Machine learning assisted real-time heat load forecasting in district heating network

Avinash Renuke [a,*], Stavros Vouros [a], and Konstantinos Kyprianidis [a]

[a] *Mälardalen University, Västerås, Sweden*
*corresponding.avinash.renuke@mdu.se

Abstract

District heating system often consists of a long, complex network of piping carrying heat from a power plant to the consumers. The supply temperature from the plant is either controlled by the operator from experience or a predefined curve based on the outdoor temperature. An optimized supply temperature which would be lower than the one obtained traditionally would lead to lower heat loss and reduced peak load on the power plant. In this paper, we investigate the machine learning models for heat load forecasting which is a crucial parameter in the optimizing process. Models are generated using supervised machine learning algorithms: Linear models (Linear Regression, Ridge and Gaussian Process Regressor), Random Forest Regressor, Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) recurrent neural network (RNN). Data-driven models are used extensively in the literature to predict heat load prediction based on the weather and the time effect on a fixed training set, however, in this study, we model the heat load in the network in real-time scenarios i.e., adaptive training and forecasting. The model is adaptively updated as well as the training of the machine learning model in real time. It provides a “plug-and-play” solution for real-time prediction without significant pre-tuning requirements. The results of all the models are compared with various time horizons i.e., 6 hrs, 10 hrs, 24 hrs and 1 week, using the district heating data obtained for the city of Vasteras in Sweden. The performance of the prediction algorithms is evaluated using Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). An algorithm with the best accuracy is selected based on the performance comparison. Also, models suitable for short-term and long-term forecasting are discussed towards the end of the article.

1. Introduction

Optimizing district heating systems using machine learning (ML) techniques has gained significant interest in recent research. ML models can help improve the efficiency and operation of district heating systems by optimizing heat distribution and consumption. ML models can be used to predict heat demand patterns in buildings or areas connected to the district heating system. By analyzing historical data, weather conditions, and other relevant factors, these models can forecast heat demand accurately. This enables more efficient planning and optimization of heat supply from the plant and distribution, reducing energy waste and costs.

The focus of this article is a combined heat and power (CHP) plant, also known as a cogeneration plant, a type of district heating (DH) system that simultaneously generates electricity and useful heat from a single energy source. The CHP plant uses a primary energy source, such as natural gas, biomass, or waste heat, to produce electricity and heat in a combined process. The primary energy source drives a generator to produce electricity, while the waste heat generated during electricity generation is captured and used for heating purposes. The waste heat produced during electricity generation in the CHP plant is captured and utilized for district heating. This waste heat is typically recovered through heat exchangers and transferred to a heat distribution network. The heat distribution network consists of insulated pipes that transport hot water from the CHP plant to connected buildings and areas. Heat exchanges and control valves regulate the flow and temperature of the heat within the network. Buildings and residential areas connect to the district heating network through heat exchangers in the consumer substation. Heat exchangers transfer the heat from the hot water in the network to the building's internal heating system, providing space heating and sometimes domestic hot water. Currently, the supply temperature from the plant is governed by the operator’s expertise and knowledge based on historical consumption data. Due to temperature delay in the piping network, the operator injects heat into the network without any clarity on actual heat demand at the end user. The drawback of this strategy is that, although satisfying consumer expectations, the return temperature in the network would be higher than necessary. This suggests that lowering the network supply temperature has a significant potential to reduce the load on DH plants in addition to the fact that the supply temperature is frequently greater than is necessary. The optimum peak supply temperature, which is lower than the historical supply temperature, would facilitate more electricity production for the CHP plant. Such energy optimization could result in a district-wide reduction in greenhouse gas emissions given the rising worries about climate change.

Heating load forecasting plays an essential role in reducing network losses and performance optimization (reduced plant supply and return temperature). Supervised machine learning (SL) techniques are widely researched for heat demand forecasting in district heating systems while Reinforcement learning(RL) is suitable for optimal control and load balancing strategies (Idowu, Åhlund and Schelén, 2014). These methods leverage historical data, weather information, and other relevant factors to predict future heat demand accurately. Researchers have studied various ML models such as Regression models, Support vector Machine (SVM), Nature inspired, Artificial Neural network (ANN) etc. (Idowu *et al.*, 2014) presented a data-driven heat load forecasting for multi-family buildings using four ML methods – SVM, ANN, Multiple Linear regression(MLR) and Classification and Regression Tree (CART). The forecasting model was evaluated for horizon values of 1, 3, 6, 12, 18 and 24-h. The SVR method was found to be the best performing followed by MLR. The authors were able to achieve a 5.6% (best-performing) Mean Absolute Percentage Error (MAWP). The use of a context-based regression approach model for average and individual user consumption is shown to be effective (Rongali *et al.*, 2015). Dalipi et al (Dalipi, Yildirim Yayilgan and Gebremedhin, 2016) have performed a comparison of ML models for heat load forecasting for multiple buildings and showed SVR to be the best-performing one. Idowu et al (Idowu *et al.*, 2016) evaluated and compared different ML methods such as SVM, FFNN, MLR and regression trees. The models are produced and evaluated using data observed from 10 district heating substations for five multi-family apartments and five commercial buildings. They included outdoor temperature, day of the week, hour of the day, historical values of thermal load and the physical parameters of a substation (supply temperature, difference between supply and return temperature and flow rate) for a forecast horizon up to 48 hr. They found that SVM gave the best prediction performance, with FFNN and MLR having similar error rates as SVM and Regression tree models with higher performance error rates. Suryanarayana et al (Suryanarayana *et al.*, 2018) have shown that the linear models can be very powerful for heat load forecasting outperforming some of the advanced models like SVR (Support Vector Regressor) and Gradient-Boosted Trees (GBT). They achieved an MAWP of as low as 8.77 with linear models and 8.07 with DNN. Ntakolia et al (Ntakolia *et al.*, 2022) did a comprehensive review of machine learning models used for heat load forecasting and concluded that ANN and SVM are found to be the most frequent models used for heat load prediction.

Accurate heating load forecasting is a precondition to the optimization and control of district heating systems. Zhao et al. (Zhao, Li and Shan, 2021) have used SVM in their study to forecast heat load to optimize the DH system using Model predictive control (MPC). They have shown a reduction in energy peak and total energy consumption of the system. Similarly, Zimmerman et al. (Zimmerman, Kyprianidis and Lindberg, 2019) demonstrated a reduction in network supply and return temperatures by using MPC with feed-forward with CHP heat load reduction of 12.5% to 13.7%. However, the weakest point of the paper was that a realistic heat-load model was missing. The machine learning models studied in the literature for heat load predictions are for fixed training sets and not adaptive for real-time scenarios which creates the motivation for the present study.

The objective of this paper is to compare machine learning models to forecast heat load for the DH system of the city of Västerås in Sweden, for adaptive and real-time forecasting. The model created should not require any pre-tuning for use in the DH system. Following machine learning techniques have been considered in this paper:

1. Linear models: linear regression (LR), Ridge Regression(RR), Gaussian Process regression(GPR)
2. Support Vector Regressor (SVR)
3. Random Forest (RF)
4. LSTM (Long-Short-Term Memory)

The model is developed for a single substation (Tillberga) and then validated with another substation (Skultuna) present in Västerås city (explained in the next section).

The paper is organized as follows. Firstly, the DH system is described followed by a modelling approach and an overview of the considered algorithms is presented. Secondly, the methodology is described, including the training, validation and test periods. The accuracy of the different models is then investigated and compared using experimental data from a real-world DH plant.

2. Methodology

In this section, to emphasize the motivation behind this work, a brief overview of the DH system is provided first. The dataset, ML models and methods, variables, and their structure are then explained.

2.1. District heating system

The DH system presented in this study is from the city of Västeås and its surrounding region. A typical DHS has three main parts - The heat generation - which usually consists of a co-generation plant and/or a heat-only boiler station, the Distribution network - which consists of insulated pipes of varying diameters carrying hot water through the entire network and the substations - where heat is transferred from the primary to the secondary network via a heat-exchanger. Fig. 1(top) shows the basic schematic drawing of a CHP plant and a DH system. The red and blue lines denote supply and return pipes respectively. Fig. 1(bottom) shows the schematic illustration of the city of Västerås with a CHP plant connected with different regions around the city.

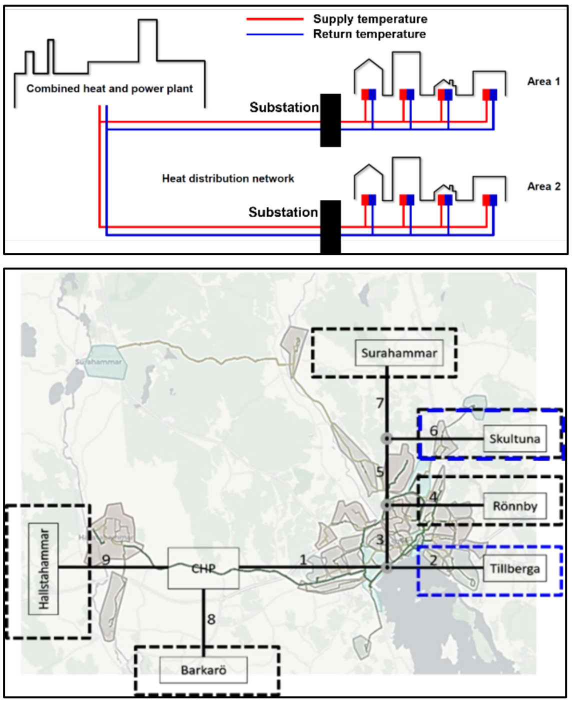

Figure 1: (top)A schematic diagram showing a district heating system network with substations connected to the regions (bottom) Schematic illustration of regions connected with the DH system to the CHP in the city of Västerås and its surrounding regions

The supply temperature of the hot water is controlled directly from the plant's control room based on the outdoor temperature and it follows mostly a given operation temperature curve. The return temperature, on the other hand, depends mainly on the customer's heat usage. The current DH system is 3rd generation where the temperature level varies between 70 and 120°C, particularly during the winter season. The heat load in district heating systems is the sum of all heat loads that are connected to the network and distribution and other losses in the network.

The heat load ML model is created using the DH system of the region Tillberga, which is approximately 14.5 km from the CHP. The reasons for choosing Tillberga are due to the fact of having access to historical network measurements for multiple years. The ML model will be validated both with Tillberga and Skultuna of the city of Västerås.

The heat consumption at the substation, Q_{sb} delivered to the region is mainly a function of the supply temperature, Ts, the return temperature, Tr, and the flow rate, m, as shown in Eq. (1).

$$Q_{sb} = m.C.(T_s - T_r) \tag{1}$$

Where, Q_{sb} is heat consumption at the substation, C is the specific heat of hot water, T_s is supply temperature and T_r is the return temperature at the substation.

2.2. Dataset

The data used in this study are provided by Mälarenergi who operates the CHP plant. These data are measured and collected by regular measurements that are part of the control system in a DH plant. The measurements consist of the time of day, supply temperature (ST), return temperature (RT), flow rate (FR), and heat load (HL) at the individual substations with 1-hour time step intervals. The data used in this study are for the winter season: January 1st 2019 to March 31st 2019 for model selection and January 1st 2022 to March 31st 2022 of Tillberga and Skultuna for testing.

Before building a machine learning model, it is important to preprocess the data and remove or replace any missing values or outliers. The data is preprocessed to remove and replace any missing data/outliers (outside ±3-standard deviation) that could cause problems with the ML model, such as biased results or inaccurate forecasting. We used interpolation to replace the missing/outlier data. Another step in the data preprocessing is to scale the data. Scaling means changing the range of data so that all the values are within a similar range. The data is scaled using the 'MinMaxScaler() (*sklearn.preprocessing.MinMaxScaler*)' function from the Python library which scales and translates each feature individually such that it is in the given range on the training set, e.g. between zero and one.

The data is then split into training, validation, and test sets. This is because we want to train our models on the training data, validate them on the validation data, and then test them on the test data.

In this model, we have used a 70-30 split for training and testing respectively to assess the model.

2.3. Input variables

The heating demand of a consumer depends on several types of factors such as climate and weather conditions, timestamp information, and historical information. In this work, we use the data which is readily available from most of the plants which are limited to weather, historical, and time-stamp information.

a. Historical heat load data
b. Weather input variables: Outdoor air temperature
c. Time-stamp variables: hour of the day and day of the week

2.4. Machine learning models

The following ML models are studied in this paper which are found to be more suitable for the heat load forecasting problem (Ntakolia et al., 2022). In this study, we have used supervised ML models which use a set of input variables to forecast the values of an output variable. In these models, we look at historical data to *train* a model to learn the relationships between variables and a *target (output),* the thing we're trying to forecast. This way, when new data comes in, we can use the input values to make a good prediction of the output (target). We have studied the four most widely used supervised learning algorithms: Linear models (LM), Random Forest (RF): a Tree-based model, SVM and LSTM: ANN based model.

Linear models: Linear ML models are simple (old-school) algorithms that make predictions based on linear relationships between the input variables and the target variable. The relationship between the inputs and the target is represented by a linear function. The following three LMs are used in this study:

a. LR: Linear regression models the relationship between the inputs (features) and target by fitting a linear equation to the observed data. It aims to minimize the sum of squared residuals between the predicted and actual target values.
b. Ridge Regression (RR): Ridge regression is a regularized version of LR that adds a penalty term to the cost function, aiming to reduce the model's complexity and prevent overfitting.
c. Gaussian Process Regression (GPR): GPR is a non-parametric Bayesian regression technique that can be used for regression. It models the relationship between the input and the target variable as a distribution over functions rather than a single function.

Random Forest (RF): RF is a tree-based algorithm that combines multiple decision trees to make predictions. It belongs to the ensemble learning family, where multiple models are combined to improve overall performance and generalization.

Support Vector Regressor (SVR): SVM, based on a statistical learning theory, are one of the most successful and widely applied ML methods, for solving regression problems. SVR is a method of SVM for regressions. SVR is a powerful approach for handling non-linear regression problems by mapping the input variables into a higher-dimensional space and finding an optimal hyperplane that fits as many data points as possible within a specified margin.

Long-Short-Term Memory (LSTM): LSTM is a type of recurrent neural network (RNN) architecture that is widely used for sequential data analysis, such as time series forecasting, natural language processing, and speech recognition. They introduce specialized memory cells and gating mechanisms that allow the network to selectively remember or forget information over time. LSTM networks have been proven effective in modelling and predicting sequences with long-term dependencies (Schaefer, Udluft and Zimmermann, 2008).

Hyperparameter analysis is a crucial step in machine learning to optimize model performance. It involves tuning the hyperparameters of a machine learning algorithm to find the best combination that yields the highest accuracy or lowest error on a given dataset. In this study, an automated grid search was conducted for each specific ML model by selecting the hyperparameters of the ML model. Grid search is a popular technique for hyperparameter optimization in machine learning. It involves exhaustively searching through a specified grid of hyperparameter values to find the combination that yields the best model performance.

Table 1 shows the hyperparameters used in each ML model.

All the ML models are assessed using open-source software like Python with packages like Scikit-Learn (*scikit-learn: machine learning in Python — scikit-learn 1.2.2 documentation*, 2023) and Tensorflow (*TensorFlow*, 2023).

2.5. Performance metrics

In this work, we considered commonly used metrics for evaluating the performance of the proposed models. These are the Root Mean-Square

Error (RMSE), the Mean Absolute Percentage Error (MAPE) and the Correlation coefficient (Corr Coef). In this paper, we have used MAPE and RMSE for performance evaluation.
RMSE is commonly used to measure the difference between a model's predicted values and actual values observed (the average prediction error over all time instants) and is computed as:

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(\hat{y_i} - y_i)^2\right)} \qquad (2)$$

Table 1: ML models and hyperparameters

ML model	Hyperparamete	Definition	Defined Parameter
LR (Linear	-	-	-
RR (Ridge Regression)	alpha (α)	the regularization parameter	0.01 - 10
GPR (Gaussian Process Regression)	Kernel type	the Kernel function	RBF(Radial Basis Function)
	length_scale	Smootheness of the kernel	0.1- 10
	alpha (α)	Regularization Parameter	0.001-10
RF (Random Forest)	n_estimators	number of decision trees in the random	100-500
	max_depth	the maximum depth of each	0-10
SVR (Support Vector Regressor)	Kernel type	the Kernel function	RBF(Radial Basis Function)
	C	the penalty paramet	0.1- 10
	epsilon	the margin of tolerance around the predicted value	0.01-1
LSTM (Long-Short-Term Memory)	units	the number of memory cells or hidden units in the LSTM layer	10-500
	activation	Activation Function	relu
	Batch Size	-	8-64
	Number of Epochs	-	50-500
	Loss Function	-	mean_squared_error
	optimizer	-	adam

MAPE estimates how close forecast values are to actual values in percentage and is computed as:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|\hat{y_i} - y_i|}{|y_i|} \qquad (3)$$

where $\hat{y_i}$ is the estimated value in a forecasting model, y_i is the measured value and n is the total number of forecasted data points in each forecast horizon.

2.6. Model workflow

As mentioned earlier, we have used two datasets from different years to build and test the models in adaptive learning and real-time scenarios. The first dataset i.e. dataset from region Tillberga from 2019 is used to test the performance of different ML models. The decision is taken based on their performance metrics and the best model is selected for adaptive learning. Figure 3 shows the workflow of the development of the heat load forecasting model. The selected model with optimized hyperparameters will be used in the dataset of the year 2022 for Tillberga and Skultuna of Västerås City. The final model is trained adaptively as the new data adds to the historical set and the training set. The training set moves as the new data is available i.e., the model which dynamically adapts to the new patterns in the data.

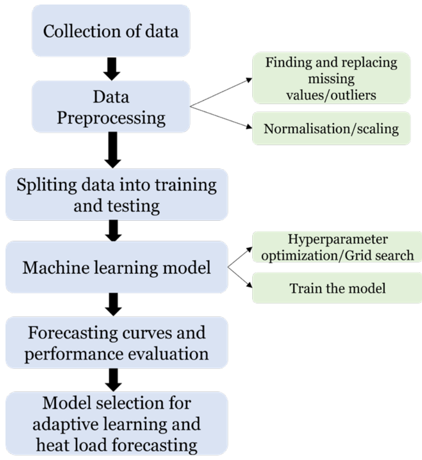

Figure 2: Workflow for machine learning model selection and heat load forecasting

3. Results

In this section, results are presented in two subsections: a) the performance of different models and selection for the testing data b) the performance of selected models in adaptive learning.

3.1 Performance comparison of ML models

The performance of all the ML models mentioned above is evaluated for the dataset of the Tillberga substation. The data is from the winter season from 1 January 2019 to 31 January 2019. Table 2 shows the performance of different ML models for the testing dataset. Amongst the linear models, GPR shows very good performance in prediction with 4.77% MAPE with LR and RR showing similar performance. GPR can capture the complex and non-linear relationship between input features and output. The high accuracy of GPR suggests that there might be a non-linear relationship between features and output.

Table 2: ML models and hyperparameters

ML model	MAPE, %	RMSE
LR (Linear Regression)	8.22	394.8
RR (Ridge Regression)	8.26	388.5
GPR (Gaussian Process Regression)	4.77	242.4
RF (Random Forest)	4.30	233.8
SVR (Support Vector Regressor)	4.28	225.8
LSTM (Long-Short-Term Memory)	5.50	279.9

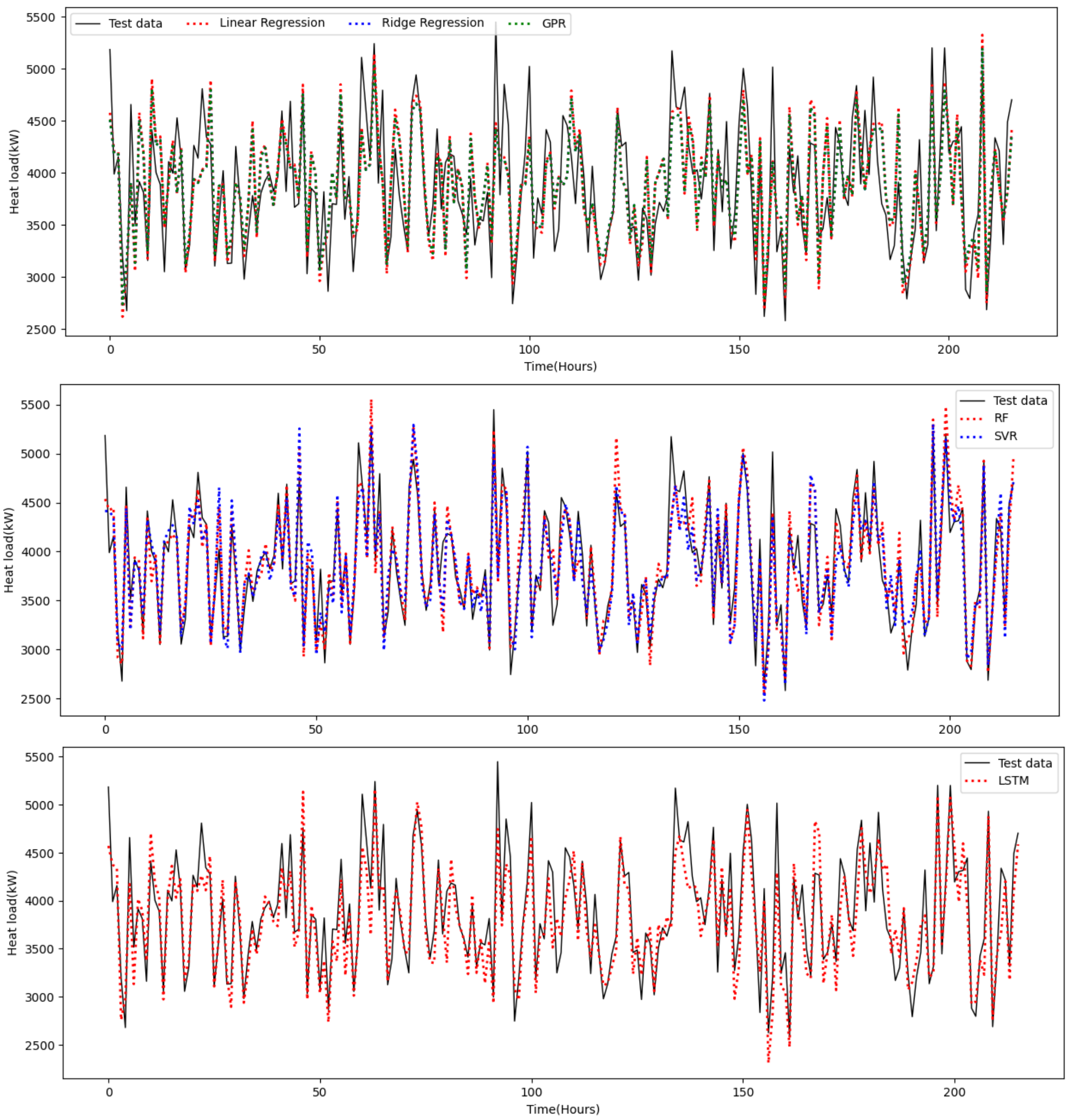

Figure 3: Performance comparison of actual and forecasted heat load for testing data for ML models

Table 1 also shows that the best-performing model is SVR with 4.28% MAPE which suggests that predicted heat loads with SVR are closer to the actual heat consumption as also seen in Figure 3. The best performance of the SVR over the other methods is associated with efficient feature space modelling and the fact that SVR is less prone to overfitting. It can be concluded that the GPR, RF and SVR can be effectively applied to the prediction of heat load in the DH system. In this study, we select SVR for adaptive training in real real-time scenario of heat load forecasting.

3.2 Adaptive modelling and performance

The SVR model is applied to the dataset of Tillberga and Skultuna region of Västerås city for adaptive learning and heat load prediction. The data is taken for the winter season from 1st January to 31st March 2022. The performance is evaluated for the adaptive training set of two weeks and four weeks. Figure 4 shows an example of the performance execution for the Tillberga region. The Fig, black line represents the measured data of the heat load of Tillberga against which predicted values (green line) are plotted. The training set is shown in the red line which adaptively moves with the addition of new data. The blue line represents the predicted horizon (6 hrs, 10 hrs, 24 hrs or 1 Week). The hours in the period are plotted on the x-axis which represents the total time duration.

Figures 5 and 6 show results of actual heat load and heat load prediction for 24 hrs, based on the SVR algorithm for Tillberga and Skultuna respectively. The figures also show the performance indicators (MAPE and RMS) for two different training set lengths i.e., 2-weeks and 4-weeks.

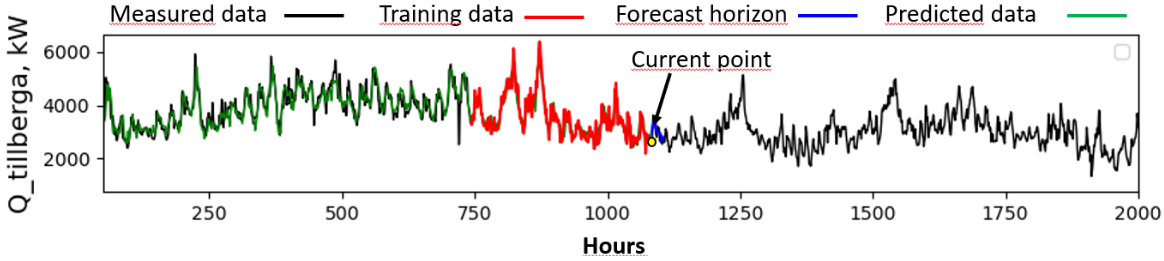

Figure 4: Adaptive learning of model to predict heat load

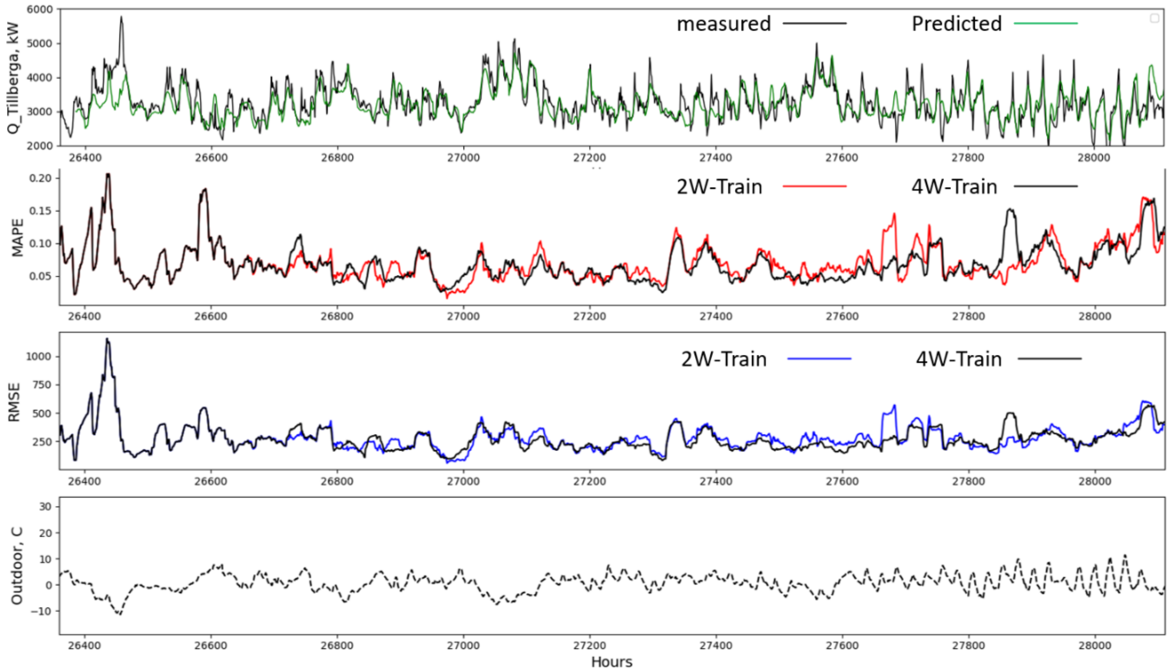

Figure 5: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Tillberga** region with data from 1st Jan 2022 to 31st March 2022 with **24 hr** of prediction horizon

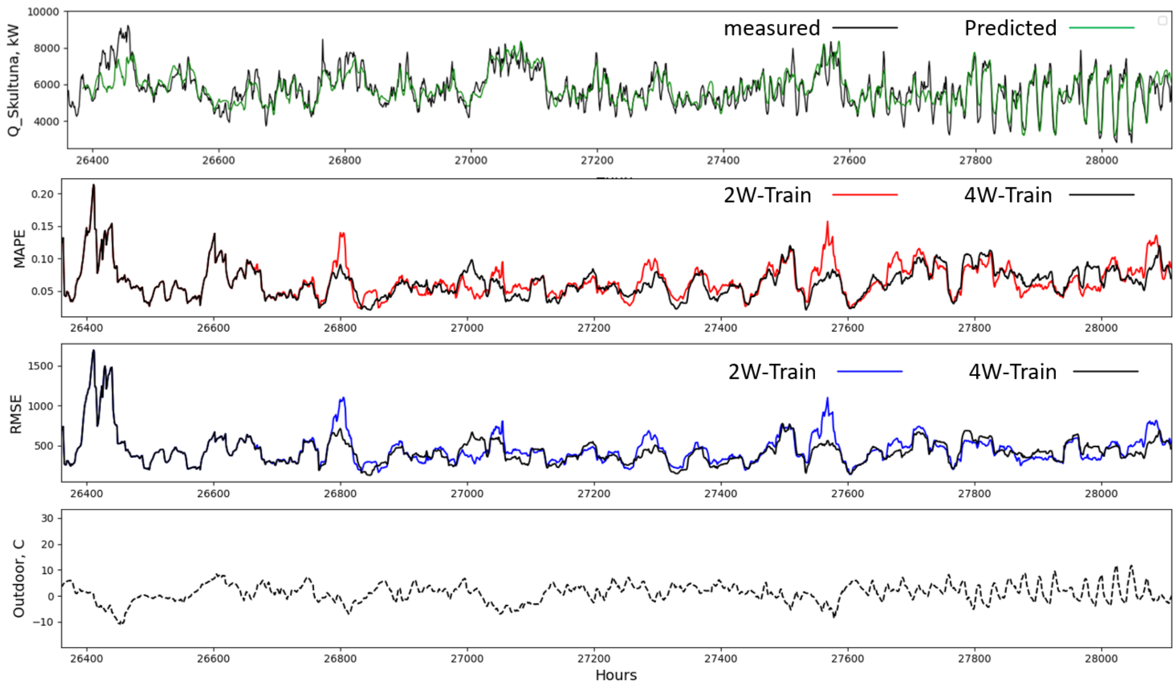

Figure 6: Forecasted heat load and Performance indicators (MAPE and RMSE) for the **Skultuna** region with data from 1st Jan 2022 to 31st March 2022 with **24-hour** prediction horizon

As can be seen from Figure 5, the predicted HL with SVR is closer to the actual energy consumption, with a MAPE value fluctuating between 3 to 10%. The high error at the initial predictions is due to the small training set available for the forecasting model. However, as the training set achieves the 2-week/4-week data, the accuracy improves. Also, the 4-week training data shows better performance than the 2-week training data which follows the intuition. The high error shown by the 2-week training data at around hour 27650 is due to the exclusion of the low outdoor temperature data. The 4-week training data captures the low outdoor temperature hence the more accurate predictions. The RMSE plot shows similar behaviour, and no deviations are observed. This confirms the good performance of the SVR for adaptive learning with good accuracy.

The model is adaptively updated as well as the training of the machine learning model in real time. It provides a "plug-and-play" solution for real-time prediction without significant pre-tuning requirements. This claim is validated by applying the SVR model to a completely different data set of another region i.e., Skultuna. Figure 6 shows the forecasted heat load and Performance indicators (MAPE and RMSE) for data from 1 Jan 2022 to 31 March 2022. The performance is quite like seen in the previous case of Tillberga region. The MAPE is within the range of 3-10%.

The performance discussed above was for 24 24-hour forecast horizon. The performance with 6 hours, 10 hours and 1 week is also evaluated and presented in the appendix section of the paper. It can be seen that for the short forecasting horizons, 6 hrs and 10 hrs, the errors are slightly higher than the long forecast horizon. The error in the long forecasting horizon seems to be more stable.

4. Conclusion

In this paper, six ML algorithms for the heat load prediction in the DH network of Västerås city are developed, compared and analyzed. The algorithms studied are LR, RR, GPR, RF, SVR and LSTM. Hyperparameter analysis is carried out to find the optimized values for each model. The performance of the models was compared using data from 2019 for the winter season. The predicted hourly results were compared with actual heat load data.

The SVR algorithm proved to be the most efficient one, producing the best performance in terms of MAPE and RMSE. The SVR model is then selected for adaptive learning and heat load forecasting in real-time scenarios. The model is tested for the actual data from the winter of 2022 of the Tillberga region. The results are also compared with shorter (2-week) and longer (4-week) training sets. The model, overall, shows good performance with MAPE ranging from 3 to 10%. To provide a "plug-and-play" solution for real-time prediction without significant pre-tuning requirements, the model is tested also with completely different regions on winter data. The performance shows similar accuracies to that in the Tillberga region. It proves that the developed SVR method is appropriate for adaptive learning and application in heat load prediction. In future, we intend to use this model to predict heat load in real-time scenario for the optimization of DH network supply and return temperature.

Acknowledgement

The authors would like to thank Mälarenergi for providing the data used for this work.

This research was performed as part of the "DISTRHEAT—Digital Intelligent and Scalable control for Renewables in HEAting neTworks" project, which received funding in the framework of the joint programming initiative ERA-Net Smart Energy Systems' focus initiative Integrated, Regional Energy Systems, with support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 775970.

References

Dalipi, F., Yildirim Yayilgan, S. and Gebremedhin, A. (2016) 'Data-Driven Machine-Learning Model in District Heating System for Heat Load Prediction: A Comparison Study', *Applied Computational Intelligence and Soft Computing*, 2016, pp. 1–11. Available at: https://doi.org/10.1155/2016/3403150.

Idowu, S. *et al.* (2014) 'Forecasting heat load for smart district heating systems: A machine learning approach', in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Venice, Italy: IEEE, pp. 554–559. Available at: https://doi.org/10.1109/SmartGridComm.2014.7007705.

Idowu, S. *et al.* (2016) 'Applied machine learning: Forecasting heat load in district heating system', *Energy and Buildings*, 133, pp. 478–488. Available at: https://doi.org/10.1016/j.enbuild.2016.09.068.

Idowu, S., Åhlund, C. and Schelén, O. (2014) 'Machine learning in district heating system energy optimization', in *2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*, pp. 224–227. Available at: https://doi.org/10.1109/PerComW.2014.6815206.

Ntakolia, C. *et al.* (2022) 'Machine learning applied on the district heating and cooling sector: a review', *Energy Systems*, 13(1), pp. 1–30. Available at: https://doi.org/10.1007/s12667-020-00405-9.

Rongali, S. *et al.* (2015) 'A context vector regression based approach for demand forecasting in district heating networks', in *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*. *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, pp. 1–6. Available at: https://doi.org/10.1109/ISGT-Asia.2015.7387141.

Schaefer, A.M., Udluft, S. and Zimmermann, H.-G. (2008) 'Learning long-term dependencies with recurrent neural networks', *Neurocomputing*, 71(13), pp. 2481–2488. Available at: https://doi.org/10.1016/j.neucom.2007.12.036.

scikit-learn: machine learning in Python — scikit-learn 1.2.2 documentation (2023). Available at: https://scikit-learn.org/stable/ (Accessed: 12 June 2023).

sklearn.preprocessing.MinMaxScaler (no date) *scikit-learn*. Available at: https://scikit-learn/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html (Accessed: 8 June 2023).

Suryanarayana, G. *et al.* (2018) 'Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods', *Energy*, 157, pp. 141–149. Available at: https://doi.org/10.1016/j.energy.2018.05.111.

TensorFlow (2023). Available at: https://www.tensorflow.org/ (Accessed: 12 June 2023).

Zhao, J., Li, J. and Shan, Y. (2021) 'Research on a forecasted load-and time delay-based model predictive control (MPC) district energy system model', *Energy and Buildings*, 231, p. 110631. Available at: https://doi.org/10.1016/j.enbuild.2020.110631.

Zimmerman, N., Kyprianidis, K. and Lindberg, C.-F. (2019) 'Achieving Lower District Heating Network Temperatures Using Feed-Forward MPC', *Materials*, 12(15), p. 2465. Available at: https://doi.org/10.3390/ma12152465.

Appendix

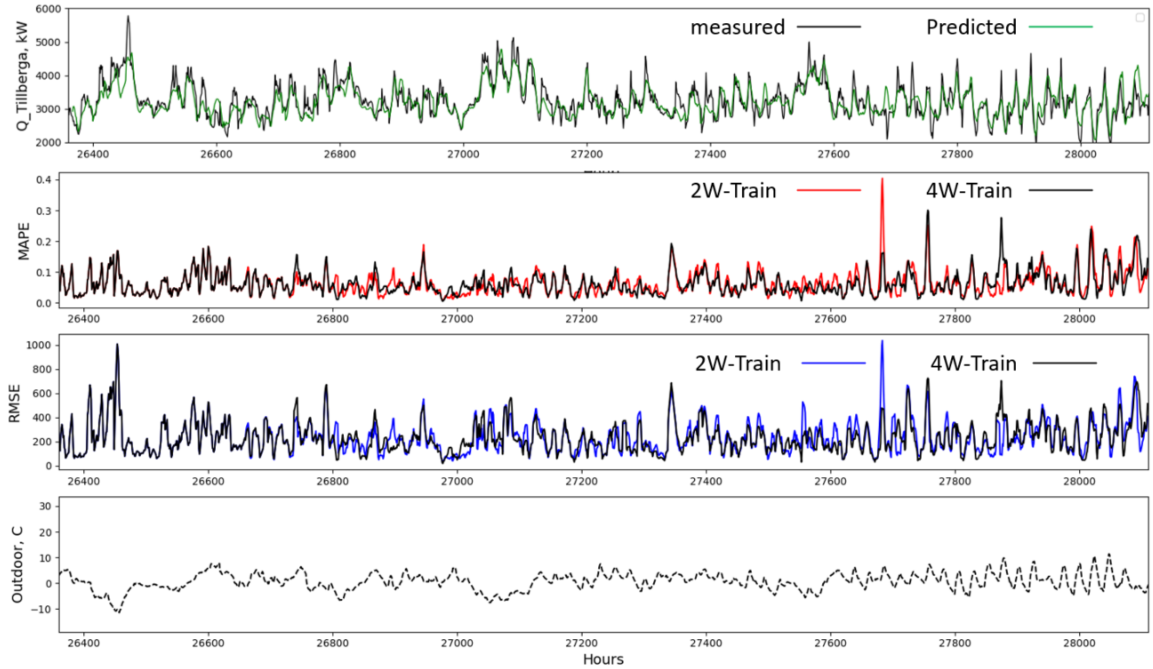

Figure 7: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Tillberga** region with data from 1st Jan 2022 to 31st March 2022 with 6 **hr** prediction horizon

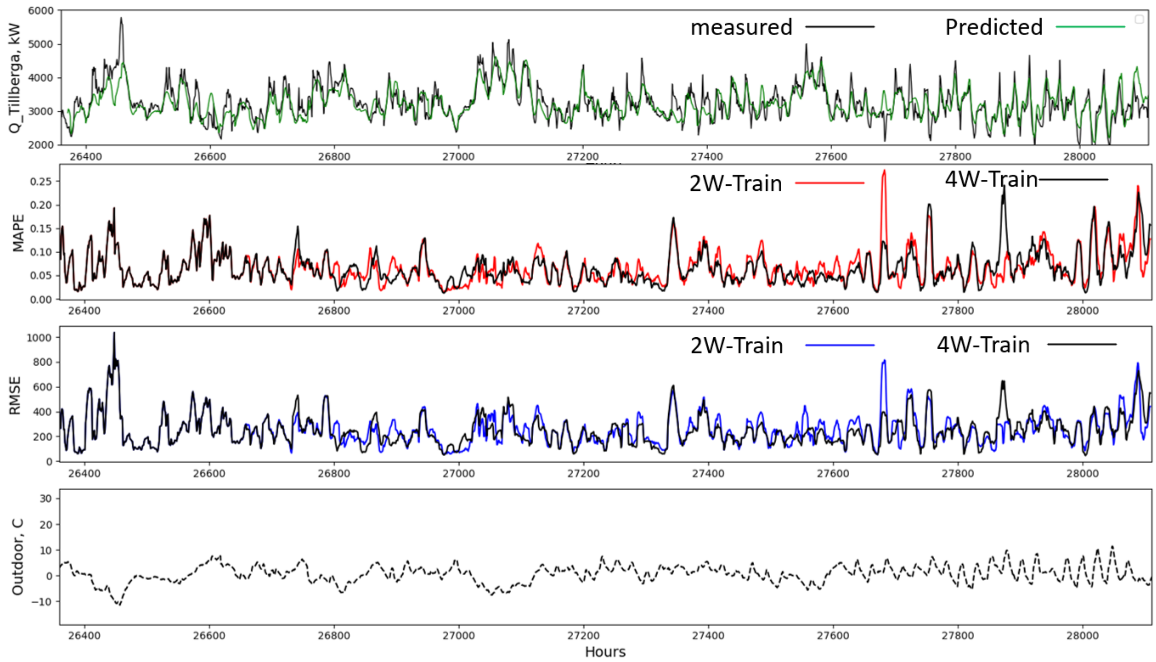

Figure 8: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Tillberga** region with data from 1st Jan 2022 to 31st March 2022 with **10 hr** prediction horizon

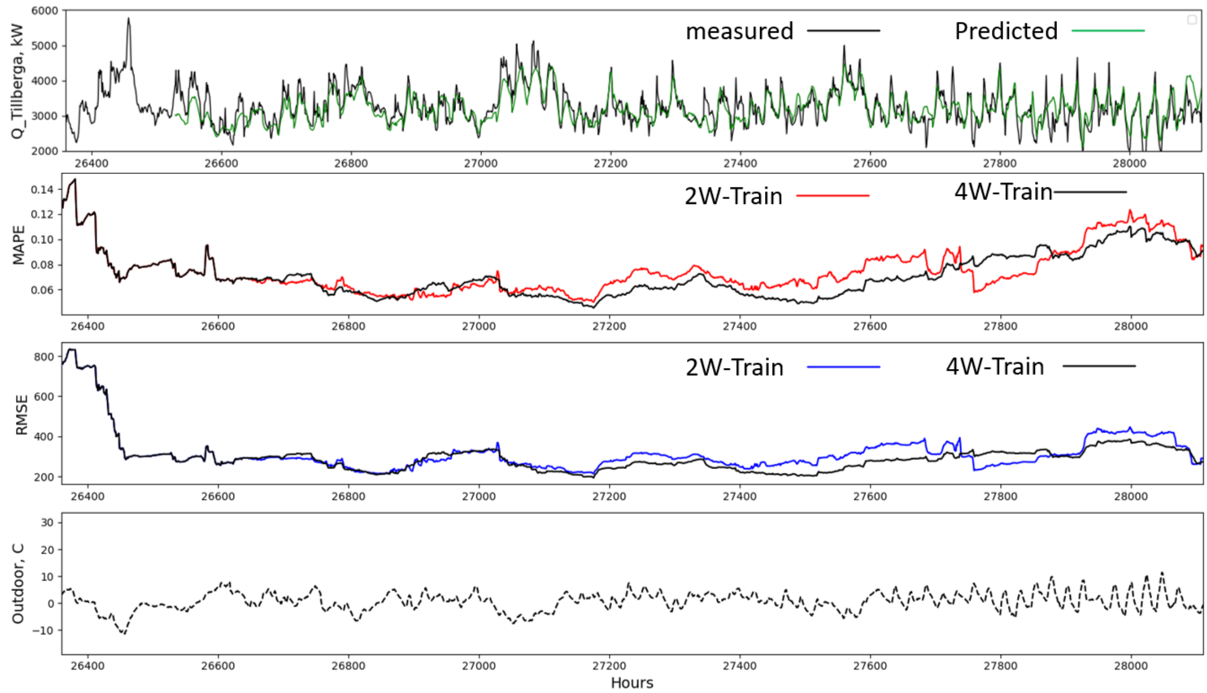

Figure 9: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Tillberga** region with data from 1st Jan 2022 to 31st March 2022 with **1-week** prediction horizon

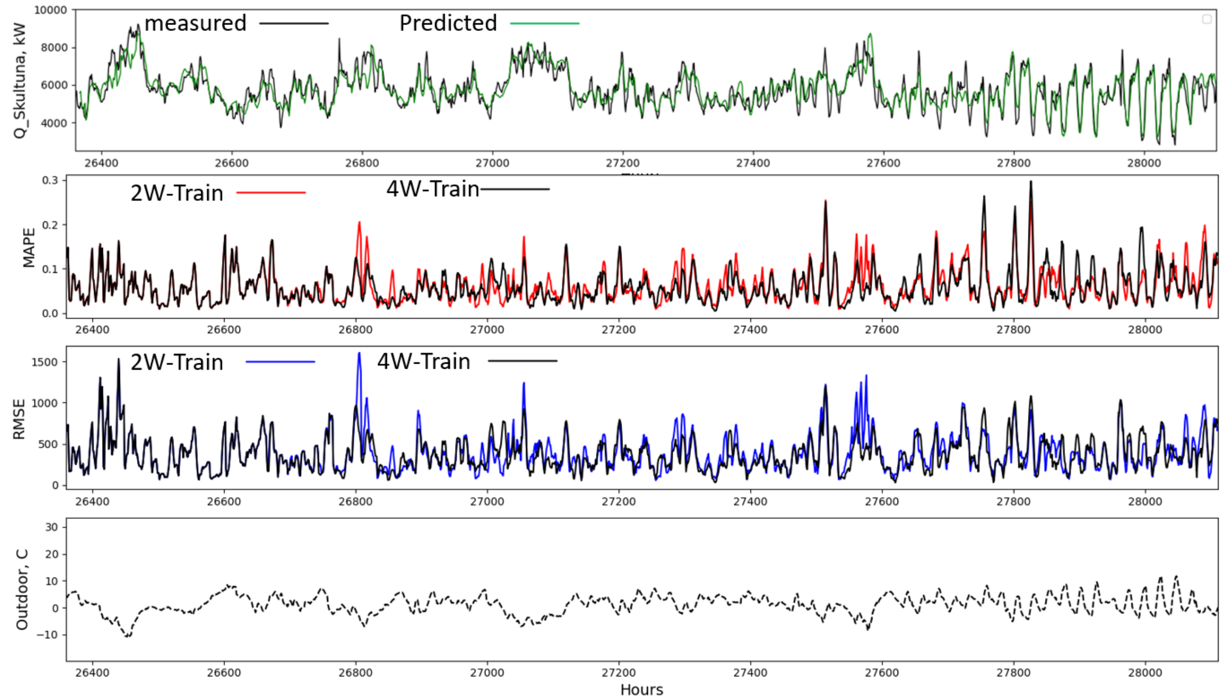

Figure 10: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Skultuna** region with data from 1st Jan 2022 to 31st March 2022 with 6 **hr** prediction horizon

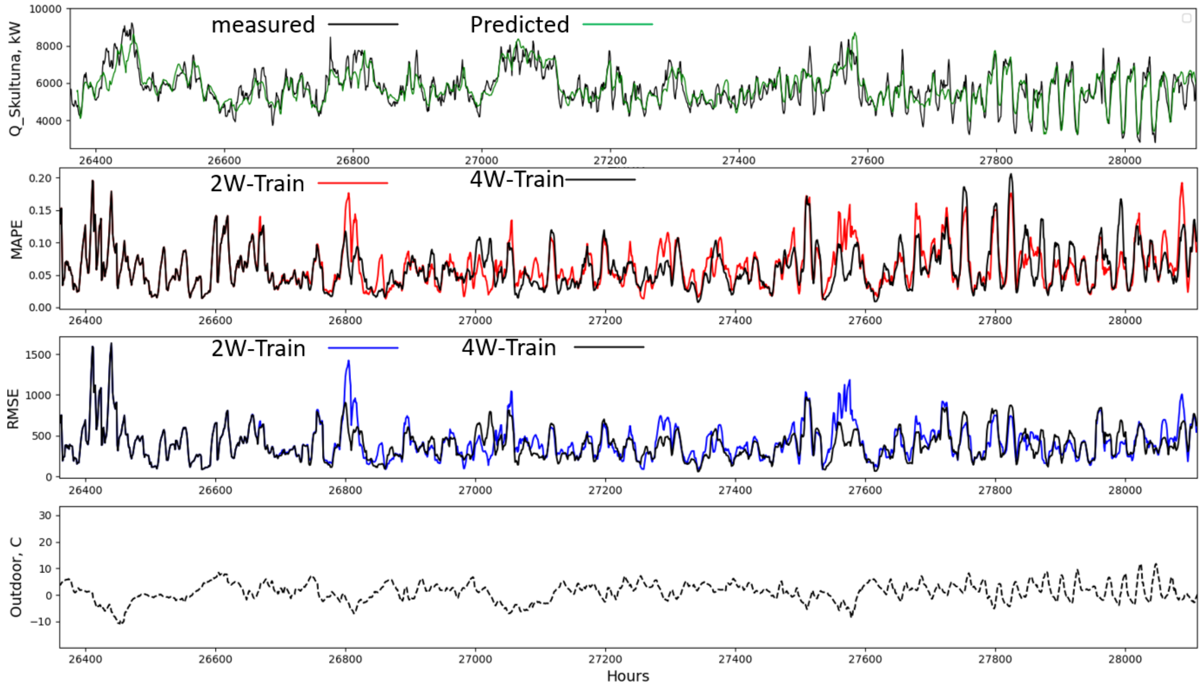

Figure 11: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Skultuna** region with data from 1st Jan 2022 to 31st March 2022 with **10 hr** prediction horizon

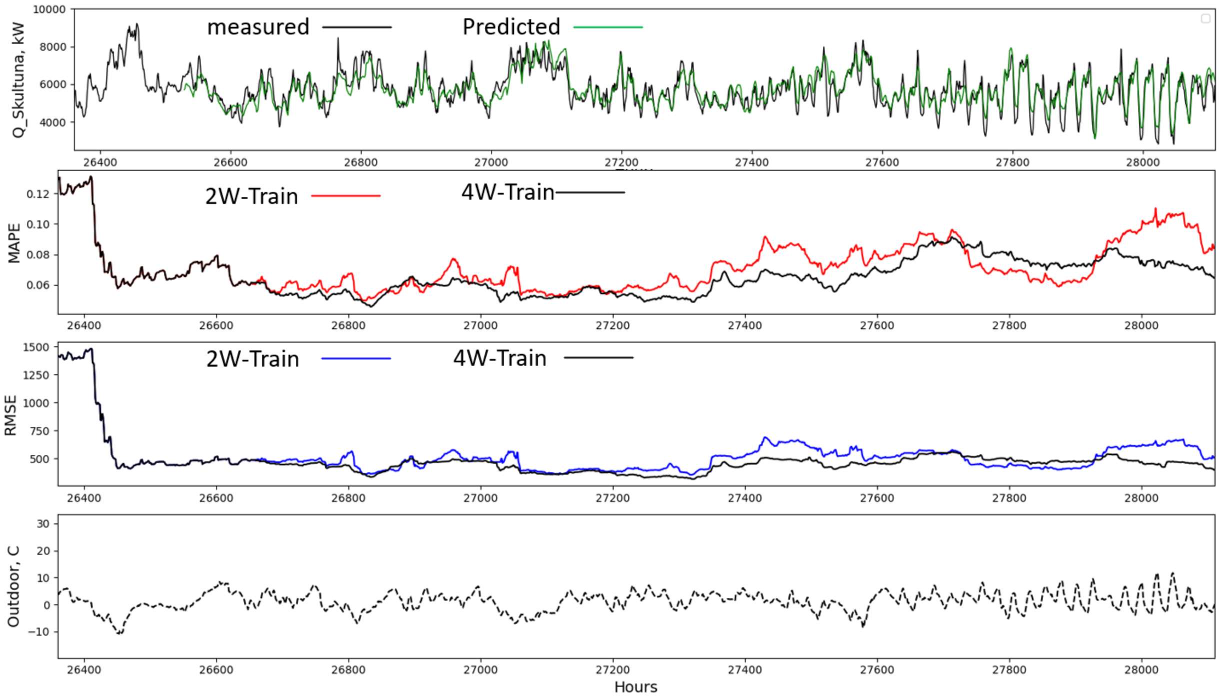

Figure 12: Forecasted heat load and Performance indicators (MAPE and RMSE) for **Skultuna** region with data from 1st Jan 2022 to 31st March 2022 with **1-week** prediction horizon