RIS model as mentioned in (5). Then plugging the test data (actual) channel values and the calculated RIS optimal phase shift vector to (6), we can obtain the maximized SNR values. From there onwards, (7) and (8) can be obtained with respect to the maximized SNR values. Therefore, we have shown that RIS could be optimized by predicting the CSI in a RIS-assisted communication system. Let us focus on the two major scenarios of our research output. The generated results shown in Fig. 4 and Fig. 5 compare key scenarios, including the outputs with optimized-phase RIS, fixed-phase RIS, without the RIS, and the integrated outputs predicted by the transformer, LSTM, and DNN models. Additionally, in the fixed-phase RIS scenario,  $\Phi$  is set as an identity matrix.

The system performance for a fixed number of RIS elements with varying transmit power is illustrated in Fig. 4. Specifically, Figure 4(a) demonstrates the variation in outage probability, while Figure 4(b) shows the variation in average rate. As the transmit power increases, the outage probability decreases, and the average rate increases. In Fig. 4(a), the lowest outage probability is observed when the RIS is optimized. It is the optimum scenario when

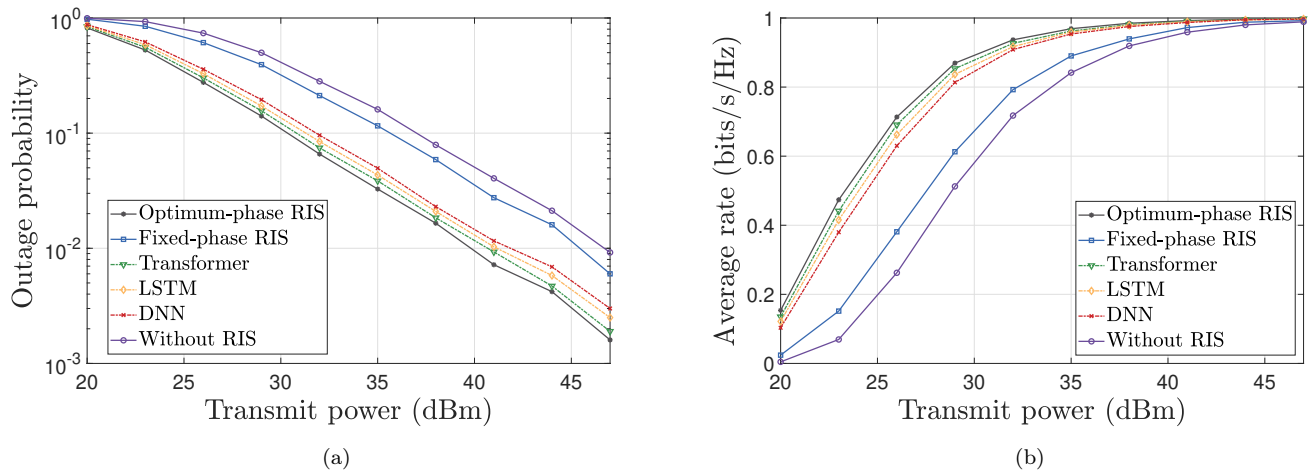


Fig. 4. System performance for a fixed number of RIS reflecting elements ( $N = 8$ ) in terms of (a) outage probability, and (b) average rate, when the transmit power is varied.

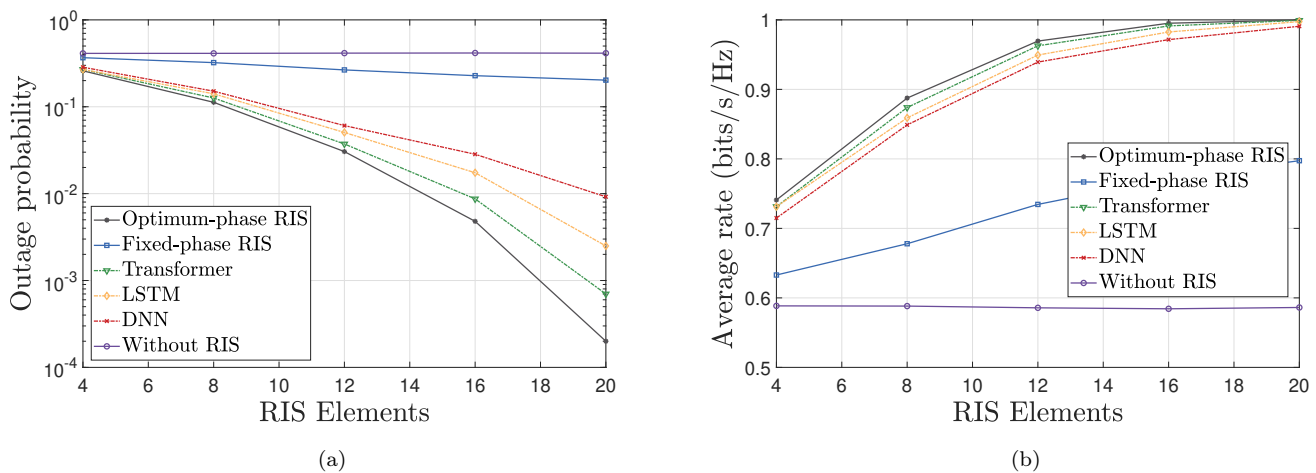


Fig. 5. System performance for fixed transmit power (30 dBm) in terms of (a) outage probability, and (b) average rate when the number of RIS reflecting elements is varied.

the CSI is accurately known at the transmitter. This is the performance upper/lower bound and is not possible in practice. The transformer model closely approaches the performance of the optimized RIS, outperforming both the LSTM and DNN models. Furthermore, to achieve an outage probability of 0.01, the optimized RIS scenario requires a transmit power of 39.79 dBm, compared to 46.66 dBm for the scenario without RIS. In addition, to achieve the same outage probability, the required transmit powers for the transformer, LSTM, and DNN models are 40.67 dBm, 41.13 dBm, and 41.85 dBm, respectively. In Fig. 4(b), the average rate is maximized when the RIS is optimized and the transformer model gets very close to the upper bound established by the optimized RIS. Moreover, it can be observed that when the transmit power is 25 dBm, the performance gap between the scenarios where the transformer model and the system without RIS is approximately 0.4 bits/s/Hz. The LSTM and DNN models also demonstrate notable performance improvements, although they fall short of the transformer model's performance.

The system performance for a fixed transmit power, while varying the number of RIS elements, is presented in Fig.

5. Figure 5(a) illustrates the variation in outage probability, whereas Figure 5(b) depicts the variation in average rate. Generally, as the number of RIS elements increases, the outage probability decreases, and the average rate increases for scenarios incorporating RIS. Figure 5(a) shows that the outage probability is minimized when the RIS is optimized. The transformer architecture yields the closest results to the optimized RIS, followed by the LSTM and DNN architectures. As per the observations, for a 12-element RIS system, the optimized RIS scenario achieves an outage probability of 0.03, while the transformer, LSTM, and DNN scenarios result in outage probabilities of 0.037, 0.05, and 0.06, respectively. In Fig. 5(b), the average rate is maximized when the RIS is optimized and the transformer architecture again demonstrates the closest prediction to the average rate achieved with the optimized RIS. When the number of RIS elements reaches eight, the performance gap between the scenario with the transformer model and the scenario without RIS is approximately 0.3 bits/s/Hz.

From the results, it is evident that the application of RIS has increased the overall system performance noticeably



and the best results are given when the phase is optimized. In every graph, prediction curves lie between the optimum phase and the system without RIS. Out of the three prediction curves, the transformer prediction curve goes very closely with the optimum phase RIS output showing that the transformers have provided the most convincing results compared to both the DNN and LSTM architectures.

## 5. CONCLUSIONS

Harnessing the full potential of RIS technology hinges on accurate CSI estimation/prediction since the optimum RIS phase shifts are a function of the corresponding composite channel's CSI. In this study, we proposed a novel sequence-to-one transformer architecture to predict the RIS-associated CSI, enabling the efficient and accurate optimization of the RIS phase shifts. For the CSI prediction task, three architectures namely DNN, LSTM, and transformers were utilized. For the time series channel sample data set created, the transformer architecture provided the lowest RMSE value outperforming DNN and LSTM methods in the scenarios discussed above. After that, the predicted CSI from ML models was fed into the RIS model for phase optimization. According to the results obtained, the transformer was the better multivariate time series prediction architecture out of the three architectures in terms of performance but at the cost of higher complexity. Optimizing the RIS phase shift based on transformer-predicted CSI was found to perform very close to the optimum case when the CSI was assumed to be accurately known. Conversely, the DNN architecture yielded the lowest performance and was also the least complex architecture. The LSTM architecture was positioned in between the other two architectures, offering a middle ground in terms of both performance and complexity. Finally, this study concluded that the proposed sequence-to-one transformer architecture provided promising results for channel prediction regarding RIS phase optimization.

## REFERENCES

- Ardah, K., Gherekhloo, S., de Almeida, A.L., and Haardt, M. (2021). TRICE: A channel estimation framework for RIS-aided millimeter-wave MIMO systems. *IEEE Signal Process. Lett.*, 28, 513–517.
- Balatsoukas-Stimming, A. and Studer, C. (2019). Deep unfolding for communications systems: A survey and some new directions. In *2019 IEEE International Workshop on Signal Processing Systems (SiPS)*, 266–271. doi:10.1109/SiPS47522.2019.9020494.
- Basar, E., Di Renzo, M., De Rosny, J., Debbah, M., Alouini, M.S., and Zhang, R. (2019). Wireless communications through reconfigurable intelligent surfaces. *IEEE access*, 7, 116753–116773.
- Björnson, E., Demir, Ö.T., et al. (2024). *Introduction to Multiple Antenna Communications and Reconfigurable Surfaces*. Now Publishers, Inc.
- Chowdhury, M.Z., Shahjalal, M., Ahmed, S., and Jang, Y.M. (2020). 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. *IEEE open j. Commun. Soc.*, 1, 957–975. doi:10.1109/OJCOMS.2020.3010270.
- Dai, L. and Wei, X. (2022). Distributed machine learning based downlink channel estimation for RIS assisted wireless communications. *IEEE Trans. Commun.*, 70(7), 4900–4909. doi:10.1109/TCOMM.2022.3175175.
- Elbir, A.M. and Coleri, S. (2022). Federated learning for channel estimation in conventional and RIS-assisted massive MIMO. *IEEE Trans. Wireless Commun.*, 21(6), 4255–4268. doi:10.1109/TWC.2021.3128392.
- Gao, S., Dong, P., Pan, Z., and Li, G.Y. (2021). Deep multi-stage CSI acquisition for reconfigurable intelligent surface aided MIMO systems. *IEEE Commun. Lett.*, 25(6), 2024–2028. doi:10.1109/LCOMM.2021.3063464.
- Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 249–256. JMLR Workshop and Conference Proceedings.
- Hashemi, R., Ali, S., Mahmood, N.H., and Latva-Aho, M. (2023). Deep Reinforcement Learning for Practical Phase-Shift Optimization in RIS-Aided MISO URLLC Systems. *IEEE IoT J.*, 10(10), 8931–8943. doi:10.1109/JIOT.2022.3232962.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Liu, S., Gao, Z., Zhang, J., Renzo, M.D., and Alouini, M.S. (2020). Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces. *IEEE Trans. Veh. Technol.*, 69(8), 9223–9228. doi:10.1109/TVT.2020.3005402.
- Mahmood, N.H., Berardinelli, G., Khatib, E.J., Hashemi, R., De Lima, C., and Latva-aho, M. (2023). A functional architecture for 6G special-purpose industrial IoT networks. *IEEE Trans. Ind. Inf.*, 19(3), 2530–2540. doi:10.1109/TII.2022.3182988.
- Nguyen, C., Hoang, T.M., and Cheema, A.A. (2023). Channel estimation using CNN-LSTM in RIS-NOMA assisted 6G network. *IEEE trans. mach. learn. commun. netw.*, 1, 43–60. doi:10.1109/TMLCN.2023.3278232.
- Tay, Y., Dehghani, M., Bahri, D., and Metzler, D. (2022). Efficient transformers: A survey. *ACM Comput. Surv.*, 55(6), 1–28.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Xia, G., Liu, H., and Long, K. (2024). Transformer-empowered parallel channel prediction for fast-paced and dynamic RIS-aided wireless communication systems. *IEEE Commun. Lett.*, 1–1. doi:10.1109/LCOMM.2024.3383427.
- Yuan, X., Zhang, Y.J.A., Shi, Y., Yan, W., and Liu, H. (2021). Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities. *IEEE Wireless Commun.*, 28(2), 136–143.
- Zhang, Q., Saad, W., and Bennis, M. (2022). Millimeter wave communications with an intelligent reflector: Performance optimization and distributional reinforcement learning. *IEEE Trans. Wireless Commun.*, 21(3), 1836–1850. doi:10.1109/TWC.2021.3107520.
- Zhang, S., Zhang, S., Gao, F., Ma, J., and Dobre, O.A. (2021). Deep learning optimized sparse antenna activation for reconfigurable intelligent surface assisted communication. *IEEE Trans. Commun.*, 69(10), 6691–6705. doi:10.1109/TCOMM.2021.3097726.

# Data Center Resource Usage Forecasting with Convolutional Recurrent Neural Networks

Miika Malin and Jaakko Suutala

*University of Oulu / ITEE / BISG, 9BISG, P.O. Box 4500, 90014  
Oulu Finland (e-mail: {firstname.lastname}@oulu.fi)*

---

**Abstract:** Energy efficiency, scalability, and reliability are increasingly important for sustainable data centers. In this paper, we focus on forecasting real-world resource usage using neural network time series models, specifically utilizing convolutional recurrent long short-term Memory (LSTM) and gated recurrent unit (GRU) architectures. In our analysis, we compare LSTM and GRU in terms of forecasting accuracy and computational complexity during model training. We demonstrate that recurrent neural networks are more accurate and robust compared to the traditional autoregressive integrated moving average (ARIMA) time series model in this complex forecasting problem. GRU achieved a 9% reduction and LSTM a 5% reduction in forecasting mean squared error (MSE) compared to ARIMA. Furthermore, the GRU architecture with a 1D convolution layer outperforms LSTM architecture in both forecast accuracy and training time. The proposed model can be effectively applied to load forecasting as part of a data center computing cluster. In this application, the proposed GRU architecture has 25% fewer trainable parameters in the recurrent layer than the commonly used LSTM.

*Keywords:* recurrent neural network, convolutional neural network, data center load forecasting, energy efficiency, sustainability, control optimization, monitoring

---

## 1. INTRODUCTION

Energy consumption reduction and resource optimization are becoming important in computational intensive data center environments aiming towards more sustainable and green systems (Bourhane et al., 2020). World wide energy consumption of data centers has been estimated rose to 205 TWh in 2018 from 153 TWh in 2005. This means ~1% of global total electricity usage (Masanet et al., 2020; Jones, 2018). The whole information and communications ecosystem causes more than 2% of emissions. This is on same level with aviation fuel emissions, and it is predicted to be even higher in future (Jones, 2018).

Optimizing energy consumption of hardware in data centers is critical, as servers and other IT equipment can typically take more than 40% of total energy-usage in data center (Shehabi et al., 2016). To be able to optimize the IT systems in proactive manner, intelligent and efficient resource usage forecast is one subject to be solved. Several different neural network architectures have been proposed for the task, and many of the solutions use recurrent neural network (RNN) based approaches (e.g in (Zhang et al., 2016; Janardhanan and Barrett, 2017; Ouhamme et al., 2021; Yuan et al., 2024)) since it has been designed to be used with sequential problems such as time series forecasting. Also more traditional models such as autoregressive integrated moving average (ARIMA). ARIMA is a common tool used by statisticians in time series forecasting (Hewamalage, 2020), and has been used to tackle the problem, e.g. in (Kumar and Singh, 2020; Calheiros et al., 2015). In this work, we are going to compare two

recurrent neural network architectures: long short-term memory (LSTM) and gated recurrent unit (GRU) together with traditional ARIMA model.

While several architectures leveraging recurrent neural networks have emerged to address load forecasting, they often overlook the critical attribute of efficiency. Notably, many of these architectures do not utilize convolutional or more efficient GRU recurrent layers. In a notable advancement, the architecture proposed by Ouhamme et al. (2021) focused on optimizing the convolutional layer of the LSTM model for efficiency. Building upon this progress, our work introduces a novel approach by integrating both GRU and convolutional layers. This synergistic combination not only enhances forecast accuracy but also improves model efficiency, a critical attribute essential for real-world integration in data center scenarios.

Efficient and dependable resource usage forecasts are essential for managing dynamic, scalable clusters, enabling the adjustment of the number of powered-on physical machines as needed. Such forecasts hold the potential to significantly reduce energy consumption in data centers (Bayati, 2018). Additionally, resource usage forecasts find application in load balancing, particularly in virtual machine (VM)-based data centers where the allocation of physical machines for VMs can be modified. This capability facilitates the optimization of resource utilization and enhances the overall efficiency of data center operations (Shaw et al., 2017).

## 2. METHODS FOR TIME SERIES FORECASTING

In this section, we explore the time series forecasting methodologies employed in this study, with a focus on predicting data center resource usage efficiently. By leveraging advanced predictive techniques, our aim is to provide a robust and efficient framework for analyzing temporal data patterns and enhancing the accuracy of our forecasts. The forecasts can be used to optimize resource allocation and improving operational adaptability in data centers.

### 2.1 Autoregressive Integrated Moving Average model

The autoregressive integrated moving average model is a widely-used forecasting method that integrates autoregression, moving average, and differencing. In this study, we employ the ARIMA model as a baseline for forecasting due to its historical prevalence in data center resource usage prediction tasks and its ability to effectively model various types of time series data.

ARIMA model can be defined as

$$y'_t = \alpha + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t, \quad (1)$$

where  $\alpha$  is the optional intercept of the model.  $\phi_i$  is the coefficient for the autoregressive part of the model, and  $\theta_i$  is the coefficient for the moving average part of the model.  $y'_t$  is the value of the differenced time series at timestep  $t$ , and  $\epsilon_i$  is the past forecast error (Hyndman and Athanasopoulos, 2019).

ARIMA model uses hyperparameters  $p, d, q$ , and following conclusions about ARIMA model hyperparameters  $p$  and  $q$  can now be derived from the Equation 1:  $p$  can be seen as the order of the autoregressive part of the model, and  $q$  is the order of moving average part of the model. The hyperparameter  $d$  is the order of differencing in time series (Hyndman and Athanasopoulos, 2019).

### 2.2 Convolution on Time Series Data

Convolution can be seen as sliding a window over the data. Due to the success of convolutional neural networks (CNNs) in image and natural language processing, convolution layers has been applied to time series analysis as well.

In CNNs processing image data, convolution involves a two-dimensional filter sliding over the width and height of the image (Goodfellow et al., 2016). In contrast, for time series data, the only dimension to slide over is time, which requires the use of one-dimensional convolution.

Two dimensional convolution starting from point  $(i, j)$  can be defined as:

$$C(i, j) = \sum_{m=0}^h \sum_{n=0}^w I_{i+m, j+n} K_{m, n}, \quad (2)$$

where  $I$  is the input data,  $K$  is the kernel of the convolution with dimensions  $h \times w$ . Kernel contains weights  $w$  for the convolution. Here,  $h$  is the height, and  $w$  is the width of the convolution window. (Goodfellow et al., 2016)

Since the 1D convolution can only slide through one dimension,  $w$  is always same as number of features in input. Definition for 1D convolution can be derived from Equation 2:

$$C(i) = \sum_{m=0}^h \sum_{n=0}^w I_{i+m, n} K_{m, n}, \quad (3)$$

where  $i$  is the row of input where the convolution starts and  $w$  is the number of features in dataset. In time series context,  $w$  can be seen as a number of features recorded in each time point. Again, in time series context this means that kernel is slid over the time dimension.

Use of the 1D convolution reduces the computational complexity from  $\sim O(N^2 K^2)$  to  $\sim O(NK)$  when comparing to 2D convolution, when input dimensions are  $N \times N$  and kernel dimensions  $K \times K$  (Kiranyaz et al., 2021).

### 2.3 Long Short-Term Memory

Long short-term memory is a recurrent neural network, which uses LSTM units and tries to solve vanishing gradient problem of recurrent neural networks (Hochreiter and Schmidhuber, 1997). This means that the architecture can find more efficiently long term dependencies from time series.

LSTM network with forget gate and biases consists of LSTM cells, where each cell has three gates:

- Forget gate

$$f_t = S(W_f x_t + U_f h_{t-1} + b_f). \quad (4)$$

- Input Gate

$$i_t = S(W_i x_t + U_i h_{t-1} + b_i). \quad (5)$$

- Output gate

$$o_t = S(W_o x_t + U_o h_{t-1} + b_o). \quad (6)$$

In Equations 4-6  $S$  is an activation function. Often  $S$  is a sigmoid function, as proposed in the original version of LSTM (Hochreiter and Schmidhuber, 1997; Hewamalage, 2020; Dey and Salem, 2017). Now output / hidden state  $h_t$  of cell at timestep  $t$  can be defined as:

$$h_t = o_t \odot T(c_t), \quad (7)$$

where

$$\begin{aligned} c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ \tilde{c}_t &= T(W_c x_t + U_c h_{t-1} + b_c). \end{aligned} \quad (8)$$

In Equations 7 and 8  $T$  is an activation function, and often  $\tanh$  function is used as proposed in the original architecture (Hochreiter and Schmidhuber, 1997). In Equations 4 - 8  $x_t$  is the input vector for LSTM cell at timestep  $t$ .  $W_x, U_x$  and  $b_x$  are weights and biases to be tuned in the training process. The operation symbol  $\odot$  is an element-wise Hadamard product. (Hochreiter and Schmidhuber, 1997; Hewamalage, 2020; Dey and Salem, 2017)

From Equations 4-8 we get that LSTM has total of  $4(n^2 + nm + n)$  optimizable parameters. Here  $n$  is the dimension of hidden state and  $m$  is the dimension of input vector. (Dey and Salem, 2017)

#### 2.4 Gated Recurrent Unit

Gated recurrent unit has been motivated by LSTM unit, but it has a simpler design (Cho et al., 2014). Unlike LSTM units, GRUs do not have an output gate as shown in Equations 9-11. This more streamlined architecture provides efficiency in both training and forecasting tasks. Since it has fewer gates, it also has fewer weights to optimize, making the backpropagation through time faster. Additionally, using the trained model to yield forecasts is more efficient due to the reduced number of calculations required.

GRU network consists of multiple GRU cells, where each cell has two gates:

- Reset gate

$$r_t = S(W_r x_t + U_r h_{t-1} + b_r), \quad (9)$$

- Update gate

$$u_t = S(W_u x_t + U_u h_{t-1} + b_u). \quad (10)$$

By using these two gates the output / hidden state  $h_t$  at timestep  $t$  can be defined as:

$$\begin{aligned} h_t &= (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \\ \tilde{h}_t &= T(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h), \end{aligned} \quad (11)$$

where  $T$  is hyperbolic tangent and  $S$  is sigmoid activation function.

All  $W_x$  are weight matrices, and all  $b_x$  are bias vectors, which constitute the learnable parameters of the model. Again  $\odot$  is element-wise Hadamard product (Cho et al., 2014; Hewamalage, 2020; Dey and Salem, 2017).

From Equations 9-11 we get that GRU has total of  $3(n^2 + nm + n)$  optimizable parameters. Again  $n$  is the dimension of hidden state and  $m$  is the dimension of input vector as in the LSTM. (Dey and Salem, 2017)

### 3. EXPERIMENTAL DESIGN

In this section, we present a detailed overview of the data, the architecture of the model built using the methodologies outlined in Section 2, and the comprehensive training process. These elements form the foundation for the CPU usage forecasting described in Section 4.

#### 3.1 Dataset Description

The dataset contains resource usage traces of 1750 Virtual Machines from Bitbrains distributed data center (Bitbrains, 2013). The usage trace length is 1 month: from August 12, 2013 to September 11, 2013. The dataset has samples in 5 minute intervals. Bitbrains has customers from various industries, resulting diverse use cases and usage from one VM to another.

Table 1. Input variables from Bitbrains data

Name	Description	Unit
cpu_usage	Central processing unit usage	%
memory_usage	Memory usage	%
disk_read	Disk read throughput	KB/s
disk_write	Disk write throughput	KB/s
net_receive	Network received throughput	KB/s
net_transmit	Network transmitted throughput	KB/s

Dataset consists of two different sets. The first set contains data for 1250 VMs in fast storage area network (SAN) and second set contains data for 500 VMs from both SAN and slower network attached storage devices. Only the traces for VMs in fast storage area network were used in our experiments.

Many of the VMs had low or nearly static load throughout the trace period. This poses challenges for forecasting, as static CPU usage is trivial to forecast and would not accurately reflect the forecast accuracy. To address this issue, a subset of machines which had average CPU usage greater than 30% was selected. From this subset, five random machines were chosen for model training and forecasting to ensure a more representative evaluation of the methods. The five selected machines were identified by their ID numbers: 220, 242, 253, 269, and 283. All the features together with descriptions which were used in the forecasting models are defined in Table 1.

We also checked the correlations between all the variables from the SAN dataset, excluding the five machines used for forecasting to ensure no data leakage. These correlations are shown in Fig. 1. The target variable, CPU usage, had the strongest correlation with memory usage with Pearson's  $r = 0.69$ , while other input variables had relatively low correlation values of  $r \leq 0.12$ . There was moderate correlation between some of the input variables: Net transmit and net receive ( $r = 0.51$ ), disk write and disk read ( $r = 0.26$ ), and net receive and disk write ( $r = 0.27$ ). These relationships are expected since network traffic tends to happen in both directions when communicating with the server, workloads can perform writes followed by subsequent reads (or vice versa), and in some applications, it makes sense to save the received data on disk.

Although CPU usage did not exhibit strong correlation with other input variables aside from memory usage, we chose to include all available features in the model. This decision was based on the understanding that correlations only measure linear relationship between variables and do not account for the time shifts essential in forecasting tasks (Hyndman and Athanasopoulos, 2019). Additionally, while there was some moderate correlation among a few input variables (e.g., net transmit and net receive, disk write and disk read), the level of multicollinearity was not substantial enough to warrant feature exclusion.

The rationale for incorporating all available features as input was to provide the model with diverse data, enabling it to capture different workload patterns more effectively. Consequently, we opted not to apply any feature selection techniques (such as principal component analysis) to the data. Instead, the correlation analysis served as an initial sanity check to ensure the validity of the data.

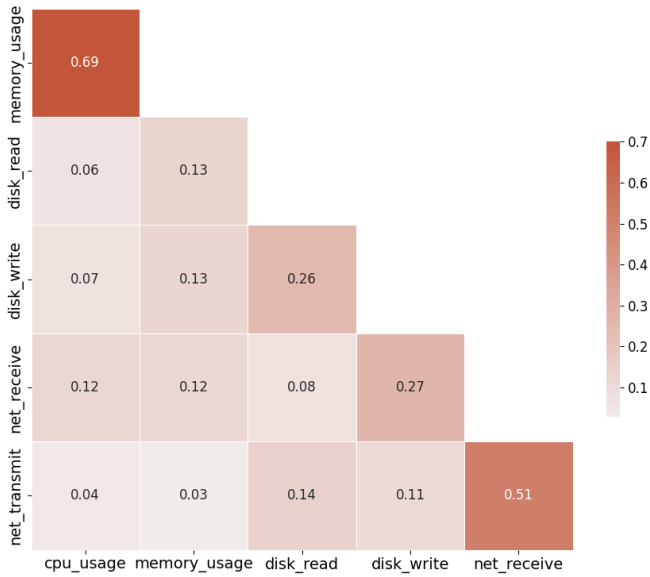


Fig. 1. Correlations between all input variables in the Bitbrains fast storage area network data. Machines used for forecasting have been excluded from the correlation analysis.

3.2 Model Architecture

The model architecture is presented in Fig. 2. Input consists of 90 timesteps of history data from 6 features which are described in Table 1. The model’s output is a forecast of CPU usage for the subsequent 6 timesteps, which corresponds to a 30-minute forecast given the 5-minute data collection intervals. The choice of a 30-minute forecasting horizon was driven by a balance between operational practicality and predictive accuracy. In data center management, it is crucial to have a sufficiently long forecasting period to enable proactive measures and resource adjustments. A 30-minute horizon provides adequate lead time to implement necessary actions such as load balancing or resource allocation. Simultaneously, this period is short enough to maintain a high level of forecast accuracy, which tends to degrade over longer horizons. CPU usage was selected as the resource to forecast because it is typically regarded as the most critical resource in a data center due to its limited availability and high demand (Zharikov et al., 2020).

1D Convolution layer was used for enhanced temporal representation of time series, thereby potentially enhancing forecast accuracy. The convolution layer had kernel with length of 6, and 35 filters. While employing a substantial amount of filters can potentially diminish model efficiency, it significantly aids in capturing diverse features inherent in the time series data. This is discussed more in Section 4.2. Since kernel with length of 6 was used with 35 filters and stride of 1 output of 1D convolution layer has dimensions (85,35). This output is fed into the recurrent layer.

The recurrent layer in our model employs either GRU or LSTM RNN units, both renowned for their ability to capture temporal dependencies effectively. In both implementations the dimension of hidden state was deliberately set to a relatively high value of 1024, and this is same as

the final output dimension of recurrent layer. This choice ensures that a comprehensive comparison of efficiency between the two RNN architectures can be conducted accurately. The rationale behind this decision lies in the understanding that a higher hidden state dimension necessitates more calculations, thereby ensuring that the results obtained are representative and robust. GRU implementation has less parameters to train than LSTM as described in Sections 2.3 and 2.4. This should make the GRU architecture more efficient than LSTM and real training time comparison between these two recurrent layers is presented in Section 4.2.

The Dense layer within the architecture requires an equal number of neurons as the desired forecast length. As detailed at the outset of this section, the forecast is for 6 subsequent timesteps. This ensures that the output dimensionality aligns precisely with the target forecast length.

Since all the 6 features were fed into 1D convolutional layer with kernel length of 6 it implies that each kernel had  $6 \times 6 = 36$  weights to be optimized. 35 filters which each had its own bias was used. This makes  $36 \times 35 + 35 = 1295$  trainable parameters total in the 1D convolution layer.

The number of trainable parameters in recurrent layer varies depending on the specific recurrent architecture employed. As told in Section 2.3 the LSTM has  $4(n^2 + nm + n)$  optimizable parameters. Again  $n$  is the dimension of hidden state and  $m$  is the dimension of input vector. In

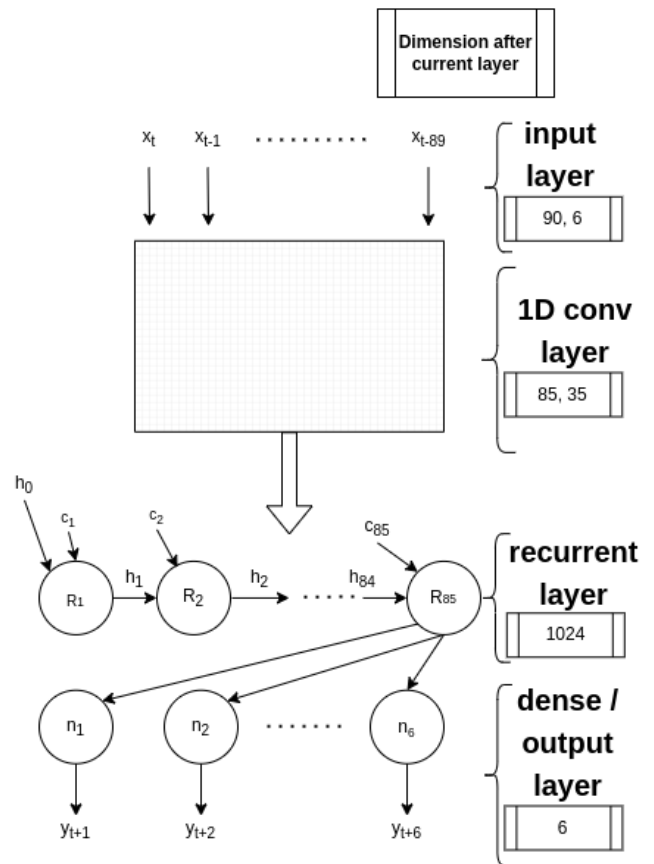


Fig. 2. Model architecture at timestep  $t$ . Dimension after each layer is shown on the right side.



this architecture this makes  $4(1024^2 + 1024 \times 35 + 1024) = 4341760$  trainable parameters for the recurrent layer with LSTM. Since Tensorflow's CUDA implementation of GRU uses two bias terms in the reset gate, the total number of trainable parameters are  $3(n^2 + nm + 2n)$ , which in this architecture is  $3(1024^2 + 1024 \times 35 + 2 \times 1024) = 3259392$  trainable parameters when using GRU on recurrent layer.

The dense layer of architecture has  $n_{output}(n_{input} + 1)$  parameters to be optimized. In this architecture it means  $6(1024 + 1) = 6150$  trainable parameters.

All these layers combined makes  $\sim 3.27\text{M}$  trainable parameters in architecture with GRU units in recurrent layer and  $\sim 4.35\text{M}$  trainable parameters with LSTM units.

### 3.3 Model Training Process

All models (GRU, LSTM, and ARIMA) were trained on data set consisting of the first 75% of data from each VM. Example of train and test split can be seen in Fig. 3.

In the training process of RNN architectures TensorFlow (Abadi et al., 2015) version 2.5 was used. 20% of the training data was further split for the validation of model and hyperparameter tuning during the training process. The hyperparameter tuning for the RNN architectures was conducted empirically through trial and error. We experimented with various combinations of key hyperparameters, including the number of layers, number of units per layer, and learning rate. The weights that yielded the highest forecast accuracy on the validation set were saved, and then later used to forecast the test in the model comparison in Section 4.1.

For the RNN architectures, min-max normalization was applied to scale all input variables to the  $[0, 1]$  range. Minimum and maximum values for each VM were calculated from the training set, and these same values were used to scale both training and test set.

The Adam optimizer combined with mean square error (MSE) loss function was used to update the weights of the RNN during the training process. Adam was chosen as optimizer because it is well suited for architectures with large amount of optimizable parameters (Kingma and Ba, 2014). The loss in both training and validation set was monitored through the training process. Example of the learning curve for the machine ID 283 can be seen in Fig. 4.

ARIMA model was optimized and trained using the `pm-darima` (Smith et al., 2017) Python package. ARIMA model parameters  $(p, d, q)$  were initially optimized using the algorithm specified in Hyndman and Khandakar (2008). Subsequently, the model was trained for each VM using its training data. The maximum depth for parameters were set to the default settings of the `pm-darima` package, specifically  $(5, 2, 5)$ . The detailed ARIMA model parameters used for each VM are provided in Table 2.

For both recurrent neural networks and ARIMA model, a distinct model was trained for each machine. In Section 4.1, the forecast experiments and results of the trained models are evaluated using the held-out test set of future time series points, as illustrated in Fig. 3.

Table 2. ARIMA model parameters for each machine

Machine ID	Parameters		
	p	d	q
220	5	1	3
242	3	1	2
253	2	1	3
269	2	1	3
283	3	1	3

Source codes and demonstration for dataset selection, ARIMA model parameter optimization and fitting, as well as GRU architecture training and forecasting are publicly available (Malin, 2024).

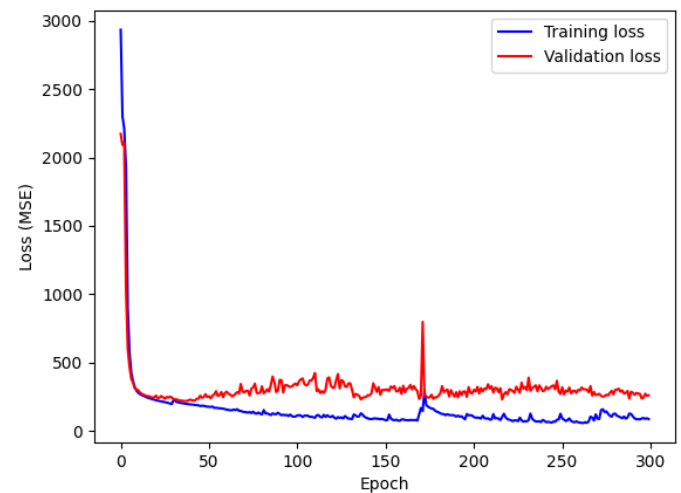


Fig. 4. Learning curve for machine ID 283.

## 4. RESULTS

In this section, we will evaluate the performance of the methods introduced in Section 2 using resource usage data from real-world scenarios and training process introduced in Section 3. This analysis aims to demonstrate the efficiency and practical applicability of the forecasting techniques in a data center environment.

### 4.1 Forecast Accuracy

In our experiments, held out test set containing last 25% of data as described in Section 3.3 was forecasted. Root mean square error (RMSE), Mean Square Error and 95th percentile of the absolute error (AE) were calculated over the forecast results for the entire test set. The results for forecast accuracy are shown in Fig. 5 and Table 3.

Prior analysis of the data provided critical insights into the problem. During this process, we discovered that some CPU usage values slightly exceeded 100%. Based on this observation, we decided to clip the forecasts of all models within the  $[0, 105]$  interval to eliminate unreasonable forecasts.

The forecast results for each VM are presented in Table 3. Error metrics were calculated from the forecast of future CPU usage in the held out test set. For every error metric, the GRU architecture achieved the best average forecast

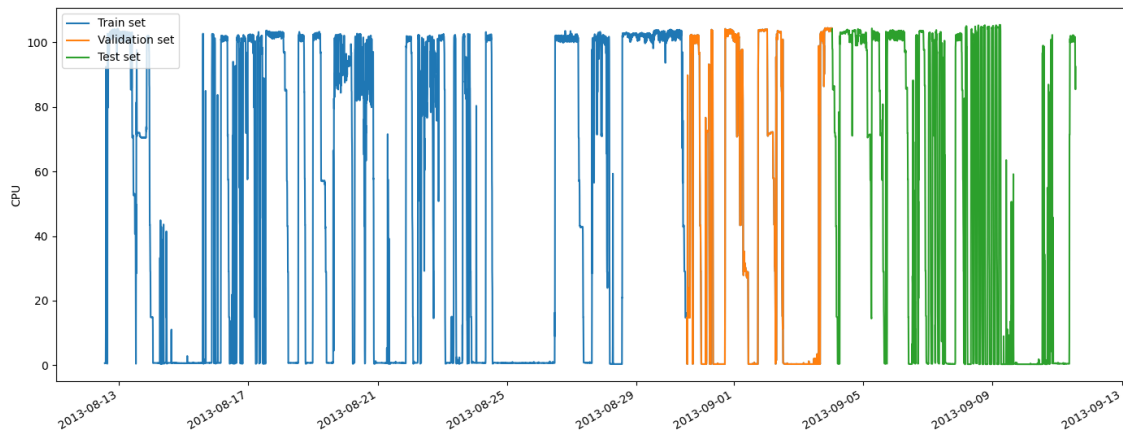


Fig. 3. Data set split example with CPU data from machine ID 253. 75% of the data was used in the training process. Last 20% of the training data was still used as a validation set to monitor metrics during training. Last 25 % of the whole dataset was held out for test set as a comparison between LSTM, GRU, and ARIMA.

Table 3. Forecasting errors for the test set of each machine. The model with the best accuracy (lowest error metric) is bolded, and the model with the lowest accuracy (highest error metric) is italicized. The mean and standard deviation are highlighted in the same manner

Machine ID	RMSE			MSE			95th AE percentile		
	GRU	LSTM	ARIMA	GRU	LSTM	ARIMA	GRU	LSTM	ARIMA
220	<b>28.72</b>	29.09	<i>30.80</i>	<b>825.05</b>	846.21	<i>948.46</i>	<b>66.38</b>	71.95	<i>84.16</i>
242	<b>25.49</b>	26.74	<i>26.81</i>	<b>650.11</b>	714.90	<i>718.69</i>	<b>62.75</b>	65.16	<i>71.29</i>
253	<b>23.14</b>	23.42	<i>24.92</i>	<b>535.55</b>	548.49	<i>621.13</i>	64.84	<b>62.63</b>	<i>71.60</i>
269	<i>20.04</i>	<b>19.37</b>	<i>19.78</i>	401.71	<b>375.11</b>	<i>391.31</i>	48.21	<b>44.48</b>	<i>47.13</i>
283	<b>20.06</b>	<i>21.00</i>	20.34	<b>402.47</b>	<i>441.01</i>	413.75	<b>48.94</b>	49.14	<i>53.30</i>
Mean	<b>23.49</b>	23.92	<i>24.53</i>	<b>562.98</b>	585.14	<i>618.67</i>	<b>58.22</b>	58.67	<i>65.50</i>
Std	<b>3.71</b>	4.00	<i>4.60</i>	<b>179.44</b>	194.44	<i>230.46</i>	<b>8.90</b>	11.47	<i>15.04</i>

performance across all machines, with an RMSE of 23.49, an MSE of 562.98 and a 95th percentile AE of 58.22. Conversely, the ARIMA model had the worst average error metrics, with an RMSE of 24.53, an MSE of 618.67 and a 95th percentile AE of 65.50. The LSTM architecture's performance fell in between these two, with average error metrics of 23.92, 585.14, and 58.67, respectively. Both the LSTM and the GRU architectures outperformed the baseline ARIMA model in forecast accuracy for every VM and across all metrics, as shown in Table 3. Additionally, when comparing the medians of the error metrics, the RNN architectures outperform the baseline ARIMA model. This can be seen from the boxplots of the error metrics in Fig. 5.

Considering all the calculated forecast error metrics in Table 3, the GRU architecture provided the lowest error in 11 out of 15 cases. In the remaining four cases where the GRU did not achieve the lowest forecasting error, the best-performing model was still an RNN architecture, specifically the LSTM. This highlights the superior performance of RNN-based models in our forecasting task.

#### 4.2 Computational Efficiency

Both RNN architectures (GRU and LSTM) were trained using the NVIDIA Tesla V100-SXM2-32GB in the super-computer Puhti at CSC – IT Center for Science. Models with architectures as described in Section 3.2 were trained for 100 epochs.

Both RNN models were trained on the same data set from machine id 357. 80% of the data was allocated for the training process, while the remaining 20% was reserved for calculating validation metrics. Although the results gathered during validation were not applied in this setting, the validation step was performed to accurately simulate the real training process of neural networks.

The total training time for the architecture with GRU units in the recurrent layer was 380.27 seconds, whereas for the architecture with LSTM units in the recurrent layer, it took 461.38 seconds. This observation supports the theory presented in Section 2.4 that GRU should be more efficient to train than LSTM.

Further improvements in efficiency could be achieved by further optimizing the convolution layer of the architecture. Methods such as max pooling in the convolution

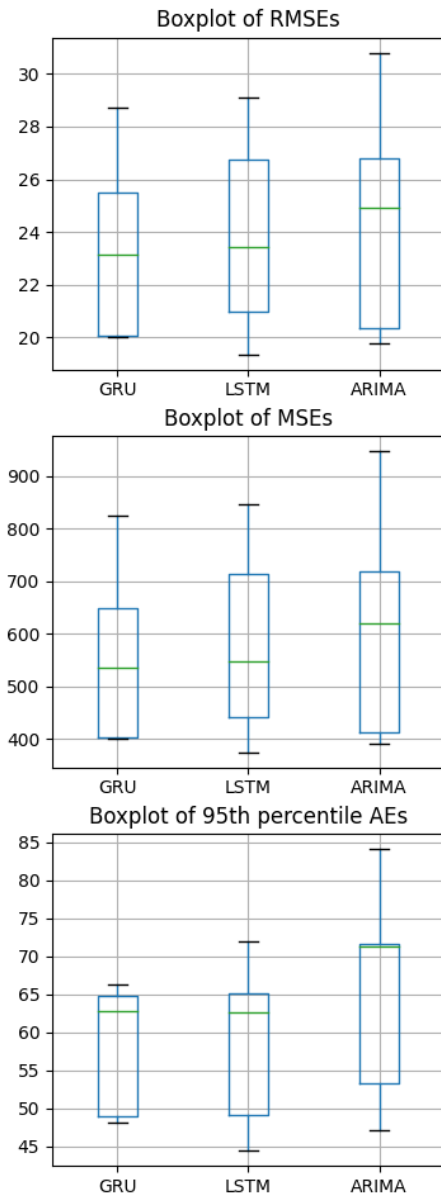


Fig. 5. Boxplot of all error metrics for the test set forecasts across different model architectures. The boxplot data holds results from all five machines used in the experiments.

layer can significantly reduce the dimensionality between the input and recurrent layers, thereby decreasing the number of weights to be tuned in the recurrent layer. However, despite initial optimization efforts, experimentation revealed that additional dimensionality reduction techniques resulted in a noticeable decrease in forecast accuracy. Therefore, while acknowledging the potential for further optimization, we decided not to pursue these techniques in this study.

Employing computationally efficient models, such as GRU, is crucial in data center load prediction. One key objective in this setting is to reduce the overall energy consumption of the computing cluster. Given that the forecasting models themselves are integral components of the system,

using efficient models helps minimize their computational overhead, thereby contributing to energy savings.

#### 4.3 Combining Accuracy and Efficiency

Table 4.3 presents a summary of the accuracy and efficiency results. The GRU model achieved a 9% reduction in forecast error, as measured by MSE, compared to the baseline ARIMA model. Additionally, the GRU model outperformed the LSTM model in terms of accuracy, despite having 25% fewer trainable parameters.

Table 4. Summary of all results. Average MSE and computational complexity is compared to the baseline ARIMA model. Rank takes both accuracy and computational complexity into account

Architecture	Average MSE	Computational Complexity	Rank
ARIMA	618.67 (Baseline)	Low	3.
LSTM	585.14 (-5.42%)	High	2.
GRU	562.98 (-9.00%)	Medium	1.

## 5. CONCLUSIONS

We have shown practical and accurate approach for forecasting IT resource usage in data centers using recurrent neural networks. In our experiments, both the LSTM and GRU architectures outperformed the baseline ARIMA model in terms of forecast accuracy.

The architecture with GRU units in the recurrent layer provided the best forecasting accuracy, as evidenced by the lowest mean and median error metrics. The LSTM units in the recurrent layer achieved the second-best forecast accuracy based on mean error metrics for the test set.

Considering that the GRU architecture not only delivered superior forecasting accuracy but also demonstrated efficient training times, we propose the use of GRU units in the recurrent layer of neural network architectures for resource usage forecasting. Reducing computational complexity in data center operations has significant practical implications. Improved forecasting accuracy and efficiency can enhance resource management, reduce energy consumption, allow dynamic scalability, and improve overall operational efficiency. These advancements can lead to cost savings and a lower environmental footprint, which are critical considerations in the field of sustainable computing.

This study provides valuable insights into forecasting data center load using RNN architectures. However, it is important to note that this work is based on the dataset from a single data center. Although the Bitbrains dataset has diverse use cases, this may limit the generalizability of the results. Future research should consider testing the models on datasets from multiple data centers to validate and potentially extend the applicability of the findings.

All the source codes for this work are publicly available (Malin, 2024), ensuring that all results can be reproduced. This will also enable future improvements to be built upon this base convolution and RNN architecture.

## ACKNOWLEDGEMENTS

The work for this paper was done as part of ArctiqDC project, with financial support from the European Regional Development Fund via the Interreg Nord program. Authors would also like to thank IT Center for Science (CSC) for providing the computing resources.

## REFERENCES

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. URL <https://www.tensorflow.org/>. [Online; accessed 28.5.2024].
- Bayati, L.M. (2018). Power management policy for heterogeneous data center based on histogram and discrete-time mdp. *Electronic Notes in Theoretical Computer Science*, 337, 5–22. doi:10.1016/j.entcs.2018.03.031.
- Bitbrains (2013). Gwa-t-12 bitbrains dataset. URL <http://gwa.ewi.tudelft.nl/datasets/gwa-t-12-bitbrains>. [Online; accessed 24.5.2024].
- Bourhane, S., Abid, M.R., Lghoul, R., Zine-Dine, K., Elkamoun, N., and Benhaddou, D. (2020). Towards green data centers. In J.L. Afonso, V. Monteiro, and J.G. Pinto (eds.), *Sustainable Energy for Smart Cities*, 291–307. Springer International Publishing, Cham. doi:10.1007/978-3-030-45694-8\_23.
- Calheiros, R.N., Masoumi, E., Ranjan, R., and Buyya, R. (2015). Workload prediction using arima model and its impact on cloud applications' qos. *IEEE Transactions on Cloud Computing*, 3(4), 449–458. doi:10.1109/TCC.2014.2350475.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734. Association for Computational Linguistics, Doha, Qatar. doi:10.3115/v1/D14-1179. URL <https://aclanthology.org/D14-1179>.
- Dey, R. and Salem, F. (2017). Gate-variants of gated recurrent unit (gru) neural networks. 1597–1600. doi:10.1109/MWSCAS.2017.8053243.
- Goodfellow, I.J., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA, USA. URL <http://www.deeplearningbook.org>.
- Hewamalage, H. (2020). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37, 388–427. doi:10.1016/j.ijforecast.2020.06.008.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735.
- Hyndman, R. and Athanasopoulos, G. (2019). Forecasting: principles and practice, 3rd edition. URL <https://otexts.com/fpp3/>. OTexts: Melbourne, Australia. OTexts.com/fpp3. [Online; accessed 13.08.2024].
- Hyndman, R. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 26, 1–22. doi:10.18637/jss.v027.i03.
- Janardhanan, D. and Barrett, E. (2017). Cpu workload forecasting of machines in data centers using lstm recurrent neural networks and arima models. *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*, 55–60. doi:10.23919/ICITST.2017.8356346.
- Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. *Nature*, 561, 163–166. doi:10.1038/d41586-018-06610-y.
- Kingma, D. and Ba, J. (2014). Adam: A method for stochastic optimization. *International Conference on Learning Representations*.
- Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D.J. (2021). 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398. doi:10.1016/j.ymsp.2020.107398.
- Kumar, J. and Singh, A.K. (2020). Cloud datacenter workload estimation using error preventive time series forecasting models. *Cluster Computing*, 23(2), 1363–1379. doi:10.1007/s10586-019-03003-2. doi:10.1007/s10586-019-03003-2.
- Malin, M. (2024). Demo for arima and gru model training. URL <https://github.com/miikamal/dc-rnn>. [Online; accessed 16.08.2024].
- Masanet, E., Shehabi, A., Lei, N., Smith, S., and Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science*, 367(6481), 984–986. doi:10.1126/science.aba3758.
- Ouhame, S., Hadi, Y., and Ulah, A. (2021). An efficient forecasting approach for resource utilization in cloud data center using cnn-lstm model. *Neural Computing and Applications*, 33. doi:10.1007/s00521-021-05770-9.
- Shaw, S.B., Kumar, C., and Singh, A.K. (2017). Use of time-series based forecasting technique for balancing load and reducing consumption of energy in a cloud data center. In *2017 International Conference on Intelligent Computing and Control (I2C2)*, 1–6. doi:10.1109/I2C2.2017.8321782.
- Shehabi, A. et al. (2016). United states data center energy usage report. no. LBNL-1005775. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Smith, T.G. et al. (2017). pmdarima: Arima estimators for Python. URL <http://www.alkaline-ml.com/pmdarima>. [Online; accessed 27.5.2024].
- Yuan, H., Bi, J., Li, S., Zhang, J., and Zhou, M. (2024). An improved lstm-based prediction approach for resources and workload in large-scale data centers. *IEEE Internet of Things Journal*, 1–1. doi:10.1109/JIOT.2024.3383512. Early Access.
- Zhang, W., Li, B., Zhao, D., Gong, F., and Lu, Q. (2016). Workload prediction for cloud cluster using a recurrent neural network. 104–109. doi:10.1109/IIKI.2016.39.
- Zharikov, E., Telenyk, S., and Bidyuk, P. (2020). Adaptive workload forecasting in cloud data centers. *Journal of Grid Computing*, 18. doi:10.1007/s10723-019-09501-2.

## Evaluation of Model Uncertainty Propagation in Mineral Process Flowsheet Designs

Henri Välikangas\*, Markku Ohenoja\*, Stéphane Brochot\*\*, Manuel González Fernández\*\*, Jari Ruuska\*, Mika Ruusunen\*.

\**Environmental and Chemical Engineering Research Unit, Control Engineering Group, Faculty of Technology, P.O.Box 4300, University of Oulu, Oulu, 90014, Finland (e-mail: [markku.ohenoja@oulu.fi](mailto:markku.ohenoja@oulu.fi)).*

\*\* *CASPEO, 45800, Saint-Jean-de-Braye, France*

---

**Abstract:** Increasing demand for critical raw materials and energy transition metals sets new targets for the mineral processing, also resulting as higher requirements for the simulation tools during process design and optimization. This study presents a framework for global uncertainty evaluation of modelled plant-wide processes, where the propagation of uncertainty sources is addressed. The uncertainties exist, for example in operational and design parameters and in material properties. The approach was demonstrated with a typical mineral processing flowsheet simulated with commercial software. First, domain knowledge was adopted to screen the parameter space and then Monte Carlo simulation was performed. After this, the generated data set was used to identify surrogate models between the uncertain inputs and process performance indicators. Finally, a global sensitivity analysis was conducted to identify the effects of uncertainties to the decision-making in process design. The results were particularly used to locate the process points where accurate information is needed for the robust process design, or where on-line measurements would be preferred to establish on-line optimization.

*Keywords:* Flowsheet simulation, Process design, Global sensitivity analysis, Surrogate model.

---

### 1. INTRODUCTION

As critical minerals have become more crucial for the operation of societies, the necessity to maximize the efficiency of all processes throughout the lifecycle, namely mining, refining, and recycling, is even more important than before. Increase in the demand of critical minerals opens new demand for circular economy system, in which the maximum efficiency is achieved by minimizing losses in all parts of the cycle (Whitworth et al., 2022). These developments increase the number of processes, where multiple minerals are present in the separation processes, which in turn increases the complexity of simulation models, and uncertainty of simulation and model-based decision making in process design and in process operation.

Mineral processes aim to extract valuable minerals from ore. The process usually consists of multiple stages, which all of them have their unique properties, and thus described with different mathematical models and uncertainties related to them. The uncertainties need to be attributed to their sources through simulations to facilitate the process optimization (Sepúlveda et al., 2014).

The lack of understanding that exists around inspected system creates a need to model the system, which itself holds inherent uncertainty, for example assumptions, process randomness and measurement errors (Caers, 2011a). Precisely, the definition of uncertainty is tied to the model uncertainty when it is quantified by sensitivity analysis (Sepúlveda et al., 2014, Arnst et al., 2021, Puy et al., 2022), although uncertainty is a

wider concept itself. According to Campolongo et al., (2000), sensitivity analysis complements uncertainty analysis.

The goal of extracting valid information is to reduce uncertainty in an influential decision-making process. Because collecting more information does not necessarily reduce uncertainty, it is important to find the parameters that best describe uncertainty (Caers, 2011b).

Sensitivity analysis is a method that can be also applied to identify how the uncertainty in model output is divided in its inputs. There, a local sensitivity analysis provides changes one input parameter at a time. Global Sensitivity Analysis (GSA) is a more robust solution compared to local sensitivity analysis (Cisternas and Lucay, 2020). It can overcome the limitations of inspecting one variable at a time, and thus enables to find relationships between the variables that would be otherwise left undiscovered (Sepúlveda et al., 2013).

GSA has been applied in mineral processing, for example, improving the milling operation by (Lucay et al., 2019) as they considered both the operational (epistemic) uncertainties and stochastic uncertainties related to feed properties. Further, a framework of deterministic process design, elimination of non-influential process variables and recognition of critical parameters through GSA was used in (Lucay et al., 2015) for a mineral concentration process. Ohenoja et al. (2023) used GSA to identify and to weight the most important process measurements in the model adaptation problem of a flotation circuit. Arancibia-Bravo et al (2022) similarly used GSA to identify critical model input parameters of copper flotation in saline systems, while (Sitorus and Brito-Parada, 2020) applied



GSA for the selection of optimal crushing equipment in multiple criteria decision-making model.

As mentioned, mineral processes are characterized by a combination of multiple processing stages making the overall flowsheet complex. The input parameter effect on the uncertainty of the global output parameters of each separation unit can be analyzed by dividing the inspected flowsheet into stages (Montenegro et al., 2015). By doing this, the propagation of the uncertainty can be analyzed. One approach to gain insight on uncertainty propagation is Monte Carlo (MC) simulation (Albert, 2020).

MC requires a sufficiently big sample size to produce sufficient resolution for the intended purposes (Helton and Davis, 2003). Thus, complex flowsheets or detailed models may set limitations to the applicability of MC. Therefore, surrogate models are also used in uncertainty evaluation. Analytical solutions to uncertainty have been inspected utilizing surrogate models, for example in (Liu et al., 2024), where the presence of two or more uncertainty factors is the source of complexity. Lu et al. (2018) showed that generalized linear models can be used to get accurate sensitivity indices, by utilizing polynomial approximations of the data.

This study presents and demonstrates a framework for global uncertainty evaluation of plant-wide processes, where the propagation of uncertainty sources is addressed. The uncertainty propagation results in this publication give insight of the uncertainty factors and sources, that would be used during the model-based process design or in operational optimization. Thus, this work aims to extend from previous GSA studies, such as (Lucay et al., 2012), where focus was on one separation process model. The uncertainty evaluation framework is demonstrated utilizing a typical mineral processing flowsheet simulated with a commercial software.

The following sections of the paper are distributed as follows; Section 2 outlines the constructed framework, and the software and mathematical methods used. Section 3 details the selected mineral processing case study, the performance of the surrogate models identified, and the sensitivity analysis results. The evaluation of the case study results, and discussion of the proposed framework is described along Section 4. Finally, Section 5 summarizes the main findings of the research.

## 2. MATERIAL AND METHODS

### 2.1 Framework

The framework is described in Fig. 1. As a starting point, the mineral processing flowsheet model was established to a simulation software. Then, the possible input and output parameters from the software used were listed with domain knowledge and inspected through a screening step. This was performed as a local sensitivity analysis by directly inspecting the variation caused by each input to one output individually. As a result of the screening step, the number of possible input parameters for the global sensitivity analysis were reduced. This is typically necessary to facilitate the MC simulation of complex flowsheets.

The final selection of parameters, and their ranges, were confirmed from the forementioned list based on domain knowledge. After the parameters had been chosen, a MC simulation was performed in the flowsheet simulation software. The resulting data set was then utilized to train the surrogate models and to perform GSA with the identified models. To improve the performance of surrogate models, and the sensitivity analysis based on them, the flowsheet was considered as blocks, where the surrogate models of previous blocks can act as inputs for the following modelled blocks.

In the demonstration of this study, the simulation software used was USIM PAC (See Section 2.2). The screening step was analyzed using spreadsheets and interviewing the experts. The MC was conducted with the embedded MC tool in the simulation software. The MC data set was then exported to Matlab® to identify surrogate models (See Section 2.3) and to perform sensitivity analysis (See Section 2.4). The studied flowsheet is presented in Section 3.1.

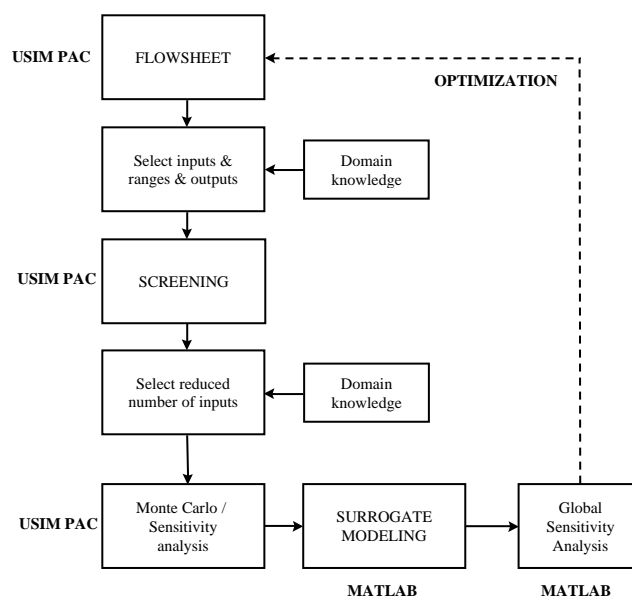


Fig. 1. Approach for estimating uncertainty propagation in flowsheet simulation.

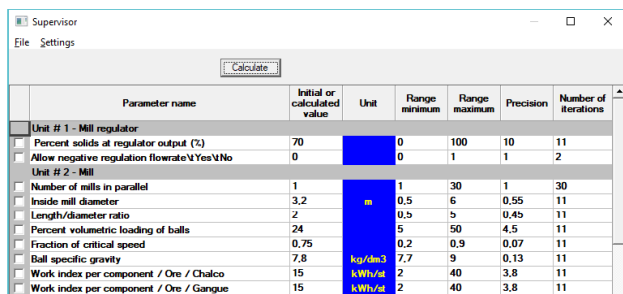
### 2.2 Simulation software

In active development since 1986, USIM PAC has been created by the BRGM's (French geological survey) Process Simulation Group. Since 2004, CASPEO, a spin-off of BRGM, has been the company behind its development and distribution. Although it has been used in several industries, USIM PAC is a process simulation software primarily intended for mineral processing and hydrometallurgical operations, where it can be used for design or optimization purposes.

The Supervisor, which is one of the optimization algorithms available in USIM PAC (Guillaneau et al., 1995) was the main calculation tool used to generate the simulation results in this work. The Supervisor algorithm can be used either as a sensitivity analysis tool or for visual optimization. It calculates user-defined parameters (soft-sensors, outputs) when some

input parameters (actuators) vary. The variation of the actuators can be defined using different methods: (1) Scanning, which generates a set of vectors by the combination of different values of each parameter in a given research domain; (2) Sensitivity Analysis, which evaluates each parameter using a user-defined range and step; (3) Monte-Carlo, which generates as many parameter values as required using a random procedure with the selected statistical distribution (Gauss, Uniform, or other) to constitute a point.

A screenshot of the Supervisor tool is depicted in Fig. 2. The output of the Supervisor tool is a file displaying the list of the values of the user-defined soft-sensors resulting from the simulations performed for each random value of the actuator or combination of actuators. As this file can be exported as a spreadsheet, the results can be easily exploited using statistical analysis tools.



Parameter name	Initial or calculated value	Unit	Range minimum	Range maximum	Precision	Number of iterations
<b>Unit # 1 - Mill regulator</b>						
Percent solids at regulator output (%)	70		0	100	10	11
Allow negative regulation flowrate Y Yes/No	0		0	1	1	2
<b>Unit # 2 - Mill</b>						
Number of mills in parallel	1		1	30	1	30
Inside mill diameter	3.2	m	0.5	6	0.55	11
Length/diameter ratio	2		0.5	5	0.45	11
Percent volumetric loading of balls	24		5	50	4.5	11
Fraction of critical speed	0.75		0.2	0.9	0.07	11
Ball specific gravity	7.8	kg/dm <sup>3</sup>	7.7	9	0.13	11
Work index per component / Ore / Chalco	15	kWh/t	2	40	3.8	11
Work index per component / Ore / Gangue	15	kWh/t	2	40	3.8	11

Fig. 2. Selection of actuators in USIM PAC Supervisor.

### 2.3 Surrogate modeling

The data acquisition from USIM PAC to the surrogate modeling was performed using the MC simulation property in USIM PAC. The resulted \*.csv was read in Matlab®, where the surrogate models were fitted using Regression Learner application. The selected modelling approach here was linear stepwise regression. The performance of the final model structure was evaluated with two different metrics: mean absolute percentage error (MAPE) and coefficient of determination ( $R^2$ ).

### 2.4 Sensitivity analysis

According to (Campolongo et al., 2000), sensitivity analysis is an integral part of the modeling process. As a quantitative method, it can decompose the variance of output variable  $Y$ . It can be used as a tool to identify noninfluential parameters, and thus be used to simplify and/or improve the uncertainty modeling.

The total sensitivity index takes into consideration all the input parameters ( $X_i$ ) and their possible combinations ( $X_{ij}$ ) and displays the average effect of the inspected input variable (Lucay et al., 2019). According to (Saltelli et al., 2007), the first order sensitivity index being similar in magnitude to the total effect index, means that there is no interaction between the inspected parameter and the rest of the parameters. Otherwise, the first order sensitivity index is always smaller than the total order index, if there is even a small interaction between the inspected parameter and other parameters.

The GSA approach was originally proposed in (Saltelli and Homma, 1996). The refined method in (Saltelli, 2002), gives a pathway to circumvent the curse of dimensionality when dealing with high factor count models, turning  $n2^k$  into  $n(2k + 2)$ , where  $k$  is a term of order and  $n$  is the sample size used to estimate one individual effect. They noted that the computation of the sensitivity indices is more straightforward in the higher order terms. The advantages of the method are the flexibility concerning the utilized models in the sensitivity analysis and the computational inexpensiveness. Thus, the method from (Saltelli, 2002) is more attractive tool for engineering applications, and was also selected to this study.

The total order index,  $S_{Ti}$ , is formed by following formula (Saltelli et al., 2007, p.164):

$$S_{Ti} = 1 - \frac{V[E(Y|X_{-i})]}{V(Y)}, \quad (1)$$

where  $i$  refers to the input parameter,  $V(Y)$  is the variance of the inspected output  $Y$ , and  $E[Y|X_{-i}]$  is the estimated conditional mean of output  $Y$  in relation to input  $X_{-i}$ .  $V(E[Y|X_{-i}])$  is the conditional variance of output  $Y$  in relation to input  $X_{-i}$ .

The sensitivity analysis was done in Matlab® utilizing Latin hypercube (LHS) sampled data based on the utilized parameter ranges. The used functions for the sensitivity analysis can be found in (Vandy, 2016).

## 3. RESULTS

### 3.1 Input screening and the studied flowsheet

The flowsheet used in simulations is presented in Fig. 3. The grinding circuit (GC) comprises a ball mill and a hydrocyclone classification with one recycle stream. The flotation circuit (FC) includes four flotation stages (Rougher, Scavenger, Cleaner 1, Cleaner 2) with two recycle streams.

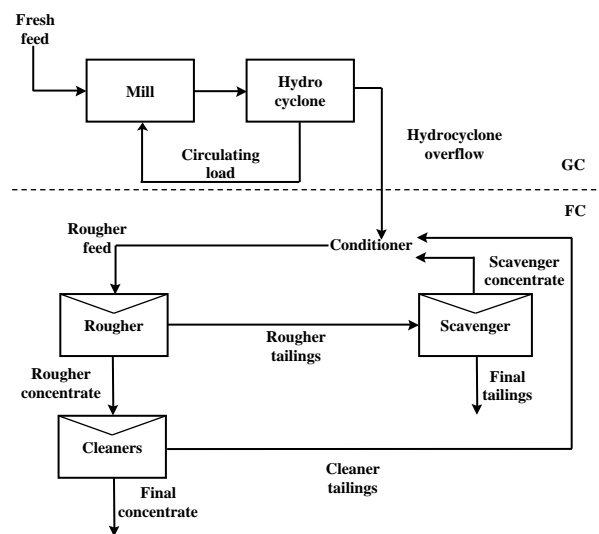


Fig. 3. Studied flowsheet.

In the studied flowsheet, the streams under interest were:

- GC product,
- Circulating load to the mill (hydrocyclone underflow),
- Rougher feed,
- Final tailings, and
- Final concentrate.

From these streams, several properties were monitored as outputs, namely:

- Ore concentration,
- Particle size,
- Ore mass flowrate,
- Total volumetric flowrate,
- Solids concentration, and
- Grade.

These represent the outputs *Y*, that are modeled and subjected to GSA.

For the screening step, a large number of input parameters were reduced to a smaller group of important input parameters with domain knowledge and simulations using the Scanning feature of USIM PAC. Table 1 presents the selected input parameters after the screening. For the grinding circuit, mill rotation speed and grinding media loading represent operational variables, whereas grindability is a material property. Similarly in flotation circuit, the pulp level and water content can be manipulated in an operational environment. The floatabilities can be considered either material properties (liberation, mineral properties) or operational variables (addition of flotation chemicals). The different flotation cells in the flowsheet, namely rougher, scavenger, and two cleaners, have unique parameters. As mentioned in Section 2.1., the surrogate model outputs from the grinding circuit also act as inputs for the flotation circuit surrogate models.

Table 1. Selected process parameters for MC simulation and surrogate modeling.

Grinding circuit	Flotation circuit
Grinding media loading	Pulp level in cell
Mill speed	Pulp water content in cell
Ore grindability	Ore floatability in cell
Gangue grindability	Gangue floatability in cell
Hydrocyclone feed diameter	Surrogate model outputs from the grinding circuit
Hydrocyclone overflow diameter	
Hydrocyclone underflow diameter	

### 3.2 Monte Carlo simulation

MC simulation was performed utilizing uniform distribution for the input parameters. The used range of the input

parameters varied from ±5% to ±10% in USIM PAC supervisor. The variation amplitude was based on domain knowledge. The number of MC simulations done in USIM PAC was 10,000. The generated data was used to fit surrogate models for the GSA. In GSA, MC was used in generating sample inputs for the surrogate models utilizing LHS design. The input range was extrapolated to ±10% for all inputs. The sample data consisted of 10,000,000 points.

### 3.3 Surrogate modeling

In total, 30 surrogate models, representing the outputs *Y*, were identified using the MC data set from USIM PAC. The model performance, in terms of MAPE, for the GC outputs and FC outputs are presented in Fig. 4 and Fig. 5, respectively. Overall, the low MAPE values (< 1.8%) indicate that the GC stream properties from the flowsheet simulation can be accurately described with surrogate models. The solids concentration in Fig. 5 lacks modeling error values for Final concentrate or Rougher feed streams, as the MC data set indicated constant output values.

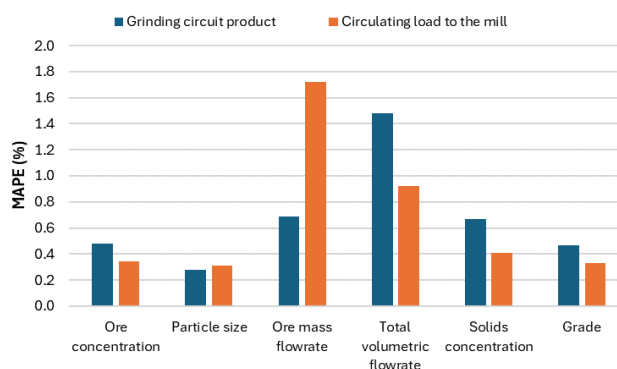


Fig. 4. Model performance (MAPE) of the identified surrogate models for GC.

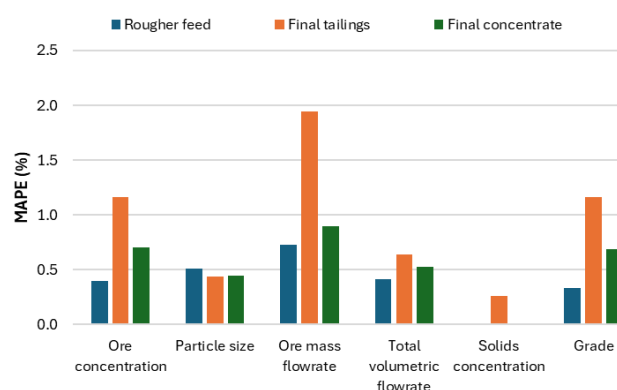


Fig. 5. Model performance (MAPE) of the identified surrogate modes for FC.

The scatter plots of the worst performing surrogate models (Circulating load to the mill and ore mass flowrate in final tailings) are presented in Fig. 6 and Fig. 7, respectively. The MAPE values inspected together with  $R^2$  give a more comprehensive understanding of the model performance. The figure shows that the  $R^2$  values are also at acceptable levels

(greater than 0.70) in this case. For the ore mass flowrate in Fig. 6, the surrogate model seems to systematically underestimate some of the values above 11.8 t/h. Thus, for this variable, another model structure could be studied to improve the modeling performance further.

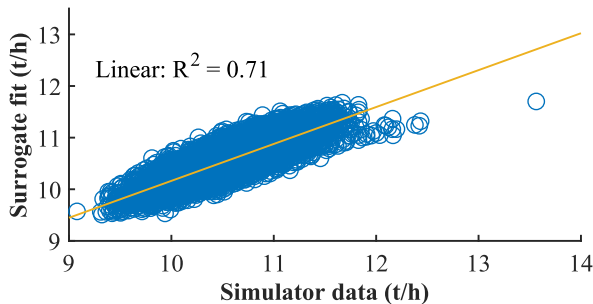


Fig. 6. Scatter plot of ore mass flowrate in the Circulating load to the mill. The scale in both axes begins at 9.

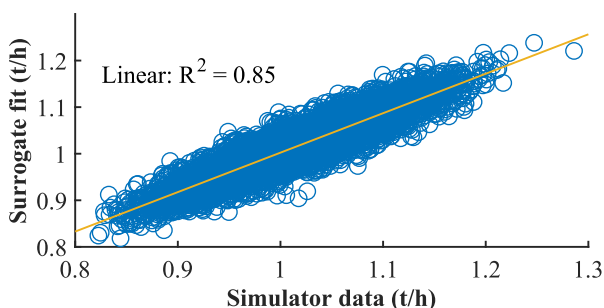


Fig. 7. Scatter plot of ore mass flowrate in the Final tailings. The scale in both axes begins at 0.8.

### 3.4 Global sensitivity analysis

The GSA was performed using identified surrogate models. In GSA, the results are interpreted individually to different flowsheet sections in order to understand the propagation routes of the uncertainties.

#### Grinding circuit

The utilized input parameters for GC were listed in Table 1. From the GSA results, some essential outputs, such as in PSD (Particle size distribution), ore concentration and grade were inspected in detail and are discussed below.

PSD is usually the key process quality parameter in grinding circuits. According to the results, PSD of the Grinding circuit product is mainly affected by the hydrocyclone parameters ( $S_{Ti}$  from 0.51 to 0.06), whilst mill and ore parameters had very small sensitivity ( $S_{Ti} < 0.04$ ) to the PSD. However, for Circulating load to the mill, the gangue grindability (0.41), mill speed (0.32) and the grinding media loading (0.25) explain the PSD variation according to the GSA.

Grinding circuit product concentration variation is best explained by gangue grindability (1st, 0.55), ore grindability (2nd, 0.45) and hydro cyclone underflow (3rd, 0.11). The same finding applies to the concentration in Circulating load to the mill, and to the grade variation in both streams.

Overall, the operational variables show only moderate sensitivity to the studied outputs in GC; Grinding media loading affects the total volumetric flowrate of GC product and PSD of circulating load. Mill speed affects the GC product solids concentration and PSD of circulating load. One explanation could be that the selected ranges of the other input parameters mask the effect of mill operational parameters in most of the studied outputs. One way to overcome this problem in sensitivity analysis would be to narrow down the parameter ranges of feed characteristics, or to sample the parameter values from different types of probability distributions as done in (Lucay et al. 2019).

#### Flotation circuit

The utilized input parameters for FC are listed in Table 1. In addition to these parameters, the outputs from the GC surrogate models are used as surrogate model inputs, and thus, in GSA. Three of those were identified to be very significant inputs parameters for all FC outputs.

In all three FC streams, the concentration in the GC product stream shows significant sensitivity to the ore concentration (or grade), and particle size ( $S_{Ti}$  between 0.47 and 0.29). This is an expected result, as variation in flotation fresh feed properties determine the flotation performance. According to the GSA results, the gangue floatability in rougher and pulp water content in cleaner 1 are also sensitive parameters ( $S_{Ti}$  between 0.10 and 0.16) to the rougher feed and final concentrate concentrations, respectively.

Interestingly, the ore concentration and particle size in Circulating load to the mill seems to explain some variation in flotation streams' properties ( $S_{Ti}$  up to 0.30 and 0.32, respectively). This might be contributed by the fact that the circulating load affects the water addition rates in the flowsheet model, which then propagates into a variation in flotation circuit. Another explanation could be that the mentioned circulating load properties are affected by mill speed, grinding media loading, and hydrocyclone parameters, which were also seen in the GSA results for the grinding circuit.

Regarding the solids concentration in final concentrate, the GSA shows that the total pulp water content in cleaner 2 is the most influencing variable ( $S_{Ti}$  0.85). This is natural, as the higher water content in the final flotation cell corresponds to the lower solids concentration of the product. Otherwise, the flotation circuit operational parameters are not among the most sensitive parameters in the simulated data set, again highlighting the need to carefully determine the input's probability distributions for the sensitivity study. Another observation from these results is that there can be limited possibilities for the operational parameters to mitigate the effect of the disturbances entering to the FC from upstream process steps.

#### Summary of GSA results

The GC circulating load parameters have a high effect ( $S_{Ti} > 0.10$ ) to the flotation circuit variation. This is due to the mill being unable to process the feed (and recycle) fast enough, thus increasing the amount of water and decreasing the solids



concentration in the feed to the FC. This is observable from the parameters with influential  $S_{Ti}$  in the circulating load being ore grindability, mill speed and mill grinding media loading.

In Fig. 8, the occurrence of the most sensitive inputs for the rougher feed properties is depicted. The frequency corresponds how often the input parameter was among the three most significant input parameters based on GSA. Similarly, Figure 9 shows the result for final tailings properties and Figure 10 for the final concentrate properties.

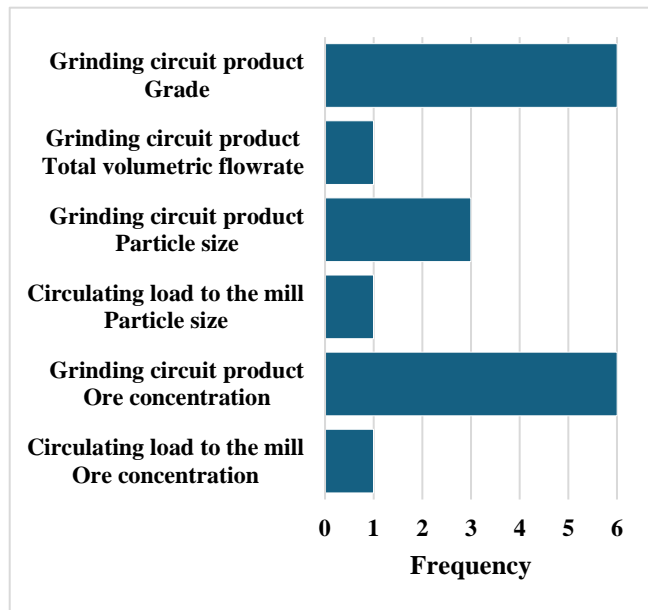


Fig. 8. Occurrence of the most sensitive inputs for rougher feed properties.

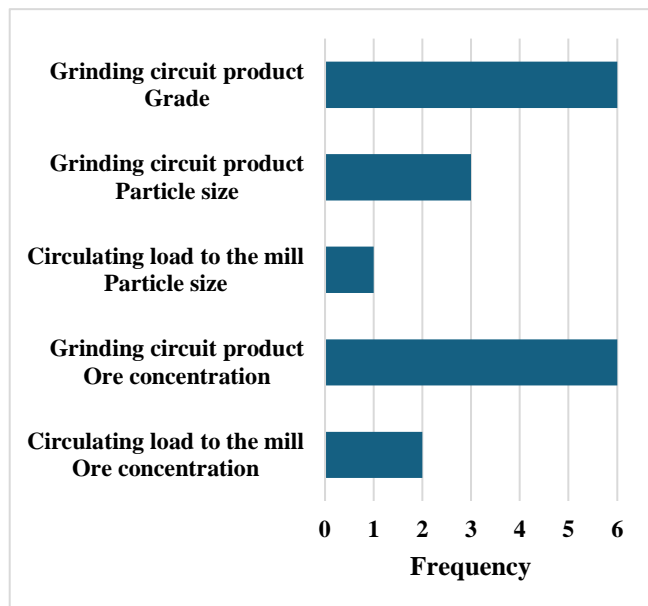


Fig. 9. Occurrence of the most sensitive inputs for FC final tailings properties.

From Fig. 8, it can be observed that the rougher feed properties are naturally sensitive to the parameters in GC product. According to the results, the input parameters that describe the

ore concentration, grade and particle size thus describe the flotation circuit performance and the final concentrate variation.

The ore feed characteristics and the ability of the GC to categorize the feed into set particle size similarly most describe the variation in the final concentrate and final tailings (See Fig. 9 and Fig. 10), which is an expected result after the rougher feed results. An outlier to this statement is the Pulp water content in cleaner 2, which is an operational input parameter in FC. As the fresh feed characteristics can't be affected by the operation of the GC, the only thing left to do is to minimize the variation in the ore size distribution by GC operation.

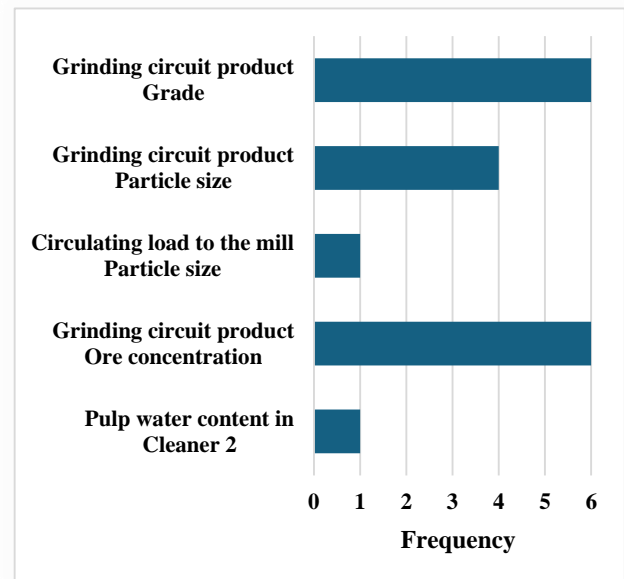


Fig. 10. Occurrence of the most sensitive inputs for FC final concentrate properties.

#### 4. DISCUSSION

In addition to the surrogate models presented, the FC outputs were also modelled using only the MC data without GC surrogates. The performance of the FC surrogates, in this case, was slightly worse ( $>0.05$  lower  $R^2$ ). This alternative would also make it more difficult to inspect the propagation of uncertainties as GC models do not act as inputs for the FC.

Production of outliers by the GSA method, or volatility, is due to its non-additive nature. Non-additivity in this context means, that the generated  $S_{Ti}$  value is not equal to the sum of the values of the component parts. This is a cost caused by the computational and straightforward nature of the used Saltelli's approach over the original Sobol's method (Saltelli et al., 2007). This limits the interpretation of the lower magnitude  $S_{Ti}$ . Thus, mainly the three most significant  $S_{Ti}$  were discussed in this study. Naturally, a more thorough GSA interpretation could involve inputs with lower influences into the analysis of uncertainty, if the volatility issue can be solved.

Like (Puy et al., 2022) concludes in their publication, Saltelli's total order method becomes inaccurate with higher dimensionality ( $k > 10$ ). The exponential growth in input



parameters of more complex flowsheets is an issue that needs to be addressed thoroughly, utilizing similar design of experiments and domain knowledge methods that were shown in this study. Domain knowledge gives insight on what parameters are influential based on previous experiences and the local sensitivity analyses on the possible unknown sensitivities in the analyzed setup.

The proposed approach for the uncertainty analysis with surrogate models resulted in a lower computational load and made it possible to use large sample size in GSA. This decreases the inherent volatility of Saltelli's method. The accuracy of Saltelli's method for the most important parameters can be improved by repeating the GSA with multiple LHS designs (Puy et al., 2022). Alternative methods for calculating total order sensitivity indices, such as Jansen, Razavi and Gupta, Janon/Monod and Azzini and Rosati, suggested by (Puy et al., 2022) could also be considered.

The results in this study suggested that the most influential  $S_{Ti}$  are related to the feed characteristics and the mill performance. Thus, for more robust operational decisions, the focus could be shifted to better measure the feed properties in on-line to minimize the effect of stochastic (inherent) uncertainty. For process design purposes, the results highlight the emphasis needed for GC design and its flexibility to tackle and decrease the uncertainties, that will otherwise strongly propagate downstream to FC streams. If stochastic uncertainties remain, only the epistemic uncertainties can be affected, and thus the focus needs to be on lesser magnitude, operational sensitive indices. To achieve the best design and operational reduction of uncertainty, maximum range of uncertainties (all  $S_{Ti} > 0.10$ ) need to be considered.

The proposed framework could benefit from the development of software interfaces establishing more automated data transfer. Dividing the process into parts is beneficial for understanding the propagation of uncertainty through the different parts of the process. At the same time, more data is generated that needs to be handled efficiently between the different software tools. Further, the interface between the sensitivity analysis results and decision-making in process design or on-line operation requires further development.

## 5. CONCLUSIONS

In this work, an uncertainty propagation evaluation framework was established to be used in model-based decision making. The proposed methods were chosen to enable rapid inspection of complex systems typically seen in flowsheet simulators. The results in this paper demonstrated that this framework can be used to identify the most influential parameters throughout the whole inspected process chain. This allows to focus further analyses to the propagation of uncertainty attributed to these identified parameters.

Regarding the case study, the propagation of uncertainty within the studied flowsheet was observed through dividing the flowsheet into the grinding circuit (GC) and to the flotation circuit (FC). By doing this, the changes in the GC were seen to have strong influence on the flotation circuit model output variation. The GC inputs were the top three most influential

inputs for all observed output parameters in the FC. The most sensitive operational FC input parameters were not found among the top three most influential input parameters but were still considered influential (total sensitivity index values  $>0.10$ ). Those parameters were gangue floatability in rougher and pulp level in cleaner.

Final sensitivity indices indicating the most sensitive parameters in the whole process were found in the fresh feed characteristics (ore concentration/grade and particle size distribution), and the ability of the mill to reduce the particle size distribution to the desired range. The operational input parameters had a lower influence in general, but that result might be due to the small range of changes applied in simulation, so further studies are required.

## ACKNOWLEDGEMENTS

This work was conducted as a part of EXCEED project (<https://exceed-horizon.eu/>), funded by the European Union's Framework Programme for Research and Innovation Horizon Europe under Grant Agreement No. 101091543.

## REFERENCES

- Albert, D. R. (2020). Monte Carlo Uncertainty Propagation with the NIST Uncertainty Machine. *Journal of Chemical Education*, 97(5), 1491–1494. doi:10.1021/acs.jchemed.0c00096
- Arancibia-Bravo, M. P., Lucay, F. A., Sepúlveda, F. D., Cortés, L., and Cisternas, L. A. (2022). Response Surface Methodology for Copper Flotation Optimization in Saline Systems. *Minerals*, 12(9). doi:10.3390/min12091131
- Caers, J. (2011a). ch3 Modeling Uncertainty: Concepts and Philosophies BT - Modeling Uncertainty in the Earth Sciences., 39–54. doi:10.1002/9781119995920.ch3
- Caers, J. (2011b). ch11 Value of Information BT - Modeling Uncertainty in the Earth Sciences., 193–213. doi:10.1002/9781119995920.ch11
- Campolongo, F., Saltelli, A., and Tarantola, S. (2000). Sensitivity Analysis as an Ingredient of Modeling. *Statistical Science*, 15(4), 377–395. doi:10.1214/ss/1009213004
- Cisternas, L. A., and Lucay, F. A. (2020). Trends in Modeling, Design, and Optimization of Multiphase Systems in Minerals Processing. *Minerals*, 1–28. doi:10.3390/min10010022
- Guillaneau, J. C., Villeneuve, J., Brochot, S., Durance, M. V., and Fourniguet, G. (1995). The supervisor of simulation: A step further to meet the process engineering needs. *Proceedings of the XIX International Mineral Processing Congress*, 1(1), 293–296. <https://www.caspeo.net/wp-content/uploads/2020/02/1995-impac-usim-pac-mineral-processing-simulator-supervisor.pdf>

- Helton, J. C., and Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety*, 81(1), 23–69. doi:10.1016/S0951-8320(03)00058-9
- Liu, J., Shi, Y., Ding, C., and Beer, M. (2024). Hybrid uncertainty propagation based on multi-fidelity surrogate model. *Computers and Structures*, 293(October 2023), 107267. doi:10.1016/j.compstruc.2023.107267
- Lu, R., Wang, D., Wang, M., and Rempala, G. A. (2018). Estimation of Sobol's sensitivity indices under generalized linear models. *Communications in Statistics - Theory and Methods*, 47(21), 5163–5195. doi:10.1080/03610926.2017.1388397
- Lucay, F. A., Gálvez, E. D., Salez-Cruz, M., and Cisternas, L. A. (2019). Improving milling operation using uncertainty and global sensitivity analyses. *Minerals Engineering*, 131(September 2018), 249–261. doi:10.1016/j.mineng.2018.11.020
- Lucay, F., Cisternas, L. A., and Gálvez, E. D. (2015). Global sensitivity analysis for identifying critical process design decisions. *Chemical Engineering Research and Design*, 103, 74–83. doi:10.1016/j.cherd.2015.06.015
- Lucay, F., Mellado, M. E., Cisternas, L. A., and Gálvez, E. D. (2012). Sensitivity analysis of separation circuits. *International Journal of Mineral Processing*, 110–111, 30–45. doi:10.1016/j.minpro.2012.03.004
- Montenegro, M. R., Gálvez, E. D., and Cisternas, L. A. (2015). The effects of stage recovery uncertainty in the performance of concentration circuits. *International Journal of Mineral Processing*, 143, 12–17. doi:10.1016/j.minpro.2015.08.004
- Ohenoja, M., Koistinen, A., Hultgren, M., Remes, A., Kortelainen, J., Kaartinen, J., Peltoniemi, M., and Ruusunen, M. (2023). Continuous adaptation of a digital twin model for a pilot flotation plant. *Minerals Engineering*, 198(April), 108081. doi:10.1016/j.mineng.2023.108081
- Puy, A., Becker, W., Lo Piano, S., and Saltelli, A. (2022). a Comprehensive Comparison of Total-Order Estimators for Global Sensitivity Analysis. *International Journal for Uncertainty Quantification*, 12(2), 1–18. doi:10.1615/Int.J.UncertaintyQuantification.2021038133
- Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2), 280–297. doi:10.1016/S0010-4655(02)00280-1
- Saltelli, A., and Homma, T. (1996). Importance measures in global sensitivity analysis of model output. *Reliab. Eng. Sys. Safety*, 52, 1–17. doi:10.1016/0951-8320(96)00002-6
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2007). *Global Sensitivity Analysis. The Primer*. Wiley. doi:10.1002/9780470725184
- Sepúlveda, F. D., Cisternas, L. A., and Gálvez, E. D. (2013). Global sensitivity analysis of a mineral processing flowsheet. In *Computer Aided Chemical Engineering (Vol. 32)*. Elsevier B.V. doi:10.1016/B978-0-444-63234-0.50153-6
- Sepúlveda, F. D., Cisternas, L. A., and Gálvez, E. D. (2014). The use of global sensitivity analysis for improving processes: Applications to mineral processing. *Computers and Chemical Engineering*, 66, 221–232. doi:10.1016/j.compchemeng.2014.01.008
- Sitorus, F., and Brito-Parada, P. R. (2020). Equipment selection in mineral processing - A sensitivity analysis approach for a fuzzy multiple criteria decision making model. *Minerals Engineering*, 150(February), 106261. doi:10.1016/j.mineng.2020.106261
- Vandy, C. (2016). *VandyChris/Global-sensitivity- analysis: Python and MATLAB codes to compute the Sobol' indices*. <https://github.com/VandyChris/Global-Sensitivity-Analysis.git>
- Whitworth, A. J., Forbes, E., Verster, I., Jokovic, V., Awatey, B., and Parbhakar-Fox, A. (2022). Review on advances in mineral processing technologies suitable for critical metal recovery from mining and processing wastes. *Cleaner Engineering and Technology*, 7, 100451. doi:10.1016/j.clet.2022.100451

# Optimizing annual-coupled industrial energy systems with sequential time dependencies in a two-stage algorithm

Marion Powilleit\*. Stefan Kirschbaum\*\*. Joram Wasserfall\*\*\*. Clemens Felsmann\*\*\*\*

\* *Society for the Advancement of Applied Computer Science (GFaI), Berlin, Germany (+49-30-814-563-527; [powilleit@gfai.de](mailto:powilleit@gfai.de)).*

\*\* *Society for the Advancement of Applied Computer Science, Berlin, Germany ([kirschbaum@gfai.de](mailto:kirschbaum@gfai.de))*

\*\*\* *Society for the Advancement of Applied Computer Science, Berlin, Germany ([wasserfall@gfai.de](mailto:wasserfall@gfai.de))*

\*\*\*\* *Institute of Power Engineering, Dresden University of Technology, Dresden, Germany ([clemens.felsmann@tu-dresden.de](mailto:clemens.felsmann@tu-dresden.de))*

---

**Abstract:** Our paper presents a two-stage algorithm designed to address year-round-coupled optimization problems encountered in energy system optimization, particularly relevant for scenarios involving seasonal storages or other conditions depending on annual integrals. We apply this algorithm to MILP and MIQCP models. The solution we propose aims to stay feasible with the original problem while getting close to optimal results. It also significantly reduces the computing time compared to solving the original problem alone. This is crucial because the original problem, when coupled, is very complex and sometimes impossible to solve.

*Keywords:* Energy system optimization, seasonal storage, year-round coupling, MILP, MIQCP

---

## 1. INTRODUCTION

For the best design and operation of energy systems, mathematical optimization methods are a well-established tool to increase efficiency, minimize costs, manage capacity restrictions/availabilities, and reduce ecological impact. In the context of energy systems for, e.g., industrial or municipal energy supply or the manufacturing sector, Mixed-Integer Linear Programming (MILP) represents the state of the art for achieving fast and satisfactory results. However, for individual problems or research purposes, nonlinear algorithms are also commonly applied. Especially the inclusion of bilinear terms can help in dealing with problems of temperature-dependence. It augments the problem to Mixed-Integer-Quadratically-Constrained Programming (MIQCP). For optimal design and operation of heat pump and heat storage systems, previous publications (Hering et al., 2021; Hering et al., 2022) and our earlier work (Powilleit et al., 2019; Wasserfall et al., 2019) worked with this subset of nonlinear optimization models.

Energy system optimization is usually applied to quasi-stationary models and therefore may not suffice for certain problems. One example of this is design optimization, which involves the optimal dimensioning of plants. In this scenario, the entire year needs to be considered as one coupled problem. Due to computational constraints, this can quickly become challenging for moderately complex systems. This challenge is addressed by calculating with aggregated time series (e.g., clustering time points or typical days), where the reduced model still yields highly accurate results (Bahl et al., 2017). More details and ideas about optimal time series aggregation give the publications of Hoffmann et al. (2020; 2022). An important aspect is that the sequence of the time steps is not kept in order with clustering aggregation techniques.

In addition to the question of optimal plant size, other circumstances also necessitate the coupling of all time steps into a single optimization problem. These include:

- optimizing the loading of seasonal storages,
- integrating CO<sub>2</sub> emission limitations over the course of a year, and
- considering specific pricing models or regulatory conditions (such as CHP-remuneration).

The integration of seasonal storages was addressed by Kotzur et al. (2018b) through the inclusion of typical charging and discharging days into their aggregation method. Baumgärtner et al. (2020) proposed an approach which decomposes the original problem into smaller subproblems. Kirschbaum et al. (2023) introduced an adapted rolling horizon technique with integer relaxation.

In this work, we propose a two-stage algorithm similar to methods used for structural optimization. The first stage with downsampling and relaxation methods is simplified while maintaining the sequence of the time steps. The second stage is time-resolved in full detail. The crucial aspects include formulating the resulting boundary conditions and selecting the result variables to be transferred to the second stage.

The second stage has to be feasible to the original problem while closely approaching optimality. Our aim is to develop a seamless algorithm capable of handling a broad range of models across both stages, making it suitable for engineering applications. This entails automating the aggregation, transfer, and formulation of the boundaries in the time-resolved stage. The algorithm should solve the problem faster than the original problem.

## 2. METHODOLOGY

The methods presented in this work apply to energy system optimization models (ESOM) with the characteristics of being quasi-stationary and equipped with hourly resolved annual demand time series and containing 5 to 15 technical components. Each time step is formulated with a mixed-integer objective function  $J$  to minimize operating costs while up to  $k$  constraints (e.g., energy balances, conversion terms, operational conditions) can either be mixed integer (MILP, if  $C_i = 0$ ) or in addition bilinear (MIQCP). The continuous optimizing variables  $x$  are usually converted power or energy flows. Integer variables  $y$  are used to describe the minimum part load or piecewise linear part load behavior.

$$\min J(x, y) = a^T \cdot x + b^T \cdot y \quad (1)$$

$$\text{s.t. } x^T \cdot C_i \cdot x + d_i^T \cdot x + e_i^T \cdot y \leq b_i \quad (2)$$

$$\text{with } x \in \mathbb{R}^n, y \in \mathbb{Z}^m, i = 1, \dots, k$$

These equations are often formulated for each time step and solved individually (quasi-stationary). However, when time-coupling conditions are necessary, such as optimizing a daily storage, a **rolling horizon** method is employed (attributable to Bellman (2021)). In this approach, multiple future time steps are linked within a control horizon (e.g., 48 h) to determine the optimal plant behavior. From this control horizon, only the first result is retained, and the process is repeated for the next step then. However, for saving computation time, it is also very common to keep a subset of the control horizon in a shorter write-back horizon (e.g., the first 4 h). The horizon continuously shifts forward throughout the year.

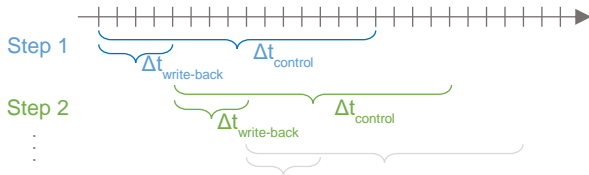


Fig. 1. Illustration of a rolling horizon optimization.

However, some issues, such as the integration of seasonal storage units or regulatory constraints, cannot be addressed by breaking the simulation into smaller parts and prevent solving with conventional receding horizon methods. These require a fully coupled optimization process. This often results in very large and complex optimization problems, which can take a long time to compute or might not even reach a solution. To tackle these cases, we propose our two-stage algorithm.

### 2.1 Two-stage algorithm: simplified first stage

In the simplified first stage, all time series  $b_i$  with the original time step width  $\Delta t$  are downsampled into a coarser grid  $\Delta t_j$  using integral-preserving averaging.

$$b_{i,t_j}^{down} = \frac{1}{\Delta t_j} \sum_{t=\Delta t_j \cdot j}^{\Delta t_j \cdot (j+1)} b_{i,t} \cdot \Delta t \quad (3)$$

Additional relaxation of the binary variables is possible: The originally binary variables  $y \in \mathbb{Z}^n$  are defined continuously as  $y \in \mathbb{R}^n$  in the bounds  $[0,1]$  to solve only an LP instead of an

MILP. These simplifications allow to solve the optimization model of a moderately complex problem in which the whole time frame is coupled in one coupled problem.

### 2.2 Two-stage algorithm: choosing transfer variables

In an intermediate step of the algorithm, it must be determined which result variables from the first stage are to be retained. These should be the crucial, year-round variables, such as the filling level of a seasonal storage or the amount of CO<sub>2</sub> emitted. These can be brought back to the original time grid through upsampling. The upsampling is optional because the results of the interpolated time steps are not intended to be used further. However, it can still be useful in special cases. The rest of the results are omitted.

From the results, only the transfer variables are selected, which should become boundary conditions for the second stage.

### 2.3 Two-stage algorithm: Fully-resolved second stage with boundary conditions

In the second stage, the fully-resolved time series and original binary variables are solved with a rolling horizon method. The transfer variables are incorporated as boundary conditions. An intuitive idea is to simply equate the results from each time step from the first stage with the ones from the second, but this leads to infeasibilities and large deviations to optimality. Therefore, we incorporate the boundary conditions into the rolling horizon.

- For integral variables (storage level, amount of CO<sub>2</sub>), in each MILP to solve, we only equate the variable  $x^*$  to the results from the first stage  $x^{*,first\ stage}$  at the last time step of the control horizon:

$$x_{t=\Delta t_{control}}^* = x_{\Delta t_{control}}^{*,first\ stage} \quad (4)$$

- For regular variables (e.g., the electrical power of a CHP), we equate the integral of all variables within the control horizon:

$$\sum_{t=i}^{\Delta t_{control}} x_i^* \cdot \Delta t = \sum_{t=j}^{\Delta t_{control}} x_j^{*,first\ stage} \cdot \Delta t_j \quad (5)$$

- We optionally apply tolerances  $\varepsilon$  to equation (4) and analogously to (5):

$$x_{t=\Delta t_{control}}^* \geq (1 - \varepsilon) \cdot x_{\Delta t_{control}}^{*,first\ stage} \quad (6)$$

$$x_{t=\Delta t_{control}}^* \leq (1 + \varepsilon) \cdot x_{\Delta t_{control}}^{*,first\ stage} \quad (7)$$

- We discovered that it is very beneficial to solvability to set the control horizon to a multiple  $n$  of the prior downsampling rate  $\Delta t_j$  and write the results back in the same length as the downsampling rate.

$$\Delta t_{control} = \Delta t_j \cdot n, \quad n \in \mathbb{Z} \quad (8)$$

$$\Delta t_{write-back} = \Delta t_j \quad (9)$$

This way, the optimality loss accepted in the first stage due to simplification is balanced out to reach approximate optimality. The selection of the values for the downsampling rate, the rolling horizon frame, and the tolerance varied are discussed in Section 3.

2.1 Demonstration of algorithm implementation

Figure 2 shows the results of the algorithm in anticipation of the third case study, the optimal filling level of a seasonal storage. The final optimum filling level of the seasonal storage is marked in green. The grey line is the result of the first downsampled stage, with only a very coarse resolution. The comparison of the two plots shows the progress: In the left one, the red line describes the result of one MILP in the rolling horizon. It starts at  $t_0$ . The filling level in the horizon meets the one from the first step (grey line) at  $t_{horizon,0}$ , which is exactly the defined boundary condition. In the second plot, the consideration starts after the write-back frame at  $t_4$ . From there, an MILP with the actualized control horizon is solved and meets the boundary condition again at  $t_{horizon,4}$ .

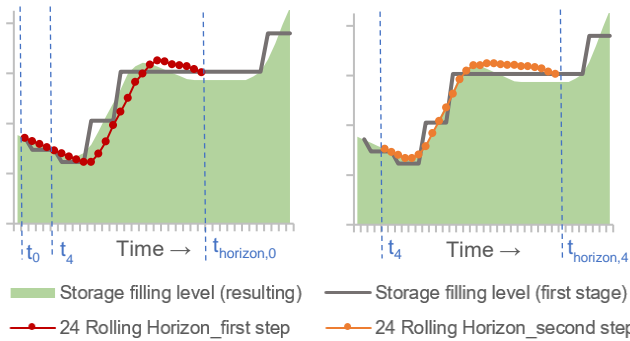


Fig. 2. Exemplary illustration to demonstrate the algorithm: first time step (left), second time step (right).

3. OPTIMIZATION RESULTS

The methodology applies to four different use cases, each of it considering different aspects. For all use cases,

- the operating costs are minimized,
- the demand profiles are extracted and adapted from real use cases or taken from published typical days and have a resolution of 8760 h, and
- energy supply from grid is always depicted with realistic prices.

Because the simulations contain many different combinations of parameters, the nomenclature for the results is standardized as follows.

Table 1: Nomenclature of result diagrams

<span style="color: red;">●</span>	The big red dot marks the reference case (if solvable): the complete solution of the annual-coupled fully-resolved problem.
<span style="color: green;">◆</span> <span style="color: blue;">◆</span>	Colored and shaped group of data points share the same simplified stage model. The legend gives the downsampling rate (“Down 4”) and adds binary relaxation if applied (“Relax”).
<span style="border: 1px solid green; padding: 2px;">48/3 - Tol 0,1</span>	The data labels mark the boundary conditions in the second stage with the control horizon of data to write back in hours (“48/3”) and the applied tolerances $\epsilon$ in % (“Tol 0,1”).
<span style="border: 1px solid red; border-radius: 50%; padding: 2px;">●</span>	The red circle marks simulations with an incomplete result set: For some time steps, no solution could be found (within the time limit).

3.1 Software in use

The models are built with the framework *EnergyFrames*, using libraries from the derivate *TOP-Energy*<sup>®</sup>. Both are proprietary in-house developments by the GFaI. As an optimization solver, Gurobi 11.0.1 is used. While striving for comparability, we calculated with a gap value of 0. However, for some simulations this was not possible in reasonable time. For illustrative purposes, we present computational times in the results, but internally validated them using the dimensionless measurement of Gurobi’s work-units.

3.2 CASE 1: OPERATIONAL OPTIMIZATION DUE TO REGULATORY RESTRICTIONS

Case 1 is a combined heating, cooling, and power system (Fig. 4), as found in small industrial systems. It has certain degrees of freedom: The heating demand of ~1 GWh/a can be met by a CHP (Combined Heat and Power system), a heat pump, or a boiler. The cooling demand of ~0.6 GWh/a can be met by an electric compression chiller or a heat-driven adsorption chiller. The electric grid can either supply power to the plants and other demands of ~1 GWh/a or receive fed-in electricity from the CHP.

The objective function is to minimize the operation costs while complying with one condition: Due to reduction goals of the operator, the whole system should not emit more than 295 t CO<sub>2</sub> per year.

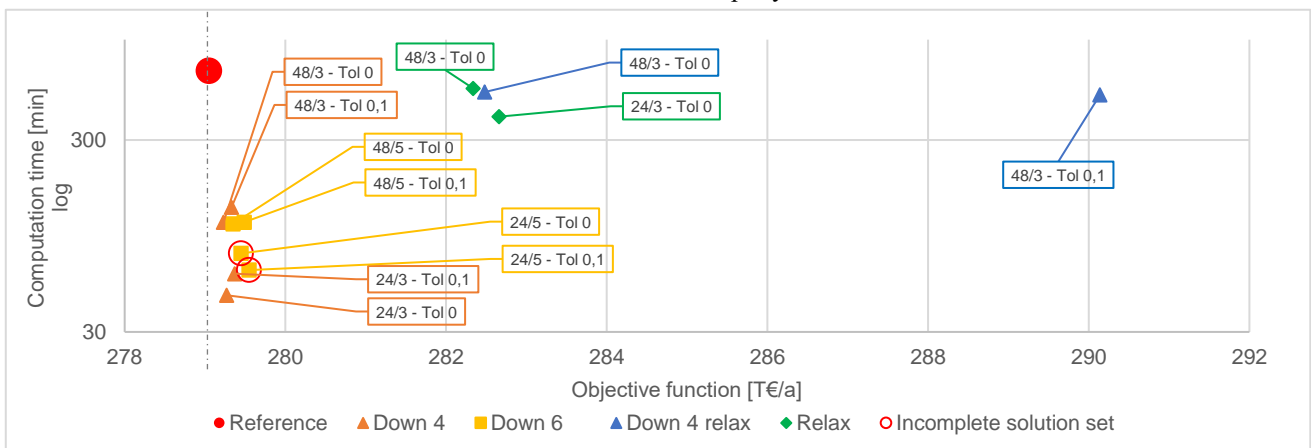


Fig. 3. Computation time vs. objective function results for the two-stage algorithm with varying parameters (Case 1).



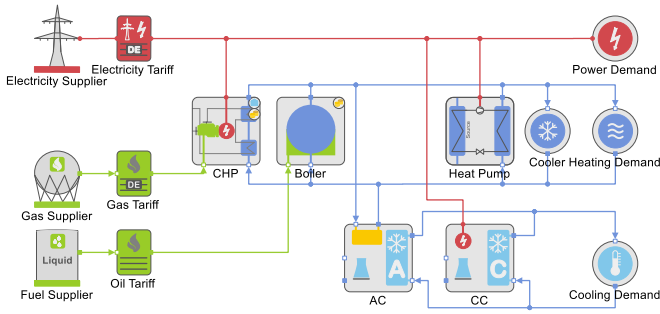


Fig. 4. Scheme of the ESOM with combined heating, cooling and power, with CO<sub>2</sub> emission restriction (Case 1).

Figure 3 shows the results of the two-stage algorithm for different parameters of downsampling rate, tolerances, and horizon frame. The calculation times are compared with each other depending on the objective function values. The aim is to minimize the calculation time with the objective function being as close as possible to the reference solution.

The transfer condition is the integral value of all emissions in each time step. This marks the boundary condition for the second stage.

The main results can be summarized as following.

- Best results can be reached with a downsampling of 4 h and a rolling horizon of 24 h because they require low computation time and still offer a near-optimal objective function.
- The influence of the tolerances in the boundary conditions is unambiguous, but usually, zero tolerance reduces calculation time because no additional degree of freedom is created.
- Binary relaxation of the first step does not show good results (▲ and ◆): Even if the first step solves much faster (~8 s instead of 1670 s), solving the individual steps of the second stage drastically augments the total calculation time and does not even find a feasible solution (○) for each time step.

To understand the reasons for the weak performance of the binary relaxation, we examine the results of the first stage and

look at the cumulated CO<sub>2</sub> emissions. The deviation of the relaxed solution to the reference is five times higher than the one from the downsampling case. To disclose that behavior, we take a deeper look into the results of the relaxed binary variables and where they violate the binary condition. We find that the main issue with the violation concerns the operation of the absorption chiller: The minimum part load of 40 % or 160 kW is violated during winter operations with loads around 15 kW (Fig. 6). The resulting boundary condition for the second stage is challenging to meet because unrealistic behavior stems from the outcomes of the first stage.

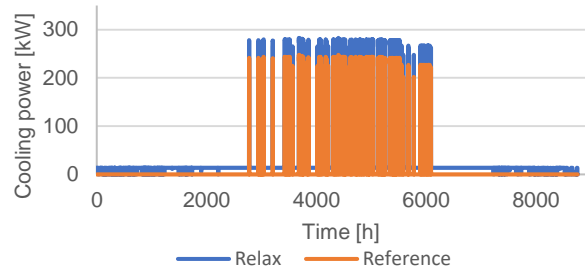


Fig. 6. Comparison of the operation of the AC (Case 1).

### 3.3 CASE 2: OPTIMAL CHP OPERATION WITH ANNUALLY RESTRICTED REMUNERATION

In this example (Fig. 7), we delve into a common question concerning the optimal operation of a small-scale CHP system that must effectively meet heat demands of ~1.2 GWh/a while also maximizing electricity sales. Under German law, remuneration is provided for every kilowatt hour of CHP electricity generated, but only for a total of 3,500 full load hours (FLH) per year. Consequently, it is not feasible to optimize each time step individually. Instead, a comprehensive approach spanning the entire year is necessary to determine the most advantageous times for distributing the load of the CHP system.

The CHP is represented with a part load behavior, while a gas boiler and an emergency cooler serve as additional degrees of freedom within the heating grid. Operating costs are minimized as an objective function.

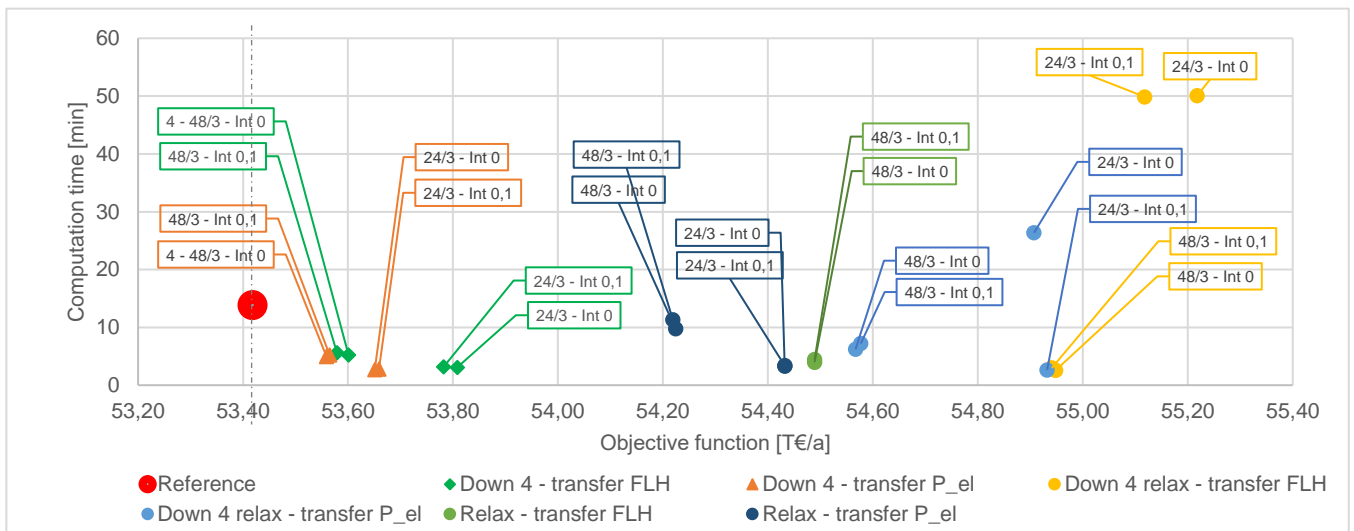


Fig. 5. Computation time vs. objective function results of the two-stage algorithm with varying parameters (Case 2).

The first stage is treated with the downsampling rate of 4 h and optional binary relaxation. In this case, we investigate the quality of two different transfer variables: First, the cumulated FLH (integral variable) is used as a transfer variable for the boundary condition. Unlike the previous case, the time-resolved integral is not necessary because a constraining sum is enough. This allows the remuneration of the FLH to be limited in the first stage and the electricity generation of the CHP to be used as a transfer variable as a second option.

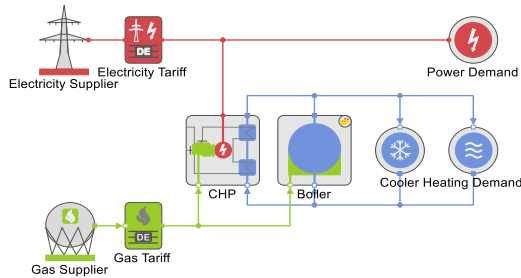


Fig. 7. Scheme of the ESOM with a CHP and annual full load hour restriction of remuneration (Case 2).

The main aspects of the results are:

- As in the previous case, the binary relaxed cases do not improve the total computation time for related reasons as Case 1.
- The other solutions are close to the objective function and reduce the calculation time by around 70 %.
- A larger rolling horizon frame serves a better objective function but increases the calculation time.
- Using the electric power as a transfer variable instead of the cumulated FLH seems to be a better option.
- The tolerance's influence is less clear than in Case 1.

### 3.4 CASE 3: OPTIMAL LOADING STRATEGY OF A SEASONAL HYDROGEN STORAGE

At the core of the fourth use case is a seasonal hydrogen storage system, which serves as a year-round coupling element to increase the use of renewable energy (Fig. 9). During the summer, 1.8 MW<sub>p</sub> Photovoltaic (PV) and electrolysis can

charge it, while a fuel cell can reconvert the stored energy into electricity to partially cover a demand of ~1.5 MWh/a. A small daily electricity storage system can compensate for short-term fluctuations. In combination with a heat pump, the waste heat from the fuel cell and electrolysis can meet the ~1 GWh/a heat demand, which otherwise would be satisfied by a natural gas boiler.

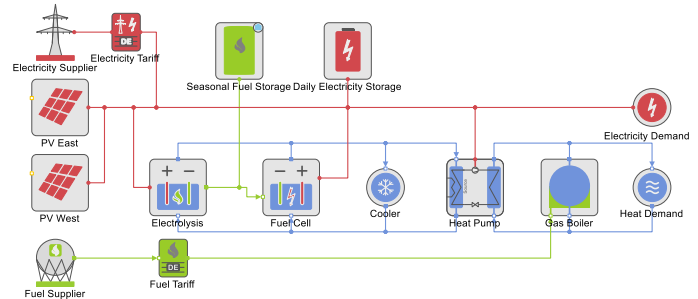


Fig. 9. Scheme of the ESOM with a seasonal storage (Case 3).

As before, the first stage is simplified by downsampling and optional binary relaxation. The transfer condition is initially intuitively the energy level of the seasonal storage (in an integral variable). However, an alternative approach is to define the charging and discharging capacity of the seasonal storage system as a transfer variable. Summarizing the results:

- Good results are achieved by transferring the filling level, employing a downsampling rate of 4 and a rolling horizon of 24 hours. The objective function closely approximates the reference case, while computation times decrease by 20 %.
- Selecting a rolling horizon frame that is too small is problematic because of increased computation times, reduced optimality, and infeasibilities at some time steps.
- Similar to previous cases, the binary relaxation method yields unsatisfying results: With computation time intervals between 200 and 1,800 seconds (not all plotted), the second stage requires too much time for the solution. In addition, it cannot find a solution for

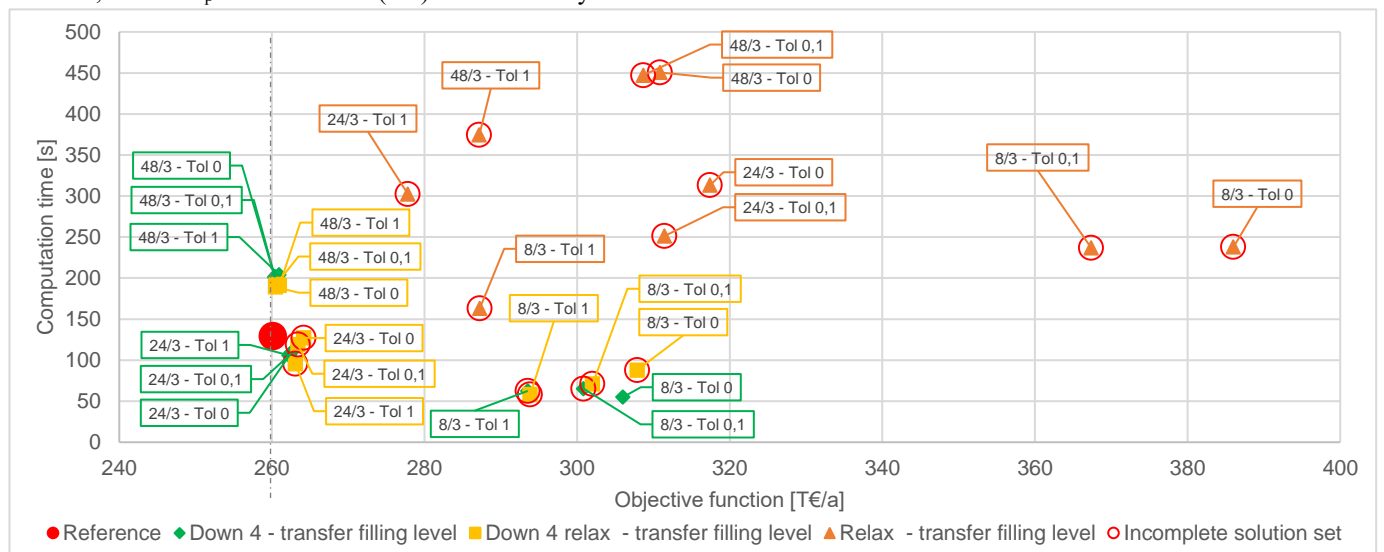


Fig. 8. Computation time vs. objective function results for the two-stage algorithm with varying parameters (Case 3).

all time steps in the maximum time allowed for solving. This leads to an incomplete solution set.

- Coupling the charging and discharging power instead of the integral variable yields highly unfavorable outcomes: Calculation times extend up to 13 hours, consistently resulting in an incomplete solution set. For reasons of clarity, these results are omitted from the plot. It is advisable to refrain from this formulation. Because the integral variable must be computed temporally resolved, no advantage to the previous case study can be expected in the first stage.
- The calculation times of the two-stage algorithm in the downsampling scenario can even exceed those of the reference case (previously observed only when binary relaxation is applied). The solution of individual time steps within a rolling horizon becomes significantly more complex in this scenario. In addition to the coupling condition, the inclusion of a daily storage further enlarges the problem size. Therefore, it is important to carefully choose the horizon frame: To avoid unnecessary complexity, it must not be too large, but must be long enough to effectively operate the daily storage.

3.5 CASE 4: OPTIMAL INTEGRATION OF A SEASONAL HEAT STORAGE TO SUPPLY HEATING OR COOLING COMPRESSION

The fourth model represents a simplified approach to a heating and cooling supply as might be implemented in municipal heat planning (Fig. 11). The core of this model is a 190 MWh seasonal heat storage that stores water at variable temperatures. On the one hand, heat can be extracted as drive heat for an electric heat pump, which raises the temperature level to satisfy the heat demand. As a degree of freedom, a gas-driven boiler can also cover the heat demand. On the other hand, a compression chiller that satisfies a cooling demand can regenerate the storage with its exhaust heat. An additional cooling unit can alternatively fulfill the cooling demand.

The COPs of both plants vary with their inlet temperature, which is reflected as the storage temperature itself. The heat demand is very high in winter, whereas summer season is dominated by cooling demand. The question is how to control the storage temperature throughout the year.

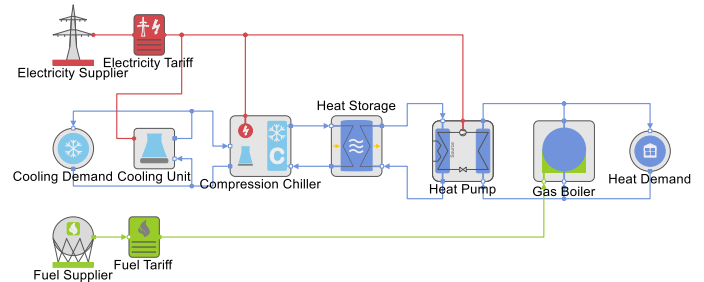


Fig. 11. Scheme of the ESOM with a seasonal heat storage and temperature dependencies (Case 4).

Even if the model is kept simple, the temperature-dependent COP introduces bilinearity into the problem and leads to a non-convex nonlinear optimization problem. That significantly increases its complexity and computational time. The most important bilinear equations are caused by the  $i$  enthalpy flows in every energy balance (with  $c_p$  being constant) and each of the both COP-dependencies with  $\alpha$  being a constant temperature correction factor. Storage mass  $m$  and heat capacity  $c_p$  are also assumed to be constant.

$$\dot{H}_i = \dot{m}_i \cdot c_p \cdot \Delta T_i \tag{10}$$

$$\Delta U = m_{storage} \cdot c_p \cdot (T_{storage} - T_{ref}) \tag{11}$$

$$\Delta U = \Delta \dot{H}_{charging} + \Delta \dot{H}_{discharging} \tag{12}$$

$$P_{el,j} = COP_j \cdot \Delta \dot{H}_j \tag{13}$$

$$COP_j = COP_{j,nom} + \alpha \cdot (T_{in,j} - T_{in,nom}) \tag{14}$$

In contrary to the other cases, even a coarse downsampling up to 168 h cannot solve the simplified stage to gap zero, but remains at gap values between 5 and 8 %. A quite reasonable gap of around 10 % is reached even in short calculation times, e.g., of 10 minutes.

Even if we could not generate a reference solution from the year-coupled original problem, the algorithm can still compare the different boundary conditions concerning the absolute value of the objective function. Figure 10 gives an overview of the results and the main aspects are summarized:

- With a downsampling of 24 h and a rolling horizon of 48 or 72 h, calculation times of less than 100 min can

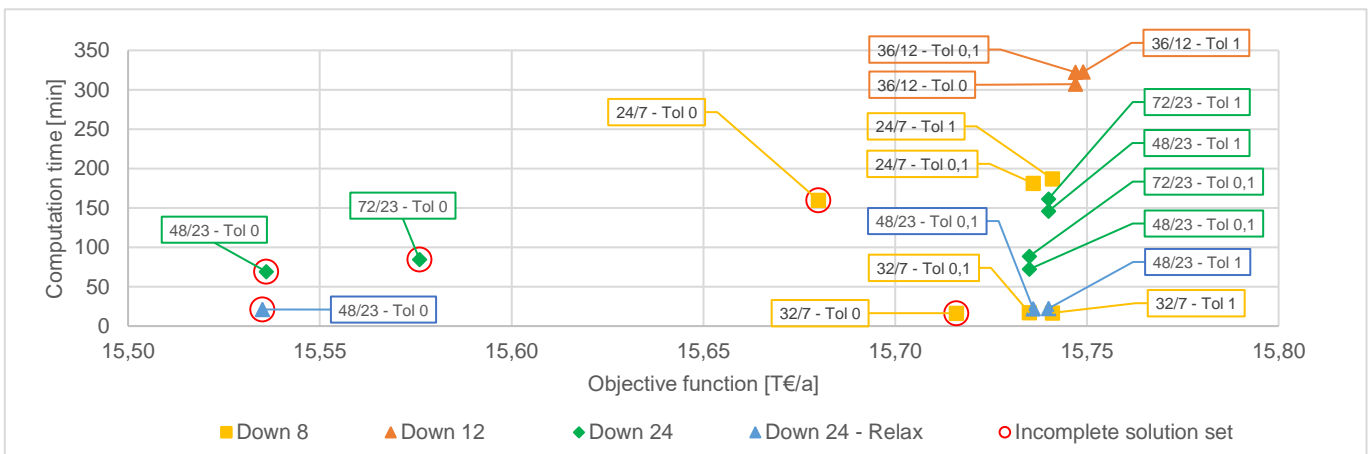


Fig. 10. Computation time vs. objective function results for the two-stage algorithm with varying parameters (Case 4).

be reached to find a feasible solution. The shortest calculation times are reached with a downsampling of 8 h, and a rolling horizon of 32 h. For example, part loads were not modeled in this simple model.

- The objective function values and filling levels do not differ significantly.
- In this case, small tolerances are recommended for the transfer variables in the boundary conditions because most cases with 0 tolerance include time steps for which no solution was found.
- Although in this case the binary relaxation provides advantages, we would not recommend it respecting the experiences of the other use cases.

Even if the model is not very sophisticated (few plants, no heat transfer laws, no part load behavior or runtime conditions), the MIQCP takes much more computation time due to the temperature dependency.

The optimal plant behavior is not only influenced by the energy prices but subsequently also by the temperature-dependent COPs. Three differently efficient heat pumps are compared for illustration (the compression chiller remains unchanged). The algorithm is applied to each heat pump configuration. Figure 12 shows the results of the storage temperatures: The higher the nominal COP, the colder the storage gets in the summer. The highly efficient heat pump can operate economically even with a low driving temperature, whereas the low efficient heat pump requires the storage to keep a higher temperature and hence lowers the full load hours of both compression chiller and heat pump by about 19 % resp. 26 %.

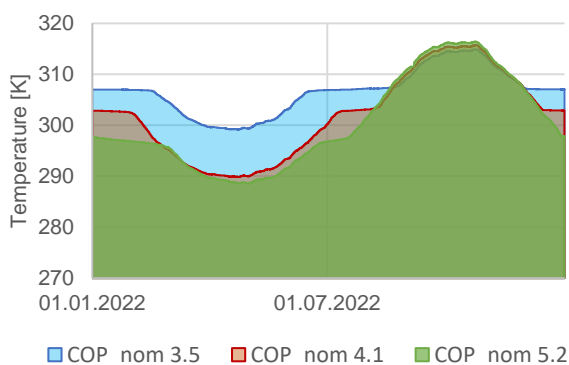


Fig. 12. Influence of nominal COPs on the seasonal storage temperature (Case 4).

#### 4. RESULTS AND DISCUSSION

For all use cases, the two-stage algorithm demonstrates promising results, yielding solutions close to optimal while also achieving significant reductions in computation time (from 15 to 90 %). In instances where the reference solution fails to solve entirely (e.g., due to memory constraints with Gurobi), a feasible solution may still be attained, although without certainty regarding its proximity to the optimum.

Generally, the selection of downsampling rates and horizon windows requires careful consideration: while coarse downsampling accelerates the solution of the first stage,

ensuring compliance with boundary conditions in the second stage may require more time. An overview about the achieved computation time saving vs. objective function losses shows Figure 13, where two of the most appropriate parameter settings were chosen. With a too high resolution, the computation time for Case 3 with the seasonal storage in combination with a daily storage may exceed the reference time. Given the absence of a reference for the MIQCP seasonal storage in Case 4, a direct comparison is not possible. However, obtaining a plausible and feasible solution in approximately 16 minutes is a promising outcome. The consistency of the similar storage charging results across different algorithm settings is also promising.

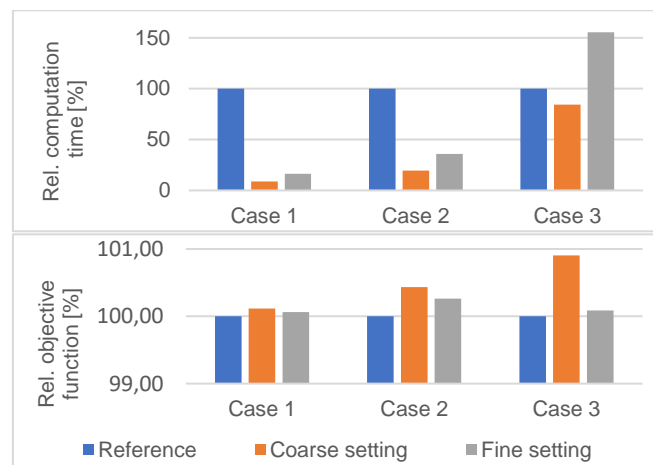


Fig. 13. Comparison of effectiveness of the two-stage algorithm.

Binary relaxation yields unsatisfactory results in this study. While the first stage solves rapidly even without downsampling, the second stage entails long computation times. It is important to note that we did not explore context-specific relaxation methods. There are likely opportunities for improvement by treating specific descriptions for which binary variables are intended differently (e.g., piecewise linear characteristics, minimal part loads or to flow direction decisions), for more details see Özbek (2022). In cases where downsampling is impossible (e.g., due to critical importance of peak power capabilities or peak power prices), revisiting this method may be warranted.

In this work, a highly capable commercial solver was utilized, allowing even the reference solution to be solved in a reasonable amount of time. However, if using open-source solvers, the parameters found here would need to be adjusted. The pure solution times were taken into account in this study. When utilized within a comprehensive program, other factors such as data handling and LP creation will inevitably come into play. Moreover, it is worth noting the advantage of managing smaller MILP files, particularly in scenarios where memory resources may be limited.

The downsampling process itself can be reconsidered: Currently, equidistant integral averaging is applied, but a more intelligent segmentation of the time series could be implemented, focusing on a higher resolution at important data points while preserving the chronological order, as discussed



by Kotzur et al. (2018a). This approach could provide a means to address issues related to peak loads or power pricing better.

## 5. CONCLUSIONS

Though the models considered here are not extensively complex (to allow a comparison with the reference solution), they nevertheless quickly escalate in complexity with the addition of a few more binary variables or constraints. At this point, year-round coupling can become unfeasible, necessitating the application of multi-stage methods.

Our aim was to develop a highly generic method capable of solving a variety of year-round coupled models within acceptable computation times, yielding feasible solutions. While it is possible to fine-tune each individual model with parameters, we believe that these results allow us to offer a general solution for MILP models in this domain. A coarser and faster solution could be achieved with a downsampling rate of 7 h and a rolling horizon of 28 h, whereas a better solution could be achieved with a downsampling rate of 4 h and a rolling horizon of 24 h. However, expressing this generality for the MIQCP case is more challenging: Here, it is advisable to examine the results of the first stage and assess their plausibility with engineering insight. Nonetheless, by doing so, very good results were achieved in this case as well.

In this work, we only addressed operational optimization problems. The method can easily be adapted to questions combined with design optimization as well when considering the aforementioned issue of peak levelling.

The formulation of MILP models is well-established in research and application, but there is room for further refinement in formulating MIQCPs to enhance their efficiency.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the support and funding provided by BMWK, which made the research project “BiliOpt - Optimierung von Energiesystemen unter Verwendung bilinearer Nebenbedingungen” (03EI1066B) possible.

Supported by:



## REFERENCES

- Bahl, B., Kümpel, A., Seele, H., Lampe, M. and Bardow, A. (2017) Time-series aggregation for synthesis problems by bounding error in the objective function. *Energy Energy*, 135, 900–912. doi: 10.1016/j.energy.2017.06.082.
- Baumgärtner, N., Shu, D., Bahl, B., Hennen, M., Hollermann, D.E. and Bardow, A. (2020) DeLoop: Decomposition-based Long-term operational optimization of energy systems with time-coupling constraints. *Energy Energy*, 198, 117272. doi: 10.1016/j.energy.2020.117272.
- Bellman, R.E. (2021) *Dynamic Programming*. Princeton University Press: Princeton, NJ.
- Hering, D., Faller, M.R., Xhonneux, A. and Müller, D. (2022) Operational optimization of a 4th generation district heating network with mixed integer quadratically constrained programming. *Energy Energy*, 250, 123766. doi: 10.1016/j.energy.2022.123766.
- Hering, D., Xhonneux, A. and Müller, D. (2021) Design optimization of a heating network with multiple heat pumps using mixed integer quadratically constrained programming. *Energy Energy*, 226, 120384. doi: 10.1016/j.energy.2021.120384.
- Hoffmann, M., Kotzur, L. and Stolten, D. (2022) The Pareto-optimal temporal aggregation of energy system models. *Applied Energy*, 315, 119029. doi: 10.1016/j.apenergy.2022.119029.
- Hoffmann, M., Kotzur, L., Stolten, D. and Robinius, M. (2020) A Review on Time Series Aggregation Methods for Energy System Models. *Energies*, 13(3), 641. doi: 10.3390/en13030641.
- Kirschbaum, S., Powilleit, M., Schotte, M. and Özbeg, F. (2023) Efficient Solving of Time-Coupled Energy System MILP Models Using a Problem Specific LP Relaxation. In: ECOS 2023 (Ed.) *36th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems*, 25-30 June, Las Palmas De Gran Canaria, Spain. ECOS 2023, pp. 2774–2785.
- Kotzur, L., Markewitz, P., Robinius, M. and Stolten, D. (2018a) Impact of different time series aggregation methods on optimal energy system design. *Renewable Energy*, 117, 474–487. doi: 10.1016/j.renene.2017.10.017.
- Kotzur, L., Markewitz, P., Robinius, M. and Stolten, D. (2018b) Time series aggregation for energy system design: Modeling seasonal storage. *Applied Energy*, 213, 123–135. doi: 10.1016/j.apenergy.2018.01.023.
- Özbeg, F. (2022) *Entwicklung einer problembezogenen Relaxierung zur effizienten Beschreibung von MILP-Modellen für Energiesysteme*. Bachelorarbeit, Berlin, TU Berlin.
- Powilleit, M., Wasserfall, J., Kirschbaum, S., Tretau, A. and Steinmann, D. (2019) Energy system optimization with short term storages respecting nonlinear temperature dependencies. In: Stanek, W. (Ed.) *32nd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2019)*, 23-28 June, Wrocław, Poland. Silesian University of Technology, Department of Thermal Engineering.
- Wasserfall, J., Powilleit, M., Kirschbaum, S. and Wrobel, G. (2019) Efficient formulation for optimizing temperature dependent energy systems with mixed-integer quadratically constrained programming. In: Stanek, W. (Ed.) *32nd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2019)*, 23-28 June, Wrocław, Poland. Silesian University of Technology, Department of Thermal Engineering, pp. 953–965.



# Numerical Methods for the Flow Fields; A Comparative Review

Jamshid Moradi, Amin Mahmoudzadeh Andwari, Juho Könnö

*Machine and Vehicle Design (MVD), Materials and Mechanical Engineering, Faculty of Technology,  
University of Oulu, FI-90014 Oulu, Finland*

---

**Abstract:** This paper provides a comparative overview of four numerical methods widely employed in computational fluid dynamics and related fields: Finite Volume (FV), Lattice Boltzmann Method (LBM), Smoothed Particle Hydrodynamics (SPH), and Spectral Methods. FV discretizes the domain into control volumes, emphasizing conservation laws and flux integrals across cell faces. It's renowned for its robustness, particularly in complex geometries. LBM is a mesoscopic approach simulating fluid dynamics through particle interactions on a lattice grid. Its intrinsic parallelism and ability to handle complex boundary conditions make it suitable for multiphase flows and porous media simulations. SPH represents fluids as a set of particles, where properties are smoothed over neighboring particles using a kernel function. SPH excels in free surface flows, astrophysical simulations, and fluid-structure interaction due to its Lagrangian nature and adaptive resolution. Spectral Methods discretize functions using orthogonal basis functions, such as Fourier or Chebyshev polynomials, enabling high-order accuracy and spectral convergence. They are preferred for problems with smooth solutions and periodic boundary conditions, like turbulence simulations and wave propagation.

*Keywords:* Fluid Flow Simulation, Mesh-free Methods, Grid-based Methods, Particle-Based Methods, High-Resolution Simulation, Parallel Scalability

---

## 1. INTRODUCTION

The realm of Computational Fluid Dynamics (CFD) stands as a revolutionary force in comprehending and engineering fluid flow phenomena. It provides a suite of numerical techniques adept at simulating and dissecting complex fluid behaviors. Amid this toolkit, four prominent methods emerge: Finite Volume (FV), Lattice Boltzmann Method (LBM), Smoothed Particle Hydrodynamics (SPH), and Spectral Methods. Each method offers distinct advantages, tailored to address various fluid flow scenarios, thus becoming indispensable across scientific and engineering disciplines.

This paper endeavors to furnish a succinct yet comprehensive overview and comparative analysis of these four numerical methodologies within the domain of fluid dynamics simulations. By delving into their fundamental principles, strengths, and limitations, this review aims to serve as a compass for researchers and practitioners in selecting the most suitable numerical technique for their specific applications.

The evolution of fluid dynamics study has been marked by a surge in numerical methods, propelled by advancements in computational power and algorithmic sophistication. FV methods, renowned for their robustness and adaptability in handling intricate geometries, have garnered widespread acceptance. Foundational works by Patankar (Barth, Herbin and Ohlberger, 2017) laid the groundwork for applying FV methods to fluid flow simulations. Subsequent refinements by Versteeg and Malalasekera (Barth, Herbin and Ohlberger, 2017) and Ferziger and Peric (Boudet, 2011) expanded the method's applicability across diverse engineering domains. FV

method operates by dividing the computational domain into small, non-overlapping control volumes. This method transforms the integral forms of conservation laws—governing the fluid's mass, momentum, and energy—into algebraic equations over these discrete volumes. By computing the fluxes of conserved quantities across the boundaries of each control volume, the FV method ensures that any flux leaving one volume enters the adjacent one, inherently conserving the quantities. This intrinsic conservation property, coupled with its flexibility in handling complex geometries and boundary conditions, makes the FV method particularly powerful and reliable for simulating a wide range of fluid flow problems, from aerodynamics in aerospace engineering to pollutant dispersion in environmental studies. The meticulous nature of flux calculation and interpolation at the boundaries, while computationally demanding, ensures the fidelity and accuracy of the simulations, thereby providing invaluable insights into the fluid behavior under various physical scenarios (Boudet, 2011). Figure 1 illustrates the Finite Volume Method (FVM), where the computational domain is discretized into small control volumes. The method computes fluxes of conserved quantities such as mass, momentum, and energy across the boundaries of each control volume, ensuring conservation laws are satisfied. A schematic diagram showing a 2D computational grid with control volumes. Each control volume is represented by a small square, with arrows indicating the fluxes of mass, momentum, and energy across the faces of the control volumes.

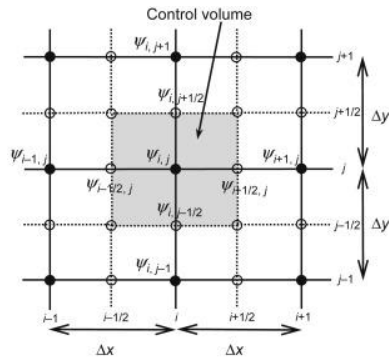


Fig. 1. Schematic of the Finite Volume Method (FVM) showing control volumes

LBM has emerged as a formidable contender to traditional Navier-Stokes solvers, particularly excelling in simulating multiphase flows and intricate boundary conditions. Succi's seminal treatise (Samanta, Chattopadhyay and Guha, 2022) provided an exhaustive introduction to LBM principles, igniting widespread interest and further advancements in the field. Recent endeavors by Aidun and Clausen (Aliu *et al.*, 2020) and Chen *et al.* (Chen *et al.*, 2020) have broadened LBM's horizon into novel territories like microfluidics and porous media flow. LBM revolutionizes fluid dynamics simulation by bridging microscopic and macroscopic scales through kinetic theory. Instead of solving the traditional Navier-Stokes equations directly, LBM models fluid flow by simulating the evolution of particle distribution functions on a discrete lattice grid. At each lattice node, particles propagate and collide according to simplified rules derived from the Boltzmann equation, with the post-collision distributions relaxed towards equilibrium states. This inherently local computation facilitates easy parallelization, making LBM computationally efficient and adaptable to modern high-performance computing architectures. The method excels in handling complex boundary conditions and interfaces, such as those found in porous media or multiphase flows, by naturally accommodating microscopic interactions and capturing emergent macroscopic behavior. As a result, LBM has found applications in diverse fields ranging from microfluidics and biomedical engineering to materials science, offering a robust, flexible, and scalable tool for exploring the intricacies of fluid dynamics in complex systems (Aliu *et al.*, 2020). Figure 2 depicts the Lattice Boltzmann Method (LBM), where fluid dynamics are simulated on a discrete lattice grid. At each lattice node, particles propagate and collide, following simplified rules that are derived from the Boltzmann equation. This method efficiently handles complex boundary conditions and interfaces. A lattice grid with nodes connected by velocity vectors. At each node, particle distribution functions are shown, along with arrows representing the streaming and collision processes.

SPH has ascended in prominence for simulating free surface flows, fluid-structure interactions, and celestial phenomena. The pioneering work of Gingold and Monaghan laid the foundation for SPH, which has since undergone refinement and widespread application across many problems. Monaghan and Kocharyan's comprehensive review (Bagheri, Mohammadi and Riazi, 2023) shed light on SPH techniques,

while subsequent breakthroughs by Price and Monaghan (Rosswog, 2020) and Rosswog (Rai and Mondal, 2021) expanded their capabilities in modeling intricate fluid dynamics scenarios.

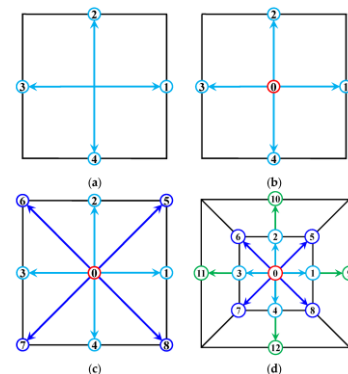


Fig. 2. Representation of the Lattice Boltzmann Method (LBM) with a D2Q9 lattice grid

SPH transforms fluid dynamics into an elegant dance of particles, free from the constraints of traditional grids. In this mesh-free Lagrangian method, the fluid is represented by discrete particles, each carrying properties such as mass, velocity, and density. These properties are smoothed over a finite distance using kernel functions, allowing for the accurate interpolation of fluid variables across the particles. As these particles move and interact, they capture the essence of fluid behavior, from subtle ripples to violent splashes, making SPH particularly adept at handling complex, transient phenomena like free surface flows, multiphase interactions, and large deformations. This flexibility extends to naturally managing moving boundaries and interfaces, which are often challenging for grid-based methods. Originating in astrophysics to model stellar phenomena, SPH has found its way into diverse applications, including oceanography, biomechanics, and industrial processes, where its ability to simulate realistic and intricate fluid motions in a computationally efficient manner makes it an indispensable tool for scientists and engineers delving into the dynamic world of fluid flows (Bagheri, Mohammadi and Riazi, 2023). As shown in Fig. 3, the Smoothed Particle Hydrodynamics (SPH) method represents the fluid as a set of discrete particles. The properties of the fluid, such as density and velocity, are interpolated using smoothing kernels, enabling the simulation of complex free-surface flows. Figure 3 provides a visualization of fluid particles within a domain, where particles are depicted with overlapping smoothing kernels, illustrating how properties are interpolated between neighboring particles.

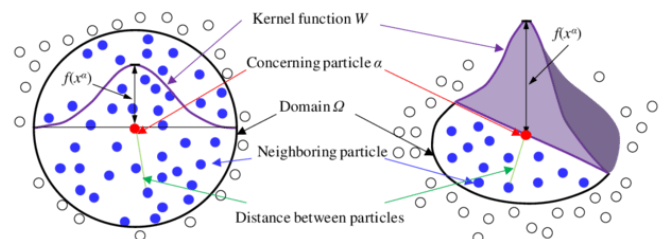


Fig. 3. Smoothed Particle Hydrodynamics (SPH) approach

Spectral Methods, rooted in mathematical and numerical analysis, offer high-order accuracy and spectral convergence properties, making them apt for problems with smooth solutions and periodic boundary conditions. Seminal contributions by Orszag (Rai and Mondal, 2021) and Canuto et al. (Caban and Tyliczszak, 2022) established the theoretical underpinnings of spectral methods, while Trefethen's exposition (Vishwanatha *et al.*, 2023) provided a contemporary synthesis of spectral techniques in fluid dynamics. Recent strides by Boyd and Fornberg (Prasad, Choi and Patil, 2022) have extended spectral methods into uncharted domains, encompassing turbulence modeling and wave propagation. The Spectral Method in computational fluid dynamics is akin to composing a symphony where fluid behavior is captured through the harmonics of global basis functions. By representing the solution of the governing equations, such as the Navier-Stokes equations, with a series of trigonometric (Fourier) or polynomial (Chebyshev) functions, this method transforms the problem into a spectral space where differentiation becomes multiplication, and complex operations simplify. The Spectral Method achieves unparalleled accuracy for smooth, periodic problems, as each function spans the entire domain, capturing even the finest nuances of the fluid's motion with minimal numerical dissipation and dispersion.

This high-fidelity approach, however, demands simple geometries and periodic or well-defined boundary conditions, often restricting its application to idealized scenarios like turbulence modeling or climate simulations. In these realms, the Spectral Method shines, revealing the intricate, often chaotic beauty of fluid dynamics in vivid detail, much like a maestro conducting an orchestra to unveil the profound complexities of a musical masterpiece (Rai and Mondal, 2021). Figure 4 illustrates the Spectral Method, which represents solutions to partial differential equations using a series of basis functions. The method provides high-order accuracy for problems with smooth solutions, as demonstrated by the convergence of the spectral approximation with increasing modes. A graph illustrating the spectral method's concept of representing a function (e.g., a sine wave) using a series of basis functions. The figure could include a comparison between the original function and its spectral approximation using different numbers of modes.

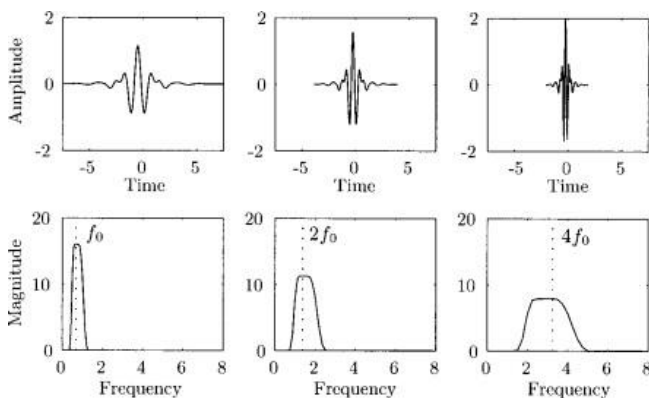


Fig. 4. Illustration of the Spectral Method, where a function is represented as a series of basis functions

In summation, the literature showcases a diverse array of numerical methods available for simulating fluid dynamics, each endowed with unique strengths and limitations. By grasping the principles and capabilities of these methods, researchers and practitioners can harness their full potential to tackle the ever-expanding complexity of fluid flow challenges in science and engineering.

## 2. NUMERICAL METHODOLOGIES

### 2.1 Finite Volume (FV) method

The method divides the computational domain into a finite number of control volumes, or cells, and calculates the values of the variables of interest (e.g., fluid velocity, temperature) at the center of each cell. The governing equations for the Finite Volume method depend on the specific physical problem being solved. However, in the context of fluid flow, the most common equations are the conservation laws, such as the continuity equation (mass conservation), momentum equations (Navier-Stokes equations for incompressible flow), and energy equation (heat transfer). Partial differential equations (PDEs) that describe the behavior of fluid flow, heat transfer, and other physical phenomena are,

continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0 \quad (1)$$

where  $\rho$  is the fluid density,  $t$  is time,  $u$  is the velocity vector,  $\nabla \cdot (\rho u)$  represents the divergence of the mass flux.

momentum equations (Navier-Stokes equations):

$$\frac{\partial (\rho u)}{\partial t} + \nabla \cdot (\rho u u) = -\nabla p + \nabla \cdot \tau + \rho g \quad (2)$$

where  $p$  is the pressure,  $\tau$  is the stress tensor,  $g$  is the gravitational acceleration.

and energy equations:

$$\frac{\partial (\rho E)}{\partial t} + \nabla \cdot (\rho E u) = \nabla \cdot (k \nabla T) + \dot{q} \quad (3)$$

where  $E$  is the total energy per unit mass (internal energy plus kinetic energy),  $k$  is thermal conductivity,  $T$  is the temperature,  $\dot{q}$  represents any internal heat sources or sinks.

These equations are discretized over each control volume in the computational domain using the FV method. The integral form of these equations over each control volume leads to a set of algebraic equations that can be solved numerically to obtain the values of the variables at each cell center. The FV method ensures conservation of mass, momentum, and energy within each control volume and is widely used in CFD simulations due to its robustness and accuracy.

### 2.2 Lattice Boltzmann Method (LBM)

The lattice Boltzmann equation is a simplified kinetic equation that describes the evolution of the distribution function  $f_i(x, e_i, t)$  representing the probability density of finding a particle

with velocity  $e_i$  at position  $x$  and time  $t$ . In its simplest form, the lattice Boltzmann equation can be written as

$$f_i(x + e_i \Delta t, t + \Delta t) - f_i(x, t) = \Omega_i \quad (4)$$

where  $\Delta t$  is the time step,  $\Omega_i$  represents the collision operator, which models the interactions between particles.

To simulate fluid flows, the lattice Boltzmann equation is typically implemented on a regular lattice grid, such as the D2Q9 lattice (2D, 9 velocity directions) or D3Q19 lattice (3D, 19 velocity directions). Each lattice point represents a fluid node, and at each node, there are discrete velocity vectors associated with the lattice directions. The evolution of the distribution functions is governed by streaming and collision processes.

The streaming process updates the distribution functions by moving particles along their respective velocity directions:

$$f_i(x + e_i \Delta t, t + \Delta t) = f_i(x, t) \quad (5)$$

The collision process models the interactions between particles and updates the distribution functions according to collision rules, which may include relaxation towards equilibrium:

$$f_i(x, t) = f_i^{eq}(\rho, u) + \omega_i(x, t) \quad (6)$$

where  $f_i^{eq}(\rho, u)$  is the equilibrium distribution function, which depends on the local fluid density  $\rho$  and velocity  $u$ ,  $\omega_i$  is the collision term.

The macroscopic fluid properties, such as density  $\rho$  and velocity  $u$ , are derived from the distribution functions. For example, the density is obtained by summing all the distribution functions at each lattice node:

$$\rho(x, t) = \sum_i f_i(x, t) \quad (7)$$

And the velocity is calculated as a weighted average of the velocity vectors:

$$u(x, t) = \frac{1}{\rho(x, t)} \sum_i e_i f_i(x, t) \quad (8)$$

Overall, the lattice Boltzmann method simplifies the simulation of fluid flows by discretizing the Boltzmann equation on a lattice grid, allowing for efficient parallel computations, and handling complex boundary conditions. It has become a popular choice for simulating a wide range of fluid flow phenomena due to its simplicity, scalability, and flexibility.

### 2.3 Smoothed Particle Hydrodynamics (SPH)

Smoothed Particle Hydrodynamics (SPH) is a mesh-free Lagrangian method used primarily for simulating fluid flows, although it can also be applied to other physical phenomena like solid mechanics and astrophysics. In SPH, the fluid domain is discretized into a set of particles, and the governing equations are expressed in terms of these particles.

The fundamental equations governing SPH include:

Continuity Equation: The continuity equation ensures mass conservation and is expressed as:

$$\frac{\partial \rho}{\partial t} = -\rho \nabla \cdot (v) \quad (9)$$

where  $\rho$  is the density of the fluid and  $v$  is the velocity of the fluid.

In SPH, this equation is approximated by summing the contributions from neighboring particles within a smoothing length  $h$  around each particle.

Momentum Equation: The momentum equation governs the motion of fluid particles and is typically written as:

$$\frac{\partial v}{\partial t} = -\frac{1}{\rho} \nabla P + \nu \nabla^2 v + f \quad (10)$$

where  $P$  is the pressure,  $\nu$  is the kinematic viscosity, and  $f$  represents external forces such as gravity.

Similar to the continuity equation, this equation is also approximated using neighboring particles within the smoothing length.

Energy Equation: The energy equation governs the thermal behavior of the fluid and is expressed as:

$$\frac{\partial u}{\partial t} = \frac{P}{\rho^2} \frac{\partial \rho}{\partial t} + \frac{v}{\rho} \nabla^2 T \quad (11)$$

where  $u$  is the internal energy of the fluid, and  $T$  is the temperature.

Like the other equations, the energy equation is also approximated using neighboring particles.

In SPH, each particle carries properties such as density, velocity, and energy, and interactions between particles are calculated using smoothing kernels that define how the influence of a particle diminishes with distance. These kernels are typically functions of the distance between particles and the smoothing length.

SPH is advantageous for simulating fluid flows in complex geometries and free surfaces as it does not require a fixed grid, and particles can move freely within the domain. However, it can be computationally expensive due to the large number of particles needed to accurately represent the fluid behavior, especially in scenarios with high spatial gradients. Nonetheless, SPH remains a popular choice for simulating fluid dynamics, particularly in scenarios where traditional grid-based methods may struggle.

### 2.4 Spectral Methods

Spectral methods are numerical techniques used for solving partial differential equations (PDEs) that arise in various scientific and engineering fields, including fluid dynamics. These methods rely on representing the solution to the PDEs as a combination of basis functions, typically chosen to be sinusoidal or polynomial functions. The equations governing spectral methods vary depending on the specific PDE being solved, but the general approach involves transforming the

differential equations into an algebraic form using the chosen basis functions.

**Basis Function Representation:** The velocity components  $u$  and  $v$  and pressure  $p$  represented using Fourier series expansions:

$$u(x, y, t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \hat{u}_{mn}(t) e^{i(mk_x x + nk_y y)} \quad (12)$$

$$v(x, y, t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \hat{v}_{mn}(t) e^{i(mk_x x + nk_y y)} \quad (13)$$

$$p(x, y, t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \hat{p}_{mn}(t) e^{i(mk_x x + nk_y y)} \quad (14)$$

where  $k_x$  and  $k_y$  are the wave numbers in the  $x$  and  $y$  directions, respectively, and  $\hat{u}_{mn}$ ,  $\hat{v}_{mn}$ , and  $\hat{p}_{mn}$  are the Fourier coefficients to be determined.

**Spatial Discretization:** discretize the spatial domain into a finite number of grids points  $(x_i, y_j)$  where the basis functions are evaluated.

**Galerkin Projection or Collocation:** Project the Navier-Stokes equations onto the space spanned by the Fourier basis functions. For example, applying Galerkin projection.

$$\begin{aligned} \int \left( \frac{\partial u}{\partial t} + u \cdot \nabla(u) \right) \cdot \phi_{mn} d\Omega \\ = -\frac{1}{\rho} \int \nabla p \cdot \phi_{mn} d\Omega \\ + \nu \int \nabla^2 u \cdot \phi_{mn} d\Omega + \int f \cdot \phi_{mn} d\Omega \end{aligned} \quad (15)$$

where  $\phi_{mn}$  represents the basis function corresponding to the  $(m, n)$ -th mode.

**Solving Algebraic Equations:** Solve the resulting system of algebraic equations to obtain the Fourier coefficients  $\hat{u}_{mn}$ ,  $\hat{v}_{mn}$ , and  $\hat{p}_{mn}$ .

**Inverse Transformation:** Reconstruct the solution to the original PDE by performing an inverse Fourier transform to obtain the spatial distribution of the solution variables  $u$ ,  $v$  and  $p$ .

**Temporal Discretization:** If the problem is time-dependent, discretize the time domain and solve the resulting system of equations iteratively over time steps using time-stepping schemes such as explicit or implicit methods.

This mathematical framework provides the basis for applying spectral methods to solve the Navier-Stokes equations for incompressible flow using Fourier series expansions. Similar approaches can be applied using different basis functions, such as Chebyshev polynomials or Legendre polynomials, depending on the problem's characteristics and desired accuracy.

Overall, spectral methods offer high-order accuracy and spectral convergence properties, making them well-suited for problems with smooth solutions and periodic boundary conditions. However, they can be computationally expensive

and may require careful treatment of boundary conditions and numerical stability issues. The specific equations governing spectral methods depend on the chosen basis functions and the formulation of the underlying PDEs.

### 3. COMPUTATIONAL DOMAIN, INITIAL CONDITIONS, AND BOUNDARY CONDITIONS

In this chapter, we outline the computational framework, including the domain setup, initial conditions, and boundary conditions, which form the foundation for accurate and efficient fluid dynamics simulations using FV, LBM, SPH, and Spectral Methods. These elements are critical in ensuring the fidelity and stability of the simulations. The revised C++ code has been modified and optimized to implement fluid simulation methods, including Finite Volume (FV), Lattice Boltzmann Method (LBM), Smoothed Particle Hydrodynamics (SPH), and spectral methods, for solving the compressible Euler equations with enhanced computational speed. The square box geometry of the computational domain is defined implicitly through the meshes or predefined particles distribution. The initial density is initialized as a matrix of constant values, and the velocities are initialized as a sinusoidal function. These initial conditions define the starting state of the fluid simulation. The boundary conditions are assumed to be periodic, meaning that the simulation domain wraps around at the boundaries.

#### 3.1 Computational Domain

The computational domain for the simulations is defined as a square box with dimensions  $[0, L] \times [0, L]$ , where  $L=1.0$  represents the characteristic length of the domain. This domain is discretized based on the specific requirements of each numerical method employed.

**Finite Volume Method (FVM):** The domain is divided into a structured grid of  $N \times N$  cells, where  $N$  is chosen to balance computational efficiency with resolution needs. The grid cells are non-overlapping, ensuring mass, momentum, and energy conservation at each cell interface.

**Lattice Boltzmann Method (LBM):** A regular lattice grid is used, with nodes arranged in a D2Q9 (for 2D) lattice configuration. The lattice spacing  $\Delta x$  is selected to satisfy the Knudsen number requirements and ensure accurate resolution of the flow features.

**Smoothed Particle Hydrodynamics (SPH):** In SPH, the fluid is represented by discrete particles distributed across the domain. The initial particle spacing is chosen to ensure adequate resolution of flow features, with a smoothing length  $h$  that is proportional to the initial particle spacing, typically  $h=1.2\Delta p$ , where  $\Delta p$  is the particle spacing.

**Spectral Method:** The computational domain is discretized using a series of orthogonal basis functions (e.g., Fourier or Chebyshev polynomials). The spatial resolution is determined by the number of modes  $M$  used in the spectral expansion, where  $M$  is chosen to capture the dominant flow features while minimizing aliasing errors.



Figures 5 to 8 illustrate the computational domain setup and boundary conditions for the FVM, LBM, SPH, and Spectral method, respectively. The FVM divides the domain into grid cells, while the LBM uses lattice nodes arranged in a regular pattern. The SPH method represents the domain as a collection of particles, and the Spectral Method discretizes the domain using a set of basis functions. Each figure presents a schematic diagram of the computational domain for the respective method. For instance, Figure 5 shows a square domain with grid cells for FVM, Figure 6 depicts lattice points for LBM, Figure 7 illustrates particles for SPH, and Figure 8 shows a domain discretized with basis functions for the Spectral Method.

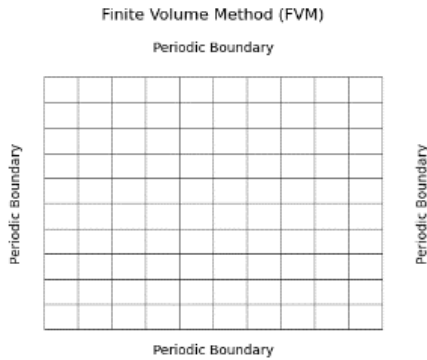


Fig. 5. Computational domain and boundary conditions for FVM

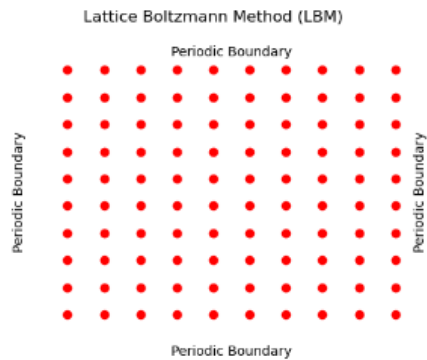


Fig. 6. Computational domain and boundary conditions for LBM

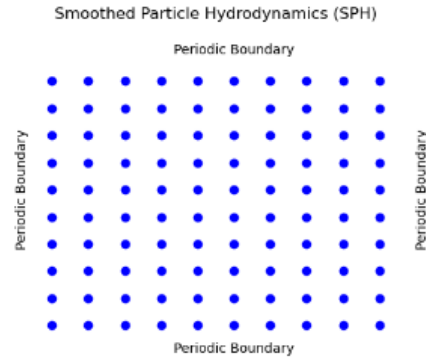


Fig. 7. Computational domain and boundary conditions for SPH

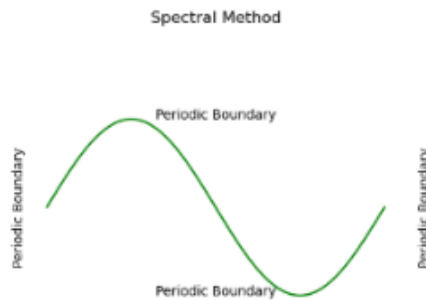


Fig. 8. Computational domain and boundary conditions for Spectral Method

### 3.2 Initial Conditions

The initial conditions are designed to represent a physically realistic starting state for the fluid flow simulations, ensuring that all subsequent dynamics are driven by the inherent physics of the system.

**Density:** The initial density field  $\rho(x,y)$  is initialized uniformly across the domain with a value  $\rho_0$ , ensuring mass conservation from the outset. This uniform initialization is perturbed slightly in some simulations to introduce instability modes, facilitating the study of flow evolution.

**Velocity:** The initial velocity field  $u(x,y)$  is prescribed as a sinusoidal function to model a shear flow or vortex pattern, given by:

$$u_x(x, y) = U_0 \sin\left(\frac{2\pi y}{l}\right) \tag{16}$$

$$u_y(x, y) = 0 \tag{17}$$

where  $U_0$  is the maximum velocity. This setup ensures a well-defined initial momentum distribution.

**Pressure:** The initial pressure  $p(x,y)$  is computed from the equation of state, ensuring consistency with the density and velocity fields. For an incompressible flow, the pressure field is adjusted iteratively to satisfy the incompressibility condition  $\nabla \cdot \mathbf{u} = 0$ .

### 3.3 Boundary Conditions

**Periodic Boundary Conditions:** Periodic boundary conditions are applied in both the  $x$  and  $y$  directions to simulate an infinite, repeating domain. This choice is particularly suitable for studying homogeneous turbulence, shear flows, and other scenarios where the effects of boundaries should be minimized.

### 3.4 Grid Independence and Sensitivity Analysis

To ensure the robustness of the computational results, a grid independence study is conducted for each method. The simulations are performed on progressively finer grids or with increasing numbers of particles/modes until the results converge within a predefined tolerance. This analysis ensures that the chosen discretization is sufficiently fine to capture the essential flow features without incurring unnecessary computational costs. Table 1 compares different grid independence studies for these four numerical methods in this study.

**Table 1. Grid independence study**

Method	Resolution Parameter	Grid 1 (Coarse)	Grid 2 (Medium)	Grid 3 (Fine)	Relative Error
FVM	Grid Size (N×N)	50×50	100×100	200×200	$\epsilon_1=8\%$ , $\epsilon_2=2\%$
LBM	Lattice Spacing ( $\Delta x$ )	$\Delta x=0.04$	$\Delta x=0.02$	$\Delta x=0.01$	$\epsilon_1=7\%$ , $\epsilon_2=1.5\%$
SPH	Particle Count ( $N_p$ )	$N_p=10,000$	$N_p=40,000$	$N_p=160,000$	$\epsilon_1=10\%$ , $\epsilon_2=3\%$
Spectral Method	Number of Modes (M)	M=32	M=64	M=128	$\epsilon_1=5\%$ , $\epsilon_2=0.5\%$

**Relative Error ( $\epsilon$ ):** Represents the error reduction between different resolutions. The error is calculated relative to the finest grid (e.g.,  $\epsilon_1$  for Grid 1 to Grid 2, and  $\epsilon_2$  for Grid 2 to Grid 3).

The grid independence study indicates that the solution becomes increasingly independent of the grid or resolution parameter as it is refined. For the FVM, the relative error between the medium and fine grid ( $\epsilon_2$ ) is significantly smaller than between the coarse and medium grid ( $\epsilon_1$ ), indicating convergence. Similarly, the LBM shows a reduction in error as the lattice spacing decreases, achieving a near-converged

solution at  $\Delta x=0.01$ . In the SPH method, increasing the particle count leads to better resolution of the flow field, with convergence observed as the particle count increases to 160,000. The Spectral Method demonstrates rapid convergence with increasing modes, with minimal error observed at  $M=128$ , which is characteristic of its high-order accuracy.

## 4. RESULTS AND DISCUSSION

Figures 9 to 12 show the density prediction in the flow domain for FV method, LBM, SPH, and spectral method, respectively. The density distribution of the SPH method shows a complete difference among the others since it simulates the flow as the discontinuous particles, although LBM follows the same concept, but it is not a meshless method as SPH.

Here are some observations for Figures 9 to 12, which relate to the density predictions using the four numerical methods:

The density distribution in Fig. 9 shows a smooth variation across the computational domain, indicating that the FVM effectively captures the flow field dynamics. The sharp gradients near boundaries are well-resolved, demonstrating the robustness of the FVM in handling complex geometries and boundary conditions. The method ensures conservation of mass, momentum, and energy within each control volume, which is evident from the consistent density patterns across the domain.

LBM provides a detailed density distribution with high spatial resolution, as shown in Fig. 10. The LBM efficiently handles complex boundary conditions, resulting in smooth transitions in density even near the boundaries of the computational domain. The periodic boundary conditions are effectively implemented, as indicated by the seamless continuity of the density field across the domain edges.

Figure 11 illustrates the density distribution obtained using the SPH method, where the fluid is represented by discrete particles. The SPH method captures intricate fluid dynamics, such as free surface flows and interactions between particles, resulting in a detailed and realistic density distribution. The particle-based nature of SPH allows for adaptive resolution, which is evident in the varying density levels throughout the domain, particularly in regions with high spatial gradients.

The Spectral Method, as shown in Fig. 12, achieves a high-order accuracy in the density prediction, with smooth and continuous density variations across the computational domain. The use of orthogonal basis functions allows the Spectral Method to resolve fine details in the density field, making it suitable for problems with smooth solutions and periodic boundary conditions. The method exhibits minimal numerical dissipation and dispersion, as evidenced by the clear and accurate density patterns, even in areas with significant flow activity.

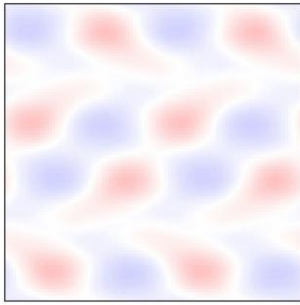


Fig. 9. Density prediction with FVM

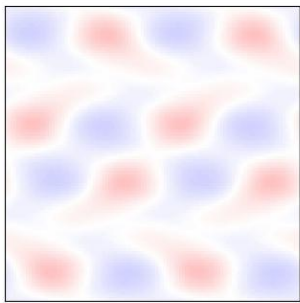


Fig. 10. Density prediction with LBM

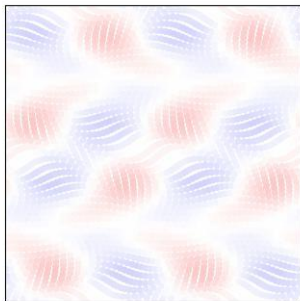


Fig. 11. Density prediction with SPH

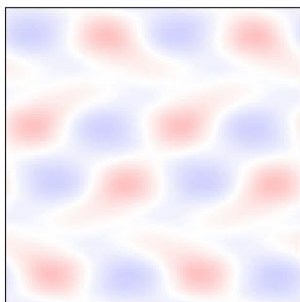


Fig. 12. Density prediction with Spectral method

To provide a comprehensive understanding of the four numerical methods discussed in this paper, Table 2 compares these techniques across various aspects such as discretization, computational efficiency, accuracy, strengths, and limitations.

Table 2. Comparison of Numerical Methods in CFD

Aspect	FVM	LBM	SPH	Spectral Method
Grid	Structured/Unstructured	Regular Lattice	Particle-based	Orthogonal Basis
Efficiency	Moderate, grid-dependent	High, parallelizable	High cost, many particles	High, but expensive
Accuracy	High, grid quality	Moderate to High	High for interactions	Very High, spectral
Strengths	Versatile, robust	Efficient parallel	Mesh-free, deformations	High accuracy, low dispersion
Limitations	Intensive for fine grids	Lattice/time step limit	Expensive, complex	Needs simple, periodic

## 5. CONCLUSIONS

The Finite Volume Method (FVM), Lattice Boltzmann Method (LBM), Smoothed Particle Hydrodynamics (SPH), and Spectral Method each have distinct advantages and limitations in computational fluid dynamics (CFD). FVM is widely used for its robustness and ability to handle complex geometries, making it suitable for various industrial applications, although it requires sophisticated meshing and can be less efficient at high resolutions. LBM, on the other hand, excels in handling complex boundary conditions and is highly efficient on parallel architectures, but it is limited by its lattice structure and interdependent time step and grid spacing. SPH is particularly effective for free-surface flows and problems with large deformations due to its mesh-free nature, but it is computationally intensive and struggles with boundary conditions. The Spectral Method provides extremely high accuracy for smooth and periodic problems, but it is less effective for problems with sharp gradients or irregular geometries. In terms of overall comparison, Spectral Methods are the best for high-precision and smooth problems, while FVM and SPH offer greater flexibility for complex, real-world applications. LBM is emerging as an efficient alternative for specific applications like multiphase flows and porous media. FVM is the most versatile for engineering purposes, SPH is ideal for simulations involving evolving boundaries, and Spectral Methods shine in scientific computations requiring high accuracy. The choice of method hinges on the specific needs of the problem, such as required accuracy, computational resources, and the nature of the physical phenomena being studied.

## REFERENCES

Elsevier Ltd, 209–214.  
[doi:10.1016/j.matpr.2023.04.341](https://doi.org/10.1016/j.matpr.2023.04.341).

- Aliu, O., Sakidin, H., Foroozesh, J. and Yahya, N. (2020) Lattice Boltzmann application to nanofluids dynamics-A review, *Journal of Molecular Liquids*. Elsevier B.V. [doi:10.1016/j.molliq.2019.112284](https://doi.org/10.1016/j.molliq.2019.112284).
- Bagheri, M., Mohammadi, M. and Riazi, M. (2023) A review of smoothed particle hydrodynamics, *Computational Particle Mechanics* [Preprint]. Springer Science and Business Media Deutschland GmbH. [doi:10.1007/s40571-023-00679-7](https://doi.org/10.1007/s40571-023-00679-7).
- Barth, T., Herbin, R. and Ohlberger, M. (2017) Finite Volume Methods: Foundation and Analysis, in *Encyclopedia of Computational Mechanics Second Edition*. Wiley, 1–60. [doi:10.1002/9781119176817.ecm2010](https://doi.org/10.1002/9781119176817.ecm2010).
- Boudet, J. (2011) Finite volume methods, in *Computational Fluid Dynamics*. CRC Press, 1–24. [doi:10.1007/978-3-319-99693-6\\_4](https://doi.org/10.1007/978-3-319-99693-6_4).
- Caban, L. and Tyliczszak, A. (2022) High-order compact difference schemes on wide computational stencils with a spectral-like accuracy, *Computers and Mathematics with Applications*, 108, 123–140. [doi:10.1016/j.camwa.2022.01.006](https://doi.org/10.1016/j.camwa.2022.01.006).
- Chen, Y., Jin, G., Zhang, P., Galdin-Torres, S.A., Scheuermann, A., and Li, L. (2020) An efficient framework for particle-fluid interaction using Discrete Element Lattice Boltzmann Method: Coupling scheme and periodic boundary condition, *Computers and Fluids*, 208. [doi:10.1016/j.compfluid.2020.104613](https://doi.org/10.1016/j.compfluid.2020.104613).
- Prasad, R., Choi, S. and Patil, M. (2022) Aerodynamic shape optimization using a time spectral coupled adjoint for nonlinear aeroelastic problems, *Aerospace Science and Technology*, 126. [doi:10.1016/j.ast.2022.107495](https://doi.org/10.1016/j.ast.2022.107495).
- Rai, N. and Mondal, S. (2021) Spectral methods to solve nonlinear problems: A review, *Partial Differential Equations in Applied Mathematics*, 4. [doi:10.1016/j.padiff.2021.100043](https://doi.org/10.1016/j.padiff.2021.100043).
- Rosswog, S. (2020) A Simple, Entropy-based Dissipation Trigger for SPH, *The Astrophysical Journal*, 898(1), 60. [doi:10.3847/1538-4357/ab9a2e](https://doi.org/10.3847/1538-4357/ab9a2e).
- Samanta, R., Chattopadhyay, H. and Guha, C. (2022) A review on the application of lattice Boltzmann method for melting and solidification problems, *Computational Materials Science*, 206. [doi:10.1016/j.commatsci.2022.111288](https://doi.org/10.1016/j.commatsci.2022.111288).
- Vishwanatha, J. S., Somanath Swamy, R. H. M., Mahesh, G., and Virupaksha Gouda, H. (2023) A toolkit for computational fluid dynamics using spectral element method in Scilab, *Materials Today: Proceedings*.

# Nonlinearity Analysis of Variables for Modelling and Control

Esko K. Juuso

*Control Engineering, Environmental and Chemical Engineering,  
Faculty of Technology, P.O.Box 4300, FI-90014 University of Oulu,  
Finland, (e-mails: esko.juuso@oulu.fi)*

---

**Abstract:** Nonlinearities become essential in various systems when the operating area widens. The linear models are special cases for narrow areas. The behaviour is often asymmetric and can become gradually steeper or flatter depending on the case. These nonlinear effects can be analysed from data distributions for chosen operating areas. Further extensions require recursive analysis. The widely used Gaussian distribution is seldom valid for a wide area. The variable specific scaling can be presented with two second order polynomial defined by five parameters interpreted as the operating point and four corner points of the feasible range. These parameters define the shape factors which may require adjusting to fill the only requirement that the functions need to be monotonously increasing. Alternative constraints provide good solutions for combining expert knowledge with the data-based analysis. If the nonlinear behaviour is analysed correctly, only linear interactions are needed in the models. As the analysis is based on the same methodology, different applications can be combined by using appropriate process data. The smooth operation and high quality of products is the main goal of all these applications, and this can be achieved by combining these indicators with process control in the same way as it has been one for smaller indicators used in condition monitoring and process control. Different parts of the methodology have been tested in versatile applications. The main benefit is that the same structures can be used in various applications since the scaling functions take care of linking to the nonlinear real world.

*Keywords:* intelligent models, nonlinear scaling, statistical analysis, nonlinear systems

---

## 1. INTRODUCTION

Pieces of the informative and reliable datasets are selected in such a way that the data may contain measurement sets from several experiment periods. The *multiple-experiment* sets and selected data periods must be handled appropriately, especially in dynamic modelling. Feedback effects, narrow operating areas and unknown disturbances cause problems in modelling. *Designed experiments* are needed if the data material is not sufficient for modelling (Hinkelmann and Kempthorne, 2008). In industrial applications, the primary goal is to extract the maximum amount of unbiased information from as few (costly) observations as possible.

*Normalisation* or *scaling* of the data is needed for measurements with considerably different magnitudes. Widely used *min-max normalisation* matches the values between the minimum and maximum to the range [0, 1]. The operating point  $c_j$  is fixed in *z-score*,

$$p_j = \frac{x_j - c_j}{\Delta c_j}. \quad (1)$$

which is calculated about the arithmetic mean,  $c_j = \bar{x}_j$ , by using the standard deviation of the variable  $\Delta c_j = \sigma_j$ , transforms the values to a distribution with mean of 0 and standard deviation 1. The arithmetic means and standard deviations are optimal for normal distributions.

*Data distributions* should be taken into account in estimating the centre point  $c_j$  and developing the scaling functions. The geometric mean and harmonic mean are useful when the sample is distributed log-normal or heavily skewed. The median and trimmed mean are two measures that are resistant (robust) to outliers. The trimmed mean ignores a small percentage of the highest and lowest values of a sample when determining the centre of the sample. Scaling with the *median* and *median absolute deviation*, i.e.  $c_j = \text{median}(x_j)$  and  $\Delta c_j = \text{median}(|x_j - \text{median}(x_j)|)$ , provides a solution, which is insensitive to outliers and the points in the extreme tails of the distribution. *Decimal scaling*, where the values are scaled by  $10^{\log_{10} \max(x_j)}$ , suits for cases where the ranges of the variables vary by a logarithmic factor. Minimum and maximum values are very sensitive to outliers.

The *outliers*, which are unusually large disturbances caused for example by temporary sensor or transmitter failures, should be removed from the data. This can be done by examining more thoroughly the data corresponding to the unusually large residual values. An observation is often considered as an outlier if the absolute value  $|p_j|$  obtained by (1) is greater than 3. The Joliffe method has been introduced to detect observations that do not confirm with the correlation structure of the data (Fortuna et al., 2007; Warne et al., 2004). A survey of outlier detection methods is reported in (Englund and Verikas, 2005). As statistical



inspection of process data tend to remove peaks which can carry precious information about system dynamics, all available information, including expert knowledge and input-output relationships, should be used (Fortuna et al., 2007).

*Nonlinear* activation functions, log-sigmoid and hyperbolic tangent, are used to generate the neuron outputs from the sum of the weighted inputs and the bias. These functions have been modified to improve the normalisation of the matching scores in multimodal biometric systems (Jain et al., 2005; Snelick et al., 2005). The *log-sigmoid function*,  $(1 + \exp(-2p_j))^{-1}$ , can be used for nonlinear scaling from the z-score values  $p_j$  to the range  $[0, 1]$ . The *double sigmoid* function extends this with different linear characteristics in the intervals  $[c_j - \Delta c_j^-, c_j]$  and  $[c_j, c_j + \Delta c_j^+]$ . The operating point  $c_j$  and the edges  $\Delta c_j^-$  and  $\Delta c_j^+$  are tuned. The sigmoid function is related to the *hyperbolic tangent tanh* ( $\frac{1}{2}p_j$ ), which scales to the range  $[-1, 1]$ . The functions introduced by Snelick et al. (2005) are based on the scaling of the min-max normalised values with two functions, which are quadratic, logistic and combined linear and quadratic: the inflection point is in the range  $[0, 1]$ .

The *clustering algorithms* can be used for compressing large datasets for modelling: the cluster centres will replace the corresponding datapoints. *Interpolation* is needed if measurements are not frequent enough or if the sampling period is not constant, e.g. various laboratory measurements are based on samples taken infrequently compared to the on-line measurements. In practice, some measurements are missing because of failures in sensors or in data acquisition. These values are either reported as missing or recognised as erroneous values. Missing data can be replaced by using imputation with constants, e.g. the feature or class mean (Enders, 2010). Outliers are handled in the same way but with extra care as their difference from the acceptable values can be fairly small. For large data sets, missing values are simply left out, since the imputation may bias the data. Multiple solutions based on clustering or model-based correction form a basis for iteration.

*Quality control* systems are developed

- to make quality control more effective and closer to real time,
- to identify calibration, measurement and communication errors as close to the observation source as possible,
- to focus on automatic quality control algorithms development,
- to develop a comprehensive flagging system to indicate data quality level,
- to make it easier for data users to identify suspicious and erroneous data, and to highlight corrected values.

Numerous methods are used real-time and non real-time for the spatial and temporal checks of meteorological data (Vejen et al., 2002).

Nonlinear effects can be presented with various functions but the large systems become highly complex combinations of special modules. In fuzzy set systems, the meanings of the variables are shown with a set of membership functions and the interactions between labels are handled with fuzzy rules.

This paper focus on nonlinearity analysis to find unified solutions for modelling and control applications (Section 2). Proposed parametric methodologies are compared with several statistical distributions (Section 3). The methodologies open new possibilities for different types of applications discussed in Section 5. Conclusions and future research are presented in Section 6.

## 2. NONLINEARITY ANALYSIS

The nonlinearity analysis is based on the data distributions in the operating area of the (sub)system. Data values are transformed to dimensionless scaled values, also called linguistic values, are set to be within a real-valued interval  $[-2, 2]$ . The basic scaling approach presented in (Juuso, 2004) has been improved later: a new constraint handling was introduced in (Juuso, 2009), and a new skewness based methodology was presented for signal processing in (Juuso and Lahdelma, 2010).

The generalised data-driven analysis extends solutions with dimensionless features and indicators. The resulting nonlinear scaling functions are compact solutions for variable specific nonlinearity handling.

### 2.1 Fuzzy systems

The origin in fuzzy set systems is seen variable specific feasible ranges which are defined by membership functions. Membership functions for finer partitions can be generated with the scaling functions (Juuso et al., 1993). The support area is defined by the minimum and maximum values of the variable, i.e. the support area is  $[\min(x_j), \max(x_j)]$  for each variable  $j, j = 1, \dots, m$ . The central tendency value,  $c_j$ , divides the support area into two parts, and the core area is defined by the central tendency values of the lower and the upper part,  $(c_l)_j$  and  $(c_h)_j$ , correspondingly. This means that the core area of the variable  $j$  defined by  $[(c_l)_j, (c_h)_j]$  is within the support area.

In early applications, the corner points were extracted from existing rule-based fuzzy systems or defined manually. The fuzzy labels were understood as membership locations corresponding values of the membership definitions within the range  $[-2, 2]$ .

### 2.2 Data-driven analysis

All the parameters are defined together in the data-driven approach. The analysis of the corner points and the centre point were earlier based on the arithmetic means or medians of the corresponding data sets (Juuso, 2004).

The nonlinearity analysis has been later extended to generalised norms defined by

$$\|\tau M_j^p\|_p = (\tau M_j^p)^{1/p} = \left[ \frac{1}{N} \sum_{i=1}^N (x_j)_i^p \right]^{1/p}, \quad (2)$$

where  $p \neq 0$ , is calculated from  $N$  values of a sample,  $\tau$  is the sample time. With a real-valued order  $p \in \mathfrak{R}$  this norm can be used as a central tendency value if  $\|\tau M_j^p\|_p \in \mathfrak{R}$ , i.e.  $x_j > 0$  when  $p < 0$ , and  $x_j \geq 0$  when  $p > 0$ . The norm (2) is calculated about the origin, and it combines two trends:

a strong increase caused by the power  $p$  and a decrease with the power  $1/p$ . All the norms have same dimensions as  $x_j$ . The norm (2) is a Hölder mean, also known as the power mean. The generalised norm for absolute values  $|x_j|$  was introduced for signal analysis in (Lahdelma and Juuso, 2008a).

For variables with only negative values, the norm is the opposite of the norm obtained for the absolute values. If a variable has both positive and negative values, each norm is an average of two norms obtained where the data sets are made positive and negative by subtracting a value  $x_L < \min((x_j))$  and a value  $x_H > \max(x_j)$ , respectively. (Juuso, 2011b)

The generalised norm values increase with increasing order, i.e.

$$(\tau M_j^p)^{1/p} \leq (\tau M_j^q)^{1/q}, \tag{3}$$

if  $p < q$ . The increase is monotonous if all the signals are not equal. The arithmetic mean, the harmonic mean and the root-mean-square (rms) are special cases where the order  $p$  is 1, -1 and 2, respectively. Norms from the minimum to the maximum corresponding the orders  $-\infty \leq p < \infty$  are presented by (2), i.e. the definition includes the  $l_p$  norms defined for  $1 \leq p < \infty$ . The geometric mean is obtain from (2) when the order  $p \rightarrow 0$ .

The computation of the norms can be divided into the computation of equal sized sub-blocks, i.e. the norm for several samples can be obtained as the norm of the norms of the individual samples:

$$\|K_s \tau M_j^p\|_p = \left\{ \frac{1}{K_s} \sum_{i=1}^{K_s} [(\tau M_j^p)_i^{1/p}]^p \right\}^{1/p} \tag{4}$$

where  $K_s$  is the number of samples  $\{x_j\}_{i=1}^N$ . In automation and data collection systems, the sub-blocks are normally used for arithmetic mean ( $p = 1$ ).

### 2.3 Dimensionless features

Distributions of the data can be analysed with dimensionless features obtained by normalising the moments  $M_j^k$ , for example by standard deviation  $\sigma_j$ :

$$\gamma_k = \frac{\tau M_j^k}{\sigma_j^k} = \frac{1}{N \sigma_j^k} \sum_{i=1}^N [(x_j)_i - c_j]^k, \tag{5}$$

where the moment  $M_j^k$  is obtained about some central value, usually arithmetic mean. Variance  $\sigma_j^2$  is the second moment  $M_j^2$ . The feature  $\gamma_3$  is called the coefficient of skewness, or briefly skewness, and the feature  $\gamma_4$  as the coefficient of kurtosis. The skewness is a measure of asymmetry:  $\gamma_3 = 0$  for a symmetric distribution. If  $\gamma_3 > 0$ , the skewness is called positive skewness and the distribution has a long tail to the right, and vice versa if  $\gamma_3 < 0$ . The kurtosis is a measure of the concentration of the distribution near its mean. The generalised moment for absolute values  $|x_j|$  was introduced for signal analysis in (Lahdelma and Juuso, 2008b).

The normalised moments (5) are generalised by using the generalised norm (2) as the central value. The standard

deviation  $\sigma_j$ , which is calculated about the origin, is used to obtain a dimensionless feature. Juuso and Lahdelma (2010) introduced a new approach based on using the generalised skewness  $\gamma_3^p$  for defining the central tendency value and the core area. The central tendency value is chosen by the point where the skewness changes from positive to negative, i.e.  $\gamma_3^p = 0$ . Then the data set is divided into two parts: a lower part and an upper part.

The same analysis is done for these two data sets. The estimates of the corner points,  $(c_l)_j$  and  $(c_h)_j$ , are the points where  $\gamma_3^p = 0$  for the lower and upper data sets, respectively. Since the search of these points is performed by using the order of the moment, the resulting orders  $(p_l)_j$ ,  $(p_0)_j$  and  $(p_h)_j$  are good estimates when additional data sets are used. The norm values can be recursively updated with (4), and a new search for the orders is done only if the values change considerably (Juuso, 2011b).

In practical applications, the data points do not always cover the whole area of operation, e.g. only the close neighbourhood of the normal operation point may be covered, or we would like to extend the model of upper part later to the lower part. Only one part may be in use in fault diagnosis. Expert knowledge is used in extending the feasible range or selecting the methodologies.

Process data often contains outliers, which must be removed before generating the feasible area, because the procedure described above is sensitive to them. This is the idea in medians and trimmed means, which are used for the data samples containing outliers. A good estimate for the support area can be obtained with the generalised norms (2) with large negative and large positive orders since these features are less sensitive to the outliers than the minimum and maximum values. Discarding values at the high and low end can be used together with the generalised norms if there are obvious outliers. Trimming does not need to be the same for the low and high values.

The operating area of each variable is defined by a feasible range represented with a trapezoidal membership function whose corner points are  $\min(x_j)$ ,  $(c_l)_j$ ,  $(c_h)_j$  and  $\max(x_j)$ . Warnings and alarms can be generated directly from the degrees of membership of the complement.

### 2.4 Nonlinear scaling functions

A nonlinear scaling function is defined as a (nonlinear) mapping of variable values inside its range to a range  $[-2, 2]$ , denoted as *linguistic range*. It more or less describes the distribution of variable values over its range which includes the normal operation in the range  $[-1, 1]$  and the areas with warnings and alarms. The values  $X_j$  are called *linguistic values* since the scaling idea originates from the fuzzy set systems: values -2, -1, 0, 1 and 2 can be associated to the linguistic labels, e.g.

$$\{very\ low, low, normal, high, very\ high\} \tag{6}$$

are defined with membership functions The number of membership functions is not limited to five: the values between these integers correspond to finer partitions of the fuzzy set system. The early applications of the linguistic equations used only integer values (Juuso, 1999).

In present systems, membership definitions are used in a continuous form consisting of two second order polynomials:

$$\begin{aligned} x_j &= f_j^-(X_j), X_j \in [-2, 0), \\ x_j &= f_j^+(X_j), X_j \in [0, 2]. \end{aligned} \quad (7)$$

The functions should be monotonous, increasing functions in order to result in realisable systems. The lower part function is defined by values corresponding linguistic levels -2, -1 and 0, and the upper part function by values corresponding linguistic levels 0, 1 and 2. The upper and lower parts should overlap at the linguistic value 0. (Juuso, 2004)

Five parameters define the coefficients of the second order polynomials,

$$\begin{aligned} f_j^-(X_j) &= a_j^- X_j^2 + b_j^- X_j + c_j, X_j \in [-2, 0), \\ f_j^+(X_j) &= a_j^+ X_j^2 + b_j^+ X_j + c_j, X_j \in [0, 2]. \end{aligned} \quad (8)$$

The scaling function is asymmetrical when the coefficients in the upper and lower part are different. The centre point,  $c_j$ , defines the operating point. Four linear equations are needed for solving the other coefficients:

$$\begin{aligned} 4a_j^- - 2b_j^- + c_j &= \min(x_j), \\ a_j^- - b_j^- + c_j &= (c_l)_j, \\ a_j^+ + b_j^+ + c_j &= (c_h)_j, \\ 4a_j^+ + 2b_j^+ + c_j &= \max(x_j). \end{aligned} \quad (9)$$

In order to keep the functions monotonous and increasing, the derivatives of functions  $f_j^-$  and  $f_j^+$  should always be positive (Fig. 1). As a second order polynomial has either a minimum or a maximum point, this requirement is fulfilled only if these points are outside the ranges  $(-2, 0)$  and  $(0, 2)$  for functions  $f_j^-$  and  $f_j^+$ , respectively. The derivatives,

$$\begin{aligned} D_j^- &= 2 a_j^- X_j + b_j^-, X_j \in [-2, 0), \\ D_j^+ &= 2 a_j^+ X_j + b_j^+, X_j \in [0, 2], \end{aligned} \quad (10)$$

are corrected to positive in the areas  $(-2, 0)$  and  $(0, 2)$ , respectively, by changing the coefficients of the polynomials (Juuso, 2004). The membership definitions are continuous functions but derivatives can have discontinuities in the centre point.

The functions are monotonous and increasing if the ratios,

$$\begin{aligned} \alpha_j^- &= \frac{(c_l)_j - \min(x_j)}{c_j - (c_l)_j}, \\ \alpha_j^+ &= \frac{\max(x_j) - (c_h)_j}{(c_h)_j - c_j}, \end{aligned} \quad (11)$$

are both in the range  $[\frac{1}{3}, 3]$ , see (Juuso, 2009). If needed, the ratios are corrected by modifying the core  $[(c_l)_j, (c_h)_j]$  and/or the support  $[\min(x_j), \max(x_j)]$ . Errors are checked independently for  $f_j^-$  and  $f_j^+$ : each error can always be corrected either by moving the corner of the core or the support. In some cases, good results can also be obtained by moving  $c_j$  in the range defined by If these constraints allow a non-empty range, the maximum of the lower limits and the minimum of the upper limit are chosen to define the limits for continuous definitions (Fig. 2).

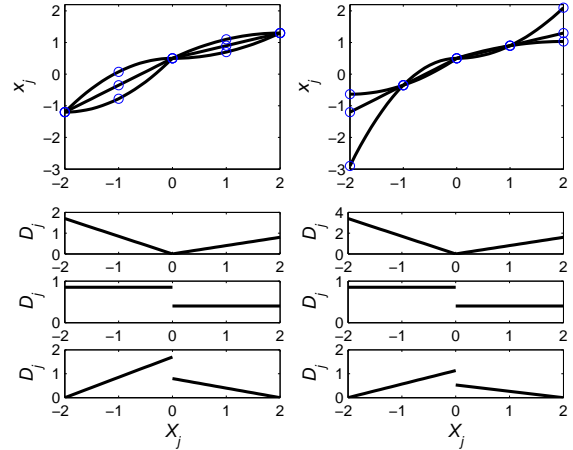


Fig. 1. Feasible shapes of the membership definitions  $f_j$  and corresponding derivatives  $D_j$ : coefficients adjusted with the core (left) and the support (right). Derivatives are presented in three groups: (1) decreasing and increasing, (2) linear and linear, and (3) increasing and decreasing. (Juuso, 2009).

The coefficients of the polynomials can be represented by

$$\begin{aligned} a_j^- &= \frac{1}{2}(1 - \alpha_j^-) \Delta c_j^-, \\ b_j^- &= \frac{1}{2}(3 - \alpha_j^-) \Delta c_j^-, \\ a_j^+ &= \frac{1}{2}(\alpha_j^+ - 1) \Delta c_j^+, \\ b_j^+ &= \frac{1}{2}(3 - \alpha_j^+) \Delta c_j^+, \end{aligned} \quad (12)$$

where  $\Delta c_j^- = c_j - (c_l)_j$  and  $\Delta c_j^+ = (c_h)_j - c_j$ . Membership definitions may contain linear parts if some coefficients  $\alpha_j^-$  or  $\alpha_j^+$  equals to one (Fig. 1).

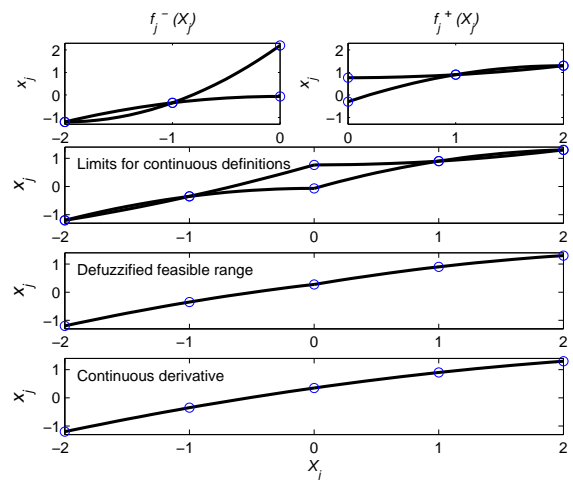


Fig. 2. Membership definitions in the core: coefficients adjusted with the centre point  $c_j$ .

The centre point is not known if the feasible range is defined manually. It can be calculated by defuzzifying the feasible range with the centre of gravity: For strongly asymmetrical feasible ranges, this value may be outside the core (Juuso, 2004). The requirement can be fulfilled

by modifying the corner points. Additional constraints can be taken into account, e.g. a good solution can be to use a locally linear function in the neighbourhood of the centre point. Then a continuous derivative is chosen at the centre point. This can be achieved by modifying the centre point or the corner points of the feasible range. There can be several acceptable modifications, for which the ratios (11) remain in the range  $[\frac{1}{3}, 3]$ .

Monotonously increasing membership definitions can be constructed by adjusting the centre point  $c_j$ , the core  $[(c_l)_j, (c_h)_j]$  and the support  $[\min(x_j), \max(x_j)]$ . An easier way for manual approach was introduced in (Juuso, 2009): first define the centre point  $c_j$ , then the core by choosing the ratios (11) from the range  $[\frac{1}{3}, 3]$ , and finally calculate the support  $[\min(x_j), \max(x_j)]$ . The norms are used together with the generalised skewness in the data-driven approach to define the centre and corner points. The ratios (11), which are checked in all data-driven cases, are also guiding the manual construction of the membership definitions.

For each variable, the membership definitions are configured with five parameters, which can be presented with three consistent sets. The working point (centre point)  $c_j$  belongs to all these sets, where the other parameters are:

- the corner points  $\{\min(x_j), (c_l)_j, (c_h)_j, \max(x_j)\}$  are good for visualisation;
- the parameters  $\{\alpha_j^-, \Delta c_j^-, \alpha_j^+, \Delta c_j^+\}$  suit for tuning;
- the coefficients  $\{a_j^-, b_j^-, a_j^+, b_j^+\}$  are used in the calculations.

The upper and lower parts of the scaling functions can be convex or concave independently. Also, simplified functions can be used: a linear membership definition needs only two parameters:  $c_j$  and  $b_j = b_j^+ = b_j^-$  or  $\Delta c_j = \Delta c_j^+ = \Delta c_j^-$ , since  $\alpha_j^+ = \alpha_j^- = 1$  and  $a_j^+ = a_j^- = 0$ ; an asymmetrical linear definition has  $\Delta c_j^+ \neq \Delta c_j^-$  and  $b_j^+ \neq b_j^-$ . Local linear functions defined by are used if appropriate.

### 3. STATISTICAL DISTRIBUTIONS

In data-based analysis, the nonlinear scaling functions are based on data samples. The parameters obtained by statistical analysis depend strongly on the statistical distribution. The functions extend the normalisation and scaling solutions from the symmetric special case defined by the z-score (1), where  $c_j = \|^\tau M_j^1\|_1$  and  $\Delta c_j = \sigma_j = \|^\tau M_j^2\|_2$ , i.e. generalised norms (2) with orders  $p = 1$  and  $p = 2$ , respectively. Other special cases, geometric mean ( $p = 0$ ) and harmonic mean ( $p = -1$ ), are used in defining the centre of the sample for log-normal or heavily skewed data. Trimmed or truncated means, medians and median absolute deviations are generally recommended for the cases with outliers. The generalised norms can also be trimmed by discarding values at the high and low end. For heavily skewed data, the discarding limits are defined by the norms with high positive and negative orders, respectively.

In the skewness based approach presented above, all the parameters are analysed from the data. As expected, the  $c_j$

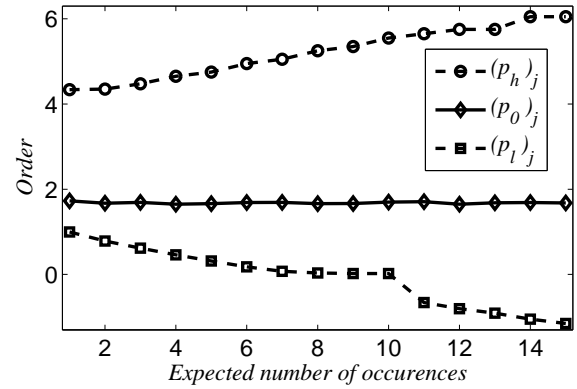


Fig. 3. Orders of the norms (Poisson)

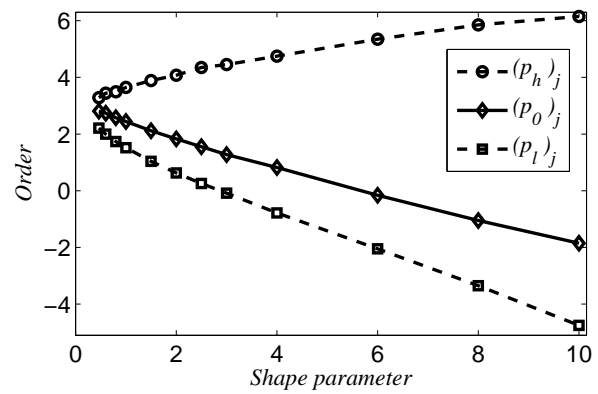


Fig. 4. Orders of the norms (Weibull,  $\lambda_W = 2$ ).

is close to the arithmetic mean ( $p = 1$ ) when the sample is taken from a normal distribution. Normalisation with the z-score is the first phase since the core is symmetrical, i.e.  $\Delta c_j^+ = \Delta c_j^- = \frac{1}{2} \|^\tau M_j^2\|_2$ . The resulting shape factors are equal,  $\alpha_j^- = \alpha_j^+ = 3$ , and the support is  $[c_j - 2\sigma_j, c_j + 2\sigma_j]$ . The size of the random sample effects on the analysis: the centre point is correctly obtained from a small sample ( $N = 1000$ ), and also the core is fairly accurate. The limits of the support area and the shape factors require larger samples, e.g. 10000 points provides fairly good estimate, but 50000 points are required for highly accurate estimates. Only a slight adjustment of the core or preferably the support is needed for these samples.

The scaling functions become asymmetrical about the centre  $c_j$  in random samples of Poisson and Weibull distributions. Orders of the norms,  $\{(p_l)_j, (p_0)_j, (p_h)_j\}$ , and shape factors,  $\{\alpha_j^-, \alpha_j^+\}$ , show strong variations in these asymmetrical distributions (Figs. 3 - 6). For the Poisson distribution, the order  $(p_0)_j$  is almost constant,  $1.68 \pm 0.03$  when the expectation number  $\lambda_P \geq 2$ , and  $(p_0)_j = 1.73$  when  $\lambda_P = 1$  (Fig. 3). For the Weibull distribution, the order  $(p_0)_j$  decreases smoothly from 2.8 to  $-1.85$  when the shape parameter increases from one to ten (Fig. 4). The order range  $[(c_l)_j, (c_h)_j]$  increases for both: from  $[1, 4.34]$  to  $[-1.15, 6.05]$  for Poisson and from  $[2.2, 3.2]$  to  $[-4.75, 6.15]$  for Weibull distributions whose scale parameter  $\lambda_W = 3$ .

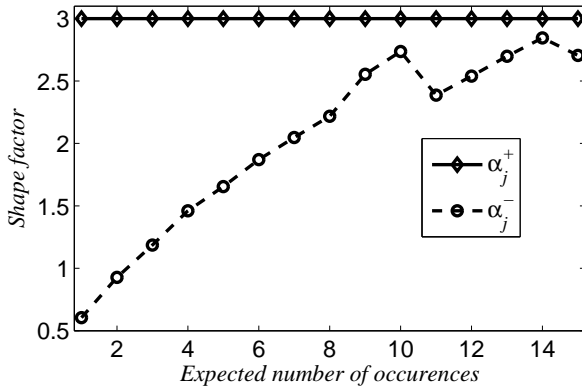
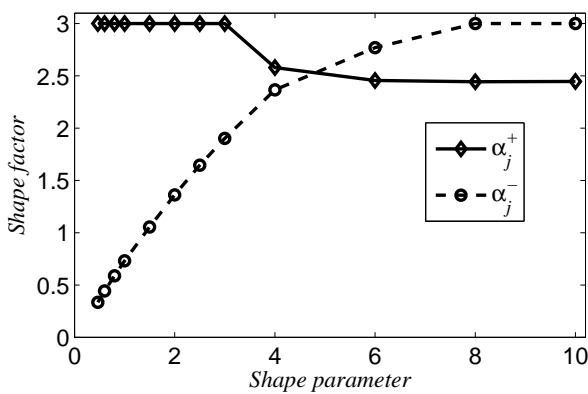


Fig. 5. Shape factors (Poisson).

Fig. 6. Shape factors (Weibull,  $\lambda_W = 2$ ).

Poisson distributions have the same shape factor  $\alpha_j^+$  as the normal distributions, but the shape factor  $\alpha_j^-$  increases from 0.6 to almost 3 when the expectation number  $\lambda_P$  increases from one to ten (Fig. 5). The core also is asymmetrical:  $\Delta c_j^+ > \Delta c_j^-$ . The difference is high, when  $\lambda_P$  is small, and becomes negligible, when  $\lambda_P > 10$  (Fig. 5). Weibull distributions are very asymmetrical when the shape parameter  $\kappa$  is small:  $\Delta c_j^+ \gg \Delta c_j^-$ ,  $\alpha_j^- \approx \frac{1}{3}$  and  $\alpha_j^+ = 3$ , when  $\kappa = 0.5$  (Fig. 6). This exponential distribution becomes more symmetrical when  $\kappa$  increases, but becomes again asymmetrical for higher  $\kappa$  values (Fig. 6). The Poisson distributions have only integer values, which causes irregular changes in orders  $(c_l)_j$  and  $\alpha_j^-$  obtained from random samples.

For all these distributions, the core area becomes wider than in the previous approaches where the mean or the median were used. Higher sensitivity around the centre point was already detected in (Juuso and Lahdelma, 2010). High positive and negative orders are used in selecting the limits for the core area if small deviations are not important. Asymmetrical scaling functions can be obtained by analysing the upper and the lower part separately.

The scaling functions consisting of two second order polynomials operate well for versatile distributions, and various sigmoid functions can be interpreted as special cases. The centre points, which define the operating point of

the model, can be defined manually. For the error, the derivative of error, the sum of error, the original error and the change of control, the centre point is zero. Also, the core and support areas can be defined manually for any membership definition. Monotonous increase needs to be checked for the manually defined functions.

The shape factors define the type of the feasible range: narrow and wide cores correspond to high and low shape factors, respectively. Also, an asymmetric core, i.e. the core can be narrow on one side of the centre point  $c_j$  and wide on the other side, is allowed. The support can depend strongly on the number of points as seen in the comparisons of different statistical distributions. Expert knowledge and physical limitations can be used in selecting the shape factors  $\alpha_j^-$  and  $\alpha_j^+$ . The factors can be set to three if the data set is fairly limited and there is no specific additional knowledge. Linear scaling functions, i.e.  $\alpha_j^- = \alpha_j^+ = 1$  are used if the material is very limited.

#### 4. NATURAL LANGUAGE

The values within the range  $[-2, 2]$  obtained by the nonlinear scaling are also called as linguistic values since they can be interpreted with linguistic terms. The linguistic terms can be interpreted as fuzzy numbers: for example values  $-2, -1, 0, 1$  and  $2$  can be associated to the linguistic labels (6) which can be made sharper or wider with powering modifiers 'extremely', 'very', 'more or less' and 'roughly', and then processed with the conjunction, disjunction and negation. Applications can have specific labels to make understanding easier, and the number of labels are not limited to these examples. The labels are only for information, the calculations are done with the numbers.

#### 5. TYPES OF APPLICATIONS

The nonlinear scaling approach expands the operating areas in many applications. The following areas are examples where compact solutions have been developed. Severity criteria are checked with the scaled values are the same for all variables. Indicators, models and control can be combined in applications (Juuso, 2018).

*Intelligent indicators* are the first applications of the combinations of the generalised norms and nonlinear scaling. Even single norms or indicators can replace and outperform large rule-based systems. Several indicators can be combined as a weighted sum. The severity criteria are the same for these combined indicators as well. (Juuso and Lahdelma, 2010)

*Statistical process control (SPC)* is an important area in utilizing data. The generalised SPC introduced in (Juuso, 2015) expands the SPC from Gaussian to non-Gaussian data sets. The analysis methods are suitable for a large set of statistical distributions. Categorical information can be studied with the same approach by using manual definitions, which means that also mixed cases can be handled. The limits can be updated in short run SPC since they are defined by the nonlinear scaling approach. The limits can even change gradually. The GSPC does not need any interruptions and even recursive approaches are possible. In these systems, the control levels are defined uniformly for the scaled values.



*Modelling and simulation* is extended to nonlinear systems by combining the nonlinear scaling and linear equations. The models can be adapted to different operating conditions by changing the parameters of the scaling functions (Juuso, 2020).

*Intelligent LE controllers* can use linear controller structures in nonlinear systems, The controllers can be adapted to different operating conditions by changing the parameters of the scaling functions. (Juuso, 2011b)

*Temporal analysis* provides indirect measurements and detection of trend episodes for high level control. For any variable, a *trend index* is calculated as a difference of the means of the scaled values obtained for a short and a long time period, respectively. The index value is in the linguistic range  $[-2, 2]$  representing the strength of both decrease and increase of the variable  $x_j$ . The same analysis can be used for detecting temporal changes of any indicators (Juuso, 2011a).

## 6. CONCLUSIONS AND FUTURE RESEARCH

This paper summarizes the main parts of the nonlinear scaling approach. Highly nonlinear asymmetrical data can be utilized in appropriate way. There is no need to assume Gaussian data outside its operating area. Different parts of the methodology has been tested in versatile applications. The main benefit is the analysis of the nonlinear behaviour. Different applications can extend the use of linear structures by enhancing them with the nonlinear scaling to take care about linking to the nonlinear real world. Future research continues this with more detailed analysis of applicability.

## REFERENCES

- Enders, C.K. (2010). *Applied Missing Data Analysis*. Guilford Press, New York.
- Englund, C. and Verikas, A. (2005). A hybrid approach to outlier detection in the offset lithographic printing process. *Engineering Applications of Artificial Intelligence*, 18(6), 759–768.
- Fortuna, L., Graziani, S., Rizzo, A., and Xibilia, M.G. (2007). *Soft Sensors for Monitoring and Control of Industrial Processes*. Advances in Industrial Control. Springer, New York. 270 pp.
- Hinkelmann, K. and Kempthorne, O. (2008). *Design and Analysis of Experiments: Introduction to Experimental Design*. John Wiley & Sons, New York, 2nd edition.
- Jain, A., Nandakumara, K., and Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern Recognition*, 38(12), 2270–2285.
- Juuso, E. and Lahdelma, S. (2010). Intelligent scaling of features in fault diagnosis. In *7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2010 - MFPT 2010, 22-24 June 2010, Stratford-upon-Avon, UK*, volume 2, 1358–1372. BINDT. ISBN=978-1-61839-013-4.
- Juuso, E.K. (1999). Fuzzy control in process industry: The linguistic equation approach. In H.B. Verbruggen, H.J. Zimmermann, and R. Babuška (eds.), *Fuzzy Algorithms for Control, International Series in Intelligent Technologies*, volume 14 of *International Series in Intelligent Technologies*, 243–300. Kluwer, Boston. doi: 10.1007/978-94-011-4405-6\_10.
- Juuso, E.K. (2004). Integration of intelligent systems in development of smart adaptive systems. *International Journal of Approximate Reasoning*, 35(3), 307–337. doi: 10.1016/j.ijar.2003.08.008.
- Juuso, E.K. (2009). Tuning of large-scale linguistic equation (LE) models with genetic algorithms. In *Adaptive and Natural Computing Algorithms, ICANNGA 2009, Lecture Notes in Computer Science*, volume 5495, 161–170. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-04921-7\_17.
- Juuso, E.K. (2011a). Intelligent trend indices in detecting changes of operating conditions. In *2011 UK-Sim 13th International Conference on Modelling and Simulation*, 162–167. IEEE Computer Society. doi: 10.1109/UKSIM.2011.39.
- Juuso, E.K. (2011b). Recursive tuning of intelligent controllers of solar collector fields in changing operating conditions. *IFAC Proceedings Volumes*, 44(1), 12282–12288. doi:10.3182/20110828-6-IT-1002.03621.
- Juuso, E.K. (2015). Generalised statistical process control GSPC in stress monitoring. *IFAC-PapersOnline*, 48(17), 207–212. doi:10.1016/j.ifacol.2015.10.104.
- Juuso, E.K. (2018). Smart adaptive big data analysis with advanced deep learning. *Open Engineering*, 8(1), 403–416. doi:10.1515/eng-2018-0043.
- Juuso, E.K. (2020). Expertise and uncertainty processing with nonlinear scaling and fuzzy systems for automation. *Open Engineering*, 10(1), 712–720. doi: 10.1515/eng-2020-0080.
- Juuso, E.K., Bennavil, J., and Singh, M. (1993). Hybrid knowledge-based system for managerial decision making in uncertainty environment. In N.P. Carreté and M.G. Singh (eds.), *Qualitative Reasoning and Decision Technologies, Proceedings of the IMACS International Workshop on Qualitative Reasoning and Decision Technologies -QUARDET'93, Barcelona, June 16 - 18, 1993*, 234–243. CIMNE, Barcelona.
- Lahdelma, S. and Juuso, E. (2008a). Signal processing in vibration analysis. In *5th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2008 - MFPT 2008, 15-18 July 2008, Edinburgh, UK*, 867–878. BINDT.
- Lahdelma, S. and Juuso, E. (2008b). Vibration analysis of cavitation in Kaplan water turbines. *IFAC Proceedings Volumes*, 41(2), 13420–13425. doi:10.3182/20080706-5-KR-1001.02273.
- Snelick, R., Uludag, U., Mink, A., and Jain, M.I.A. (2005). Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3), 450–455.
- Vejen, F., Jacobsson, C., Fredriksson, U., Moe, M., Andresen, L., Hellsten, E., Rissanen, P., Pálsdóttir, and Arason (2002). Quality control of meteorological observations automatic methods used in the nordic countries. Oslo, Norway.
- Warne, K., Prasad, G., Rezvani, S., and Maguire, L. (2004). Statistical and computational intelligence techniques for inferential model development: a comparative evaluation and a novel proposition for fusion. *Engineering Applications of Artificial Intelligence*, 17, 871–930.

# GPU acceleration of average gradient method for solving partial differential equations

Touko Puro\* Aarne Pohjonen\*\*

\* *Aalto University, Department of Computer Science, Konemiehentie 2, 02150 Espoo (e-mail: touko.puro@aalto.fi).*

\*\* *University of Oulu, Materials and mechanical engineering department, Pentti Kaiteran Katu 1 (e-mail: aarne.pohjonen@oulu.fi)*

---

**Abstract:** Previously presented method of calculating local average gradients for solving partial differential equations (PDEs) is enhanced by accelerating it with graphics processing units (GPUs) and combining a previous technique of interpolating between grid points in the calculation of the gradients instead of using interpolation to create a denser grid.

For accelerating the calculation with GPUs, we have ported the original naive Matlab implementation to C++ and CUDA, and after optimizing the code we observe a speedup factors more than two thousand, which is largely due to the original code not being optimized.

*Keywords:* GPU acceleration, scientific computing, numerical methods, partial differential equations, deformable grids

---

## 1. INTRODUCTION

To intelligently control the formation of material microstructure in sub-micrometer level, modern computational tools are needed. The full field models, such as the level set Hallberg (2013), phase field Steinbach and Salama (2023) and cellular automata Seppälä et al. (2023), offer capabilities to explicitly simulate the microstructure formation Pohjonen (2023). Inclusion of the relevant physical phenomena and their numerical modelling requires the solution of partial differential equations (PDEs).

There are several approaches to obtaining the numerical solution, such as the finite elements, finite differences, finite volume etc. The finite element method is perhaps the most advanced, but it's implementation is not straightforward. Finite differences in the standard implementation is limited to structured grids, which grids would be capable of solving the equations in Eulerian framework. There are approaches to simulate solid mechanics in the Eulerian meshes and they could provide certain advantages such as capability of simulating material distortions without the need of re-meshing, since the material flow through the node points can be simulated. However, more often solid mechanics simulations involving deformations are based in Lagrangian approach, which naturally describes the flow of the material and the material point dependent field variables Basaran (2008).

To simulate the movement of material points within the Lagrangian approach, and to solve the equations in the deformed grid, a triangular two-grid method Pohjonen (2024a) was previously proposed which achieved this purpose in a way which is easy to implement.

In the current work we present enhancements made to the previous version as well as the parallelization of the solution with multi-GPU methods. These improvements pave the way for numerically efficient models that can incorporate the most important physical phenomena affecting microstructure.

## 2. METHODOLOGY

### 2.1 Parallelization

For accelerating our PDE solver we have chosen to use MPI (Message Passing Interface) for communication between processes and CUDA for GPU acceleration. The technologies were chosen since previous codes Pekkilä et al. (2022) built on top of MPI and CUDA have been able to achieve impressive performance for multi-GPU stencil computations, with our PDE solver belonging to this family of *iterative stencil loop* (ISL) -algorithms Li and Song (2004). MPI and CUDA also work well together with MPI having CUDA-aware implementations where the user can send data directly from and to GPU memory. This is convenient for the user and it also provides optimal performance for the user by routing the GPU data through the fastest interconnect Potluri et al. (2013) and pipelining the GPU data movement with the communication.

Furthermore technologies that do not require handwritten kernels to offload computation to GPUs are not competitive in performance, especially at large data sizes Khalilov and Timoveev (2021).

Parallelizing an ISL-algorithm with both MPI and CUDA is conceptually straightforward. The whole domain is split into local subdomains such that each process will process

its own subdomain that it is responsible for updating. In order to compute the required stencil at each grid point the processes communicate grid points that are part of their subdomain to each other as required by the data dependency of the stencil. These regions of points that are communicated to other processes are called *halo regions*. Locally on each process the incoming halo regions coming from other processes are stored to regions called *ghost zones* that surround the local subdomain of the process.

The amount of communication needed for a single stencil iteration depends on the radius of the stencil. The radius of a stencil is defined as the maximum Chebyshev distance from the central grid point to all neighbouring grid points that are required for calculating the stencil at the central grid point.  $N$  iterations of a stencil of radius  $R$  requires rectangular halo regions of width  $N \times R$  for the two-dimensional case.

GPU acceleration is also straightforward since each update at a grid point is independent from updates at each other point. For achieving close to optimal performance for memory-bound applications the GPU acceleration becomes considerably more complex since it becomes important to ensure good cache reuse to alleviate the bottleneck of data movement from global memory. However we observed our application being compute bound on the hardware used, a single RTX A2000 8GB Laptop GPU, to benchmark against the original Matlab implementation, so we were able to achieve relatively good performance with quite simple kernels.

## 2.2 Numerical method

Here we present the implemented numerical method, which was in its initial form first presented in Pohjonen (2024b) as well some optimizations to the equations that now yield the results with less total compute. The optimized algorithms were discovered during implementation work required to accelerate the code.

Each grid point  $p_i$  and its surrounding neighbours can be grouped as triplets, which can be visualized as triangles surrounding the areas  $A_i$  shown in in Fig. 1.

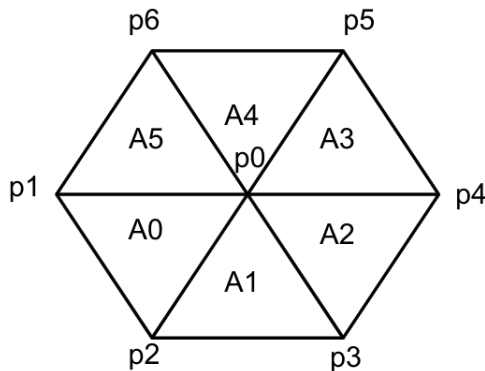


Fig. 1. Each grid point  $p_i$  and its surrounding gridpoints are grouped as triplets, which form triangular regions  $A_i$ . Each triplet defines a plane whose coefficients yield the average gradients of these regions.

Each triplet defines a plane, where the plane coefficients are calculated based on the field values and the positions of the grid points. The coefficients yield the average gradient in each of the areas  $A_i$ . The average gradient in the whole hexagonal region, composed from the surrounding triangles, is then obtained as the weighted average of the gradients by using the areas of the regions as weights.

The equations for the coefficients of the planes are:

$$a = \frac{(u_1 - u_0)y_2 + (u_0 - u_2)y_1 + (u_2 - u_1)y_0}{(x_1 - x_0)y_2 + (x_2 - x_0)y_1 + (x_2 - x_1)y_0} \quad (1)$$

and

$$b = \frac{(u_1 - u_0)x_2 + (u_0 - u_2)x_1 + (u_2 - u_1)x_0}{(x_1 - x_0)y_2 + (x_2 - x_0)y_1 + (x_2 - x_1)y_0} \quad (2)$$

The equations for the areas of the triangles are:

$$A_i = \frac{|(p_i - p_0) \times (p_{i+1} - p_0)|}{2} \quad (3)$$

And finally the equations for the partial derivatives are:

$$\partial_x u|_{p_0} = \frac{\sum_{i=0}^{i=5} A_i a_i}{\sum_{i=0}^{i=5} A_i} \quad (4)$$

$$\partial_y u|_{p_0} = \frac{\sum_{i=0}^{i=5} A_i b_i}{\sum_{i=0}^{i=5} A_i} \quad (5)$$

The first simplification is that since the weighting coefficients are the ratios of the areas of the triangles to the sum of all of the areas the divisor of 2 cancels out in and for the weights we can use the equation:

$$W_i = |(p_i - p_0) \times (p_{i+1} - p_0)|, \quad (6)$$

where the divisor is now the sum of the weights.

Another simplification is that if one calculates the coefficients of the planes with the center grid point being the origin we have  $x_0 = y_0 = u_0 = 0$ , which allows us to simplify some terms in the formula for the coefficients:

$$a = \frac{u_1 y_2 - u_2 y_1}{x_1 y_2 - x_2 y_1} \quad (7)$$

$$b = -\frac{u_1 x_2 - u_2 x_1}{x_1 y_2 - x_2 y_1}, \quad (8)$$

where  $x_1, x_2, y_1, y_2$  are coordinates in the coordinate frame where the center grid point is the origin.

Now importantly, if one chooses the grid points in a way that the cross product responding to the weights is positive, the denominator terms for the equations for  $a$  and  $b$  are equal to the weights. This means that one can skip division and multiplication by the cross products since they cancel out. With this insight the new formula for the partial derivatives becomes:

$$\partial_x u|_{p_0} = \frac{\sum_{i=0}^{i=5} a_i}{\sum_{i=0}^{i=5} W_i} \quad (9)$$

$$\partial_y u|_{p_0} = \frac{\sum_{i=0}^{i=5} b_i}{\sum_{i=0}^{i=5} W_i}, \quad (10)$$

where  $a_i$  and  $b_i$  are computed without the division with the cross product. Having to do less division gives a noticeable performance improvement, since division is a more costly operation than multiplication and addition.

### 2.3 Interpolation with plane equations

Previously a two-grid approach where a denser grid was created out of the original grid by interpolating between each grid point, was presented Pohjonen (2024a). During each timestep the interpolated denser grid is used to calculate the partial derivatives and the results are copied back to the coarse grid, from which a new dense grid is created for the next timestep. The two-grid interpolation method was found to help with instability that was caused near maxima and minima of the grid values. However the two-grid interpolation method comes at a cost namely that one has four as many grid points to compute the partial derivatives compared to the original coarse grid.

This motivated investigation to whether the added accuracy of the two-grid interpolation grid could be achieved without the need for a coarser grid. Since the interpolation points always lie between points in the original grid one can calculate the partial derivative values that would be at the interpolation points using the coarser mesh. One creates the interpolated points and triangles locally but they are only used locally to calculate the partial derivatives and are not stored anywhere. A visualization of the formed triangles can be seen in Fig. 2.

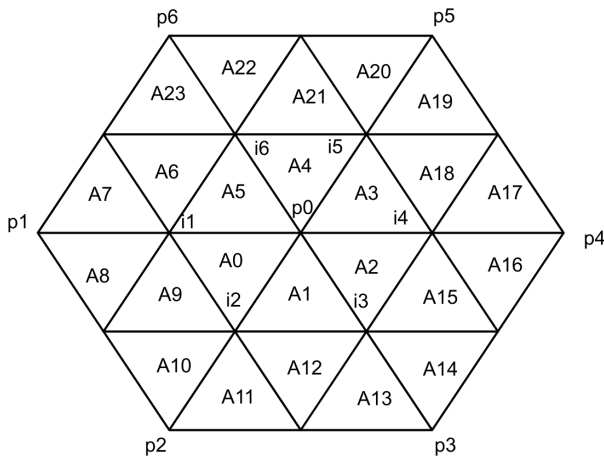


Fig. 2. The interpolated triangles drawn out. Triangles from  $A_0 - A_5$  follow the same indexing scheme as in the non-interpolated scheme. The rest of the triangles are indexed in reverse clockwise order.  $p_0 - p_6$  are the original grid points with  $i_1 - i_6$  being the interpolated points where the first partial derivatives are calculated.

Now as an example  $\partial_x u|_{i_1}$  would be the sum of the coefficients coming from planes:  $A_0, A_5, A_6, A_7, A_8$  and  $A_9$ . One can similarly also calculate the first order partial derivatives at the interpolated points  $i_2, i_3, i_4, i_5, i_6$ . This approach works, but it adds significantly more compute since now there are four times more planes to calculate.

Thus an important observation is that all of the added planes are copies of the innermost six planes, which are copies of the original planes between the non-interpolated points. This is because a linearly interpolated point between two points on a plane stays on the same plane as the original points.

Similarly the area of the interpolated triangles is exactly one fourth of the original triangles so one can use the original areas to calculate the weights since the common factor anyway cancels out. Taking all of this into account the equation for partial derivatives at the interpolated points simplifies, after cancelling another common factor of three out, to:

$$\partial_x u|_{i_{j+1}} = \frac{a_j + a_{j+1}}{W_j + W_{j+1}} \quad (11)$$

$$\partial_y u|_{i_{j+1}} = \frac{b_j + b_{j+1}}{W_j + W_{i+j}}, \quad (12)$$

where the indexing  $j \in \{0, 1, 2, 3, 4, 5\}$  forms a periodic sequence such that  $a_6 = a_0, b_6 = b_0$  and  $W_6 = W_0$ .

Thus one can calculate the first order partial derivatives at the interpolated points with a modest number of added compute. This is a worthwhile trade-off for being able to use a four times smaller grid and eight times smaller grid in two- and three- dimensions respectively. Also, since second order derivatives are only needed at the original coarse grid points one can now calculate second order partial derivatives during the same iteration in which the first order partial derivatives are calculated. This effectively halves the amount of needed memory traffic since the coordinates and field derivatives do not have to be refetched from global memory but can be directly accessed from local memory, either being stored in registers or cache. One also saves compute since the second order partial derivatives are not unnecessarily computed at the interpolation points.

For computing the second order partial derivatives one would in general need the interpolated coordinates because the innermost planes are not anymore the same since they depend on the values at the interpolated points. However, one can show that using the original coordinates gives exactly half of the correct results so instead of calculating the interpolated points one can simply scale up the result computed with the original grid points by a factor of two.

### 2.4 Rectangular grid

The numerical method was originally implemented using a regular rectangular grid Pohjonen (2024b). The regions and grid points for a rectangular grid are visualized in Fig. 3. The equations are exactly the same except now we calculate and add up plane coefficients for four planes instead of six.

The motivation for using a hexagonal grid came from the fact that with a hexagonal grid each interpolation point can be uniquely identified with linear interpolation between two grid points. This is not possible for a rectangular

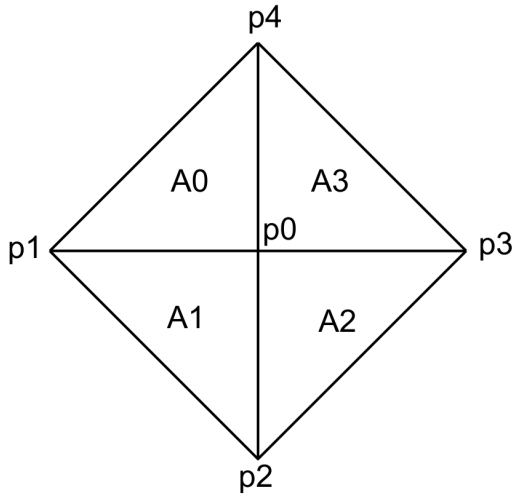


Fig. 3. Visualization of the local regions and grid points for a rectangular grid.

grid and requires for example bilinear interpolation of the surrounding four points for some interpolation points.

The rectangular grid has the advantages that it requires less computation and results in simpler memory access patterns. The rectangular grid requires less compute simply because it requires calculating plane coefficients for four planes instead of the six required for the hexagonal one. The memory access pattern of the hexagonal grid is more complex since technically it requires the use of two different stencils or the union of these stencils and the stencil used depends on the central grid point, whereas the rectangular grid requires only a single stencil that is used for all grid points. A union of two stencils contains all neighbouring grid points that are included in either stencil.

The hexagonal grid requires two stencils since as an example the x-index offset to get the upper left grid point  $p_6$  of the local hexagon depends on the y-index of the center point. Of course our treatment of x- and y -indexes are arbitrary and one can interchange how they are used. The rectangular grid requires only a single stencil since the offsets from index of the central grid point are always the same. Visualizations of the different used stencils can be seen in Fig. 4.

For a code specifically designed for our numerical approach the added complexity of which stencil to use does not really matter but it makes it harder to implement the numerical method optimally in GPU-computing libraries like Pekkilä (2019), where it would otherwise be simple.

Importantly, the communication for both meshes is the same since the union of the two stencils for the hexagonal mesh and the single stencil for rectangular mesh both have a radius of one.

The new improved algorithm motivated us to test could it be used with the original rectangular grid to achieve the improved accuracy of the two-grid approach.

The new algorithm works exactly the same as for the hexagonal grid. Due to cancellation of common terms the

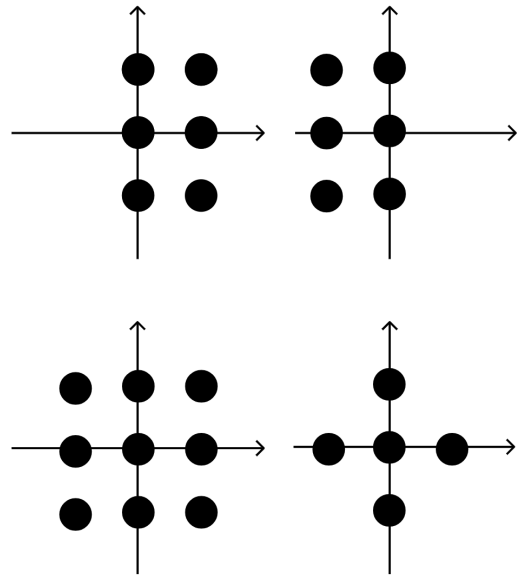


Fig. 4. Visualizations of the different stencils. The updated grid point is located at the origin. The upper stencils are the two stencils required for the hexagonal grid, the lower left stencil their union and the lower right stencil the one required for a rectangular mesh.

equations (11) and (12) can be used to compute the partial derivatives at the interpolated points  $j \in 0, 1, 2, 3$ , where the periodic sequence naturally is now that  $a_4 = a_3, b_4 = b_3$  and  $W_4 = W_3$ . The interpolated points and regions can be seen in Fig. 5

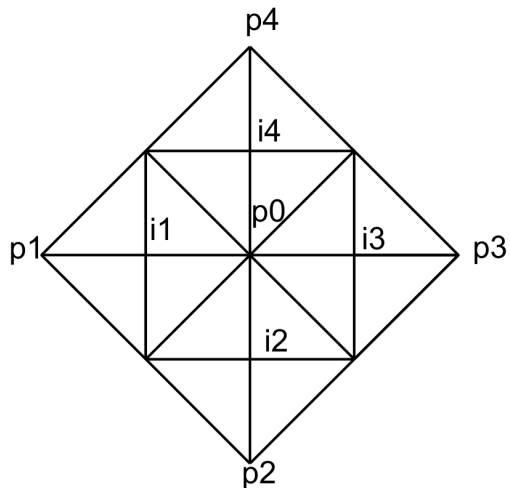


Fig. 5. Interpolated points  $i_{1-4}$  and interpolated regions drawn out.

Even though the equations are the same for using the rectangular grid the method is subtly different in the sense that now the computed gradient values at the interpolated points are different between neighbouring grid points, whereas they are the same with the hexagonal grid. We did not find this discrepancy causing problems when using the new algorithm with the rectangular grid.



### 2.5 Reusing partial results

Since on the tested hardware the GPU kernels were found to be compute bound we investigated approaches that can reuse partially computed results. The observation that the planes  $A_2, A_3$  at grid point  $x, y$  are the same as the planes  $A_0, A_5$  at  $x + 1, y$  motivated the idea of a single thread computing multiple contiguous results in the  $x$ -direction. Thus we implemented a kernel where each  $nx, ny$  thread block computes  $nx \times BlockSize, ny$  results with thread  $i, j$  calculating the results at points  $i, j, i + 1, j, \dots, i + BlockSize, j$ , where  $BlockSize$  is a tunable parameter. In order to avoid non-contiguous memory accesses between threads in a warp, the whole thread block first loads all needed  $nx \times BlockSize + 1, ny + 1$  input values into shared memory in a contiguous manner from which threads can access the values in a non-contiguous pattern without a performance penalty. Similar to the planes the interpolation point  $i_4$  at grid point  $x, y$  corresponds to  $i_1$  at the grid point  $x + 1, y$  so the computed gradient values at the interpolation point can be also reused. A visualization of the algorithm for a  $2 \times 2$  thread block with  $BlockSize = 4$  can be seen in Fig. 6.

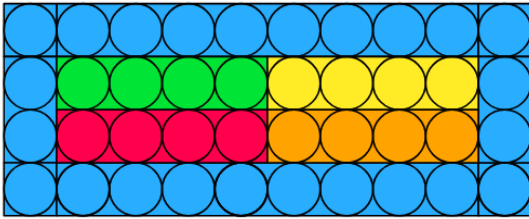


Fig. 6. Visualization of the algorithm for a  $2 \times 2$  thread block with  $BlockSize = 4$ . Blue regions are those that are only loaded and read from shared memory. Green, yellow, red and orange points are computed by threads  $(0, 0), (0, 1), (1, 0)$  and  $(1, 1)$  respectively.

One could also reuse partial results also in the  $y$ -direction, but we leave answering whether reusing partial results in both dimensions increases performance to future work.

### 2.6 Communication

For optimizing the communication of the halo regions between the processes we use the important technique of overlapping computation and communication. MPI supports an asynchronous API where one can immediately continue execution after the sending of a message has started and wait later for the arrival of the message. This allows us to compute the update at those grid points that do not depend on the ghost zones at the same time as the ghost zones are being received. Finally when the ghost zones are received we can compute the update at the outer regions that depend on the ghost zones. This overlaps most of the computation with the communication, which makes the runtime of a grid update be  $\max(W, C)$  instead of the naive  $W + C$ , where  $W$  is the time needed for the computation and  $C$  is the time needed for the communication. In Fig. 7 one can see a visualization of the different subdomain regions of a stencil of radius one.

Our approach for overlapping communication and computation is equivalent to the one in Pekkilä et al. (2022),

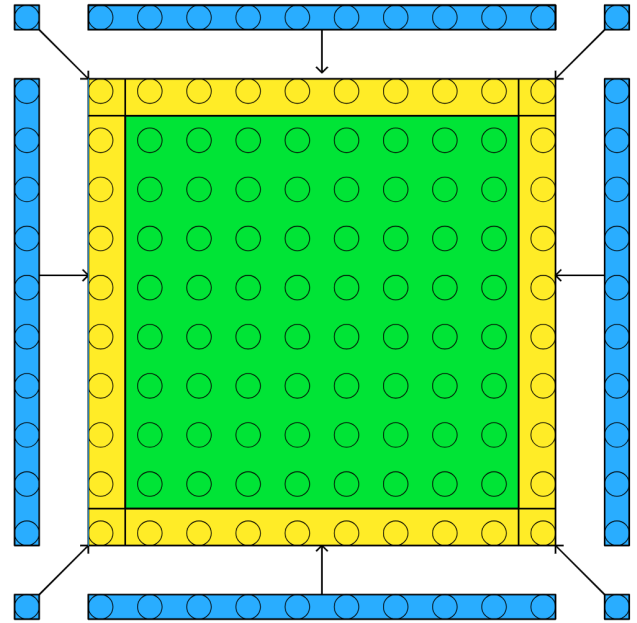


Fig. 7. Regions of a single stencil update for a stencil of radius one. Grid points in green do not depend on the blue ghost zones coming from other processes and their computation can be overlapped with the communication of the blue ghost zones. The yellow outer regions depend on the blue ghost zones and their update has to wait for the arrival of the ghost zones.

except we have a two dimensional simulation instead of a three dimensional one. For performing the original algorithm that requires two iterations of the grid to compute second-order derivatives is also suboptimal from the viewpoint of communication, because it needs halo regions of width two since the grid is iterated twice for a single update. Being able to calculate the update in a single iteration means we can use halo regions of width one, which halves the amount of communication needed.

## 3. RESULTS

### 3.1 Comparison to analytical solution

To make sure the accelerated version and the new algorithm have the same accuracy as the original code we tested the numerical solution similarly to Pohjonen (2024b) and Pohjonen (2024a).

More specifically we solved the diffusion equation (13) from an initial point concentration and compared it to the analytical solution (14) MIT (2024).

$$\partial_t u = D(\partial_{xx} u + \partial_{yy} u) \quad (13)$$

$$u(x, y, t) = \frac{M}{4\pi t D} \exp\left(-\frac{(x - x_c)^2 + (y - y_c)^2}{4Dt}\right), \quad (14)$$

with  $x_c, y_c$  being the origin of the initial point concentration. Values of  $D = 1$  and  $M = 0.1$  were used.

To make sure the code works on a deformed grid we randomly perturbed each grid points coordinates by  $0.16(r -$

$0.5) \cdot a, 0.16(r-0.5) \cdot h$ , where  $r$  is a random number between 0 and 1 and  $a$  is the x-distance between neighbouring grid points and  $h$  is the y-distance between neighbouring grid points. Comparison of the analytical solution to the result of the simulation can be seen in Fig. 8.

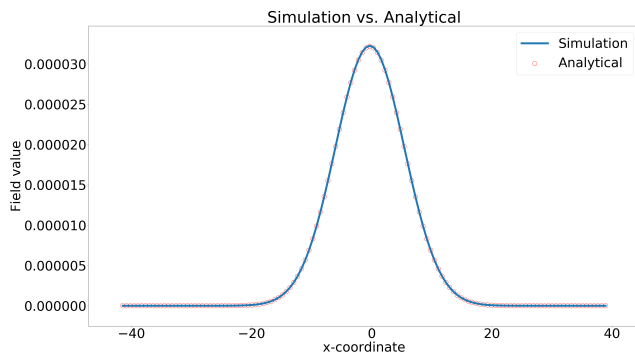


Fig. 8. Simulation result compared to the analytical result. The plotted values are a one-dimensional slice along the x-dimension in the middle of the grid.

We also tested the new algorithm using the rectangular grid and found it to give as precise answers as the hexagonal grid for this simple test case.

### 3.2 Benchmarks

We present timings of the different GPU kernels in Fig. 9. The benchmarks were performed on a single RTX A2000

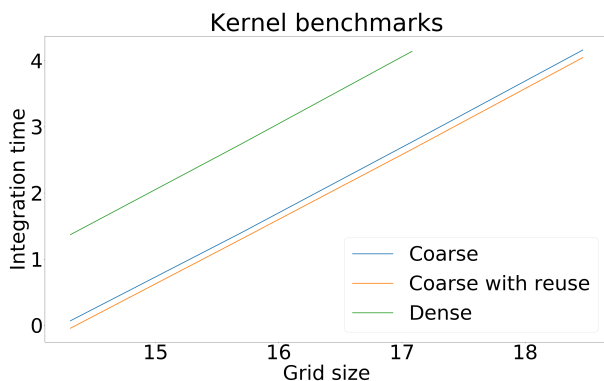


Fig. 9. Timings of different kernels. Integration time is measured as the log seconds taken to integrate 200 timesteps and grid size is measured as the log amount of grid points.

8GB Laptop GPU for easy comparison of our accelerated implementation with the original one.

We observe that the difference between computing with the coarse and dense grid is close to the factor of four one would assume from the grid sizes. A benefit of using the smaller grid and not having to store the calculated first order derivatives means that one needs less memory for the simulation and one is able to simulate larger grids. As an example in the benchmark results one can simulate  $64 \cdot 160 \times 64 \cdot 160$  grid with the coarse grid but with the dense grid one runs out of memory. Also there is quite a constant performance increase of ten percent between

the coarse grid kernel and the coarse grid kernel reusing partial results. *BlockSize* of 11 was found to give optimal performance on the tested hardware.

We do not present detailed timings for the original Matlab implementation but see it adequate to mention that the best benchmark results for the Matlab implementation were when it was more than 2000 times slower. Thus it is clear the the performance increase made possible with the accelerated version makes significantly larger simulations possible. We also do not compare the performance of the GPU implementation against the CPU implementation because we have spent considerably more effort on optimizing the GPU implementation, which would make the comparison unfair and misleading.

### 3.3 Scaling

Not only is the single GPU performance important, but it is important that we get good scaling when increasing the amount of GPUs we use. The theoretical optimal scaling is that when we increase the number of devices by a factor of  $N$  we get a speedup factor of  $N$  also.

The reported scaling benchmarks were performed on the CSC supercomputer Puhti, which has four Xeon Gold 6230 Nvidia V100 -GPUs per node with peak bandwidth of 200 Gpbs between nodes.

For ISL-algorithms the scaling starts to deviate from the theoretical optimum when the network bandwidth becomes the performance limiter Pekkilä et al. (2022). Since the amount of needed compute scales as  $O(N^2)$  and the amount of needed communication scales as  $O(N)$  for a subdomain of size  $N \times N$ , we can get good scaling by keeping the subdomain sizes large enough.

Strong scaling, meaning how well does the code scale when we add more GPUs to a simulation of fixed size, results of the simulation can be seen in Fig. 10. From the figure one

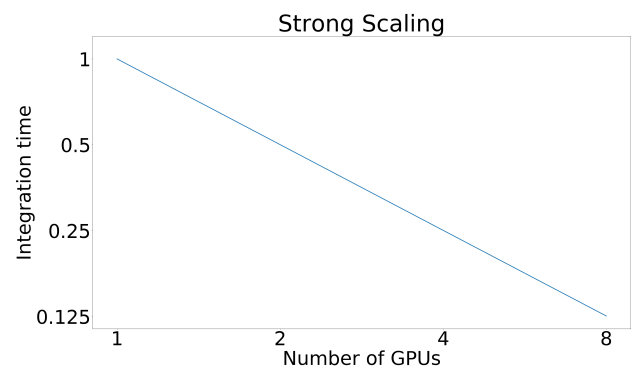


Fig. 10. Strong scaling of  $60 \cdot 160 \times 60 \cdot 160$  grid. Integration time is measured as the time taken to integrate 200 timesteps relative to the time taken on a single GPU.

can see that we have good strong scaling up to 8 GPUs, with the required time going down linearly as expected.

Weak scaling, meaning how well does the code scale when we add more GPUs with a fixed subdomain size, results of the simulation can be seen in Fig. 11. From the plot

one can see that we get good weak scaling with the time required staying effectively constant.

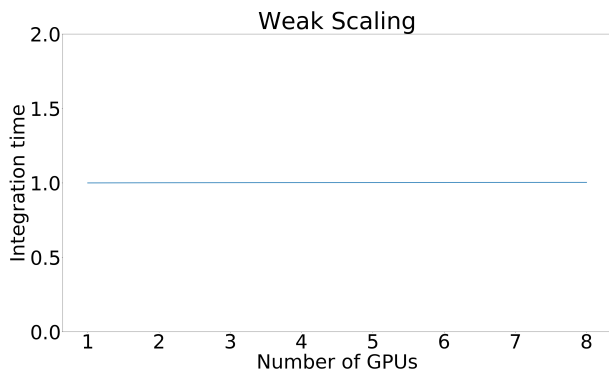


Fig. 11. Weak scaling for subdomains of size  $60 \cdot 160 \times 60 \cdot 160$ . Integration time is measured as the time taken to integrate 200 timesteps relative to the time taken on a single GPU.

#### 4. CONCLUSIONS

We have presented optimizations of the original code and the used numerical method from the viewpoints of GPU acceleration, algorithmic improvements and how the method is mathematically formulated. Together these optimizations achieve an impressive speedup factor of over 2000. This enables more accurate simulations of material microstructure, for example simulations of three dimensional cases, that were previously too prohibitive due to the needed additional computation required. Additionally our implementation scales well to multiple GPUs which enables larger simulations, where more computational resources are needed.

We also investigated the use of a simpler rectangular mesh which would be easier to use and would require less compute. The accuracy of the rectangular mesh has to be studied in harder test cases.

In future studies we plan to conduct larger simulations of material microstructure. After the original implementation work we have reimplemented the algorithm and the solver in the aforementioned GPU-computing library Astaroth.

Because in the new implementation the numerical algorithm is more separated from the rest of code it is more suitable to be the reference implementation. The reference implementation, which can be accessed from [this link](#), also includes preliminary 3d implementation of the method that will be expanded on future publications.

#### ACKNOWLEDGEMENTS

We acknowledge Matthias Rheinhardt for useful discussions that clarified the tradeoffs of using a Lagrangian compared to an Eulerian mesh for material simulations involving displacements. Arne Pohjonen acknowledges financial support from the AID4GREENEST project (AI powered characterization and modelling for green steel technology) funded by Horizon Europe Research and Innovation Program, Grant Agreement No. 101091912

#### REFERENCES

- Basaran, S. (2008). *Lagrangian and Eulerian descriptions in solid mechanics and their numerical solutions in hpk framework*. Ph.D. thesis, University of Kansas.
- Hallberg, H. (2013). A modified level set approach to 2D modeling of dynamic recrystallization. *Modelling and Simulation in Materials Science and Engineering*, 21(8), 085012. doi:10.1088/0965-0393/21/8/085012. doi:10.1088/0965-0393/21/8/085012.
- Khalilov, M. and Timoveev, A. (2021). Performance analysis of CUDA, OpenACC and OpenMP programming models on TESLA V100 GPU. In *Journal of Physics: Conference Series*, volume 1740, 012056. IOP Publishing.
- Li, Z. and Song, Y. (2004). Automatic tiling of iterative stencil loops. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 26(6), 975–1028.
- MIT (2024). Lecture Notes on diffusion, accessed 10.6.2024, Massachusetts Institute of Technology. URL <http://web.mit.edu/1.061/www/dream/THREE/THREETHREORY.PDF>.
- Pekkilä, J. (2019). Astaroth: A library for stencil computations on graphics processing units. URL <https://aaltodoc.aalto.fi/server/api/core/bitstreams/c73ad7b3-47a2-4c23-b802-7721366fb961/content>.
- Pekkilä, J., Väisälä, M.S., Käpylä, M.J., Rheinhardt, M., and Lappi, O. (2022). Scalable communication for high-order stencil computations using CUDA-aware MPI. *Parallel Computing*, 111, 102904.
- Pohjonen, A. (2023). Full field model describing phase front propagation, transformation strains, chemical partitioning, and diffusion in solid–solid phase transformations. *Advanced Theory and Simulations*, 6(3), 2200771. doi:10.1002/adts.202200771. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.202200771>.
- Pohjonen, A. (2024a). Application of two-grid interpolation to enhance average gradient method for solving partial differential equations. *The Journal of Physics: Conference series*, 2701(2), 97–111.
- Pohjonen, A. (2024b). Solving partial differential equations in deformed grids by estimating local average gradients with planes. *The Journal of Physics: Conference series*, 2701(2), 97–111.
- Potluri, S., Hamidouche, K., Venkatesh, A., Bureddy, D., and Panda, D.K. (2013). Efficient inter-node MPI communication using GPUDirect RDMA for InfiniBand clusters with NVIDIA GPUs. In *2013 42nd International Conference on Parallel Processing*, 80–89. IEEE.
- Seppälä, O., Pohjonen, A., Mendon, V., Podor, R., Singh, H., and Larkiola, J. (2023). In-situ SEM characterization and numerical modelling of bainite formation and impingement of a medium-carbon, low-alloy steel. *Materials & Design*, 230, 111956. doi:10.1016/j.matdes.2023.111956. URL <https://www.sciencedirect.com/science/article/pii/S0264127523003714>.
- Steinbach, I. and Salama, H. (2023). *Lectures on phase field*. Springer Nature.