Intelligent epidemiological models for COVID-19

Esko K. Juuso

Control Engineering, Environmental and Chemical Engineering, Faculty of Technology,
University of Oulu, Finland
esko.juuso@oulu.fi

Abstract

The coronavirus COVID-19 is affecting around the world with strong differences between countries and regions. Extensive datasets are available for visual inspection and downloading. The material has limitations for phenomenological modelling but data-based methodologies can be used. This research focuses on intelligent modelling on the basis of these datasets. The methodology has been tested in the analysis of daily new confirmed COVID-19 cases and deaths in six countries. The datasets are studied per million people to get comparable indicators. Nonlinear scaling brings the data of different countries to the same scale and linear interactions represent the varying operating conditions well. The same approach operates for both the confirmed cases and deaths and can be used for any country or group of people. The effects of the vaccinations were clearly shown at the end of the analyzed period. During the pandemic, the scaling functions expanded for the confirmed cases but remained practically unchanged for the confirmed deaths which is consistent with increasing testing. Limitations are seen if there are too many interacting things, e.g. several infection transmission chains which are in different stages. The feasibility analysis needs to be extended to the modelling with inputs. The presented approach is promising for this wider analysis.

Keywords: intelligent methods, temporal analysis, dynamic modelling, digital twins, COVID-19

1 Introduction

The coronavirus COVID-19 is affecting around the world. There are strong differences between countries and regions. People of all ages can be infected but older people and people with pre-existing medical conditions are more vulnerable to becoming severely ill. The risk is presented with three parameters:

- Transmission rate evaluated by the number of newly infected people,
- Case fatality rate (CFR) based on the percent of cases that result in death,
- Vaccine performance as a prevention measure.

An online interactive dashboard is hosted by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University for visualising and tracking reported cases of *coronavirus disease 2019 (COVID-19)* in real time (CSSE, 2021; Dong et al., 2020). Transmission dynamics is difficult to explain since the characteristics of a novel disease include many uncertainties. The open evidence review (Jefferson et al., 2021) makes information about active research on modes of transmission available.

The effective reproduction number (R) of an infectious disease is used for modelling. The tracking of the parameter is done by assuming a model structure. An example of this approach is presented in (Arroyo-Marioli et al., 2021) where the Kalman filter and a SIR model has been used for tracking R for COVID-19.

The steady-state simulation models are linear *multiple input, multiple output (MIMO)* models $\vec{y} = F(\vec{x})$, where the output vector $\vec{y} = (y_1, y_2, \ldots, y_n)$ is calculated by a linear function F from the input vector $\vec{x} = (x_1, x_2, \ldots, x_m)$. Statistical modelling in its basic form uses linear regression for solving the model coefficients. Linear methodologies are suitable for large multivariable systems and can be extended with quadratic and interactive terms response surface methodologies (Box and Wilson, 1951). Principal components compress the data by reducing the number of dimensions with a minor loss of information (Jolliffe, 2002). Partial least squares regression (PLS) is an extension of these ideas (Gerlach et al., 1979). Known semi-physical models of inputs are important in linear modelling, see (Ljung, 1999). In linear parameter varying (LPV) models, an exogeneous variable measured during the operation, modifies the local linear models (Ljung, 2008).

Dynamic data-driven modelling with parametric models, also known as identification (Ljung, 1999), is the key methodology in the dynamic modelling. These models use the static mapping and NARX/Nonlinear AutoRegressive with eXogenous structures with a finite number of inputs and outputs. The dynamic structures are reduced in dynamic models based on fuzzy set systems or neural networks (Babuška and Verbruggen, 2003).

Generalized norms are used in data analysis to extract features from waveform signals collected from the statistical databases (Lahdelma and Juuso, 2011). 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 for the norms of individual samples. This means that norms can be recursively updated (Juuso, 2011). The same methodologies can be

DOI: 10.3384/ecp21185505

used for analysing the data distributions in less frequent data, e.g. daily COVID-19 data. Distributions of the variables provide useful information about fluctuations, trends and models. This has been used in temporal analysis for all types of measurements, features and indices. Recursive updates of the parameters are needed in prognostics. (Juuso, 2020)

Dynamic LE models use the static mapping and NARX input models in the same way as fuzzy set systems and neural networks. The main difference is that the input and output variables are processed by a nonlinear scaling method, which originates from the membership functions used in fuzzy systems. (Juuso and Leiviskä, 1992; Juuso, 2004) Constraints handling (Juuso, 2009) and data-based analysis (Juuso and Lahdelma, 2010), improve possibilities to update the scaling functions recursively (Juuso, 2011). Different fuzzy approaches can be efficiently combined with LE models where the interactions between the scaled variables are linear (Juuso, 2014).

This research aims to develop unified intelligent models for analyzing the fluctuations, trends and severity of the corona situations. Parametric systems are used to adapt the solution for varying operating conditions caused by local areas and groups of people. Recursive updates are used in the parametric models.

2 COVID-19 data

This research uses the complete COVID-19 dataset maintained by Our World in Data. The collection of the COVID-19 data is updated daily and includes data on confirmed cases, deaths, hospitalizations, and testing. Raw data on confirmed cases and deaths for all countries is sourced from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Data visualizations rely on work from many people and organizations (Ritchie et al., 2020).

The Our World in Data has created a new description of all our data sources available at the GitHub repository where all the data can be downloaded. These datasets were used as a data source in this research. The collection of data is presented as tabular data where every column of a table represents a particular variable, and each row corresponds to a given record of the data set for a specific country on a certain day. Each record consists of one or more fields, separated by commas. The data can be visualised in the COVID-19 DataExplorer for individual countries. Several countries can be compared by selecting them for the view. The maps available in DataExplorer help in focusing on the analysis.

The analysis uses confirmed COVID-19 cases whose number is lower than the number of actual cases. The main reason for that is limited testing which also varies between countries and time. Therefore, the analysis is done country-wise. The pandemic introduces an increasing number of new COVID-19 cases but countries also make progress in reducing the speed towards zero new cases (Figure 1). However, the increase can start again as can be seen in the data of different countries. The pandemic can restart if it is active somewhere. The difficult periods vary between countries.

A part of the pandemic cases leads to hospitalizations and deaths. Both increases and reductions can be seen in the daily new confirmed COVID-19 deaths (Figure 2. During the outbreak of the pandemic, the calculated case fatality rate (CFR) was a poor measure of the mortality risk since it depends on the number of tests and at that time there were few tests. The true number of cases was much higher. Later the number of tests has increased strongly, but not in all countries.

An increasing number of variants and mutations has effects on the number of cases. Vaccinations were just started during the studied period. All these have strong effect on the dynamics of the pandemic. The problems become more case specific but can in the same time activate in many locations.

The research focused on the temporal analysis is aimed on finding situations for more detailed modelling and action planning.

3 Methodologies

The modelling needs to be adapted in the appropriate situations. The unified analysis requires that all the features are in the same scale. In this research, this is done by combining the nonlinear scaling and the intelligent temporal analysis. This methodology allows recursive updates of the scaling functions.

3.1 Nonlinear scaling

The nonlinear scaling brings various measurements and features to the same scale by using monotonously increasing scaling functions $x_j = f(X_j)$ where x_j is the variable and X_j the corresponding scaled variable. The function $f()$ consist of two second order polynomials, one for the negative values of X_j and one for the positive values, respectively. The corresponding inverse functions $X_j = f^{-1}(x_j)$ based on square root functions are used for scaling to the range [-2, 2], denoted as linguistification. The monotonous functions allow scaling back to the real values by using the function $f()$. (Juuso, 2004)

The parameters of the functions are extracted from measurements by using generalized norms and moments. The support area is defined by the minimum and maximum values of the variable, i.e. a specific area for each variable $j, j = 1, \ldots, m$. 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.

The corner points are defined by iterating the orders, p,

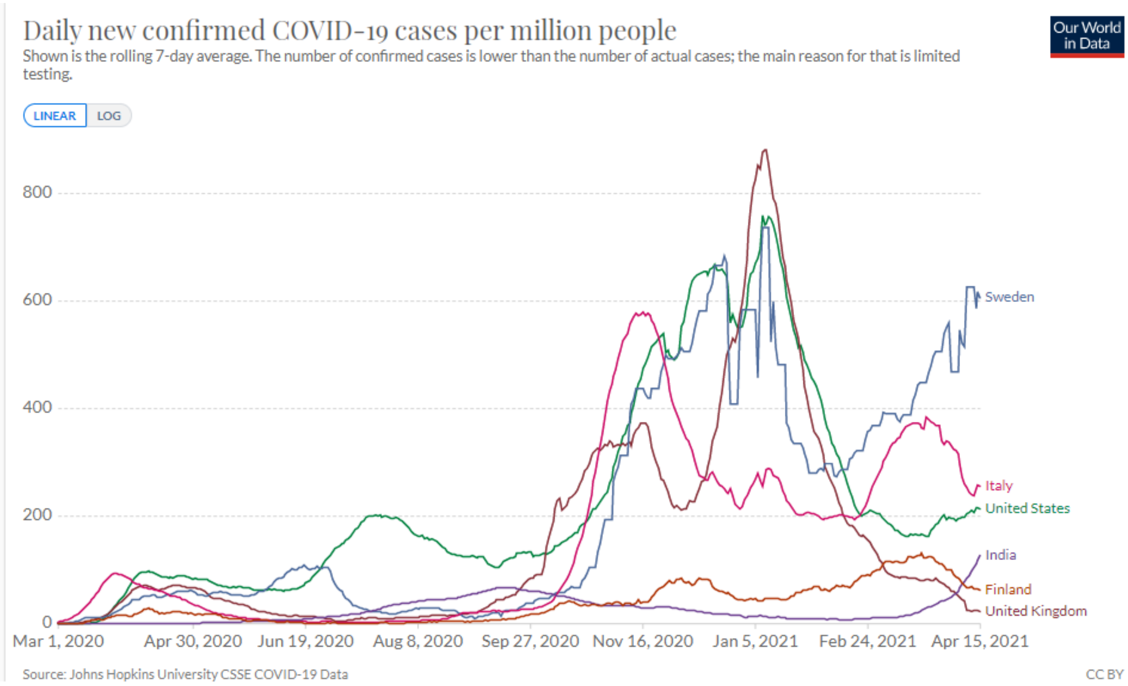

Figure 1. Daily new confirmed COVID-19 cases per million people, rolling 7-day averages collected from (Our, 2021) for selected countries.

of the corresponding generalised norms

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

where $p \neq 0$, is calculated from N values of a sample, τ is the sample time. This provides possibilities to recursively update the scaling functions since the generalized norms can be recursively updated. The iteration is based on the generalized skewness (Juuso and Lahdelma, 2010).

The scaled values should preserve the directions of the temporal changes with time. To achieve this, the scaling functions should be monotonously increasing. This is achieved by limiting the ratios,

$$\begin{aligned} \alpha_j^- &= \frac{(c_l)_j - \min(x_j)}{c_j - (c_l)_j}, \\ \alpha_j^+ &= \frac{\max(x_j) - (c_h)_j}{(c_h)_j - c_j}, \end{aligned} \tag{2}$$

in the range $[\frac{1}{3}, 3]$. The corner points are adjusted if these limitations are not filled. There are several alternatives to select the points to tune. (Juuso, 2009)

The second order polynomials,

$$\begin{aligned} f_j^- &= a_j^- X_j^2 + b_j^- X_j + c_j, \quad X_j \in [-2, 0), \\ f_j^+ &= a_j^+ X_j^2 + b_j^+ X_j + c_j, \quad X_j \in [0, 2], \end{aligned} \tag{3}$$

are monotonously increasing if the coefficients are defined as follows:

$$\begin{aligned} a_j^- &= \tfrac{1}{2}(1 - \alpha_j^-)\, \Delta c_j^-, \\ b_j^- &= \tfrac{1}{2}(3 - \alpha_j^-)\, \Delta c_j^-, \\ a_j^+ &= \tfrac{1}{2}(\alpha_j^+ - 1)\, \Delta c_j^+, \\ b_j^+ &= \tfrac{1}{2}(3 - \alpha_j^+)\, \Delta c_j^+, \end{aligned} \tag{4}$$

where $\Delta c_j^- = c_j - (c_l)_j$ and $\Delta c_j^+ = (c_h)_j - c_j$.

3.2 Steady-state LE modelling

The nonlinear scaling transforms the nonlinear problem $\vec{y} = F(\vec{x})$ to a linguistic equation (LE) model represented by a compact equation

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

where the functions f_j and f_{out} are scaling functions. Each variable j has its own time delay n_j compared to the variable with latest time label. In the general case, the the weight factors

$$w_{ij} = -\frac{A_{ij}}{A_{i\,out}}. \tag{6}$$

The coefficients A_{ij} and $A_{i\,out}$ can be set to one or to a chosen value by modifying the scaling functions.

The directions of the interactions analyzed with these methodologies are aimed to be valid in a wider area than the scaling functions of the individual model variables. The quadratic effects are embedded in the scaling approach and the model can handle various kinds of interactions. Subsets of scaled variables can form linguistic principal components (LPCs) and PLS regression could be useful in modelling. The LPV modelling further extends the feasible areas of the model parameters.

3.3 Dynamic LE modelling

In NARX models, the input and output values are chosen according to appropriate system orders. In the regressor

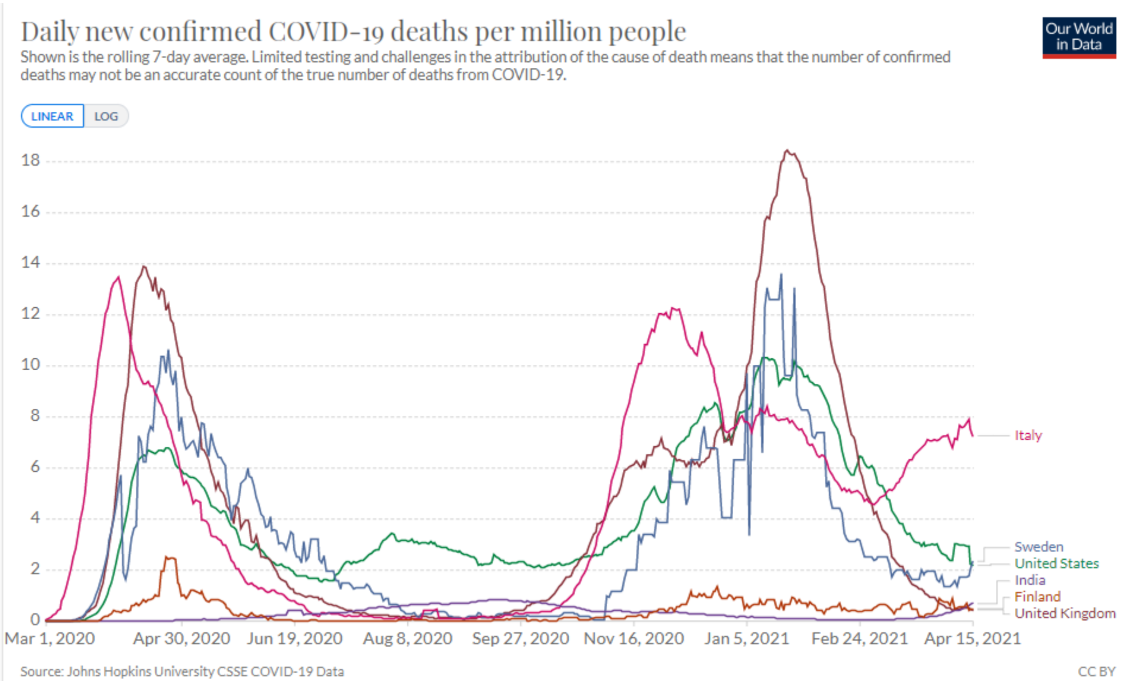

Figure 2. Daily new confirmed COVID-19 deaths per million people, rolling 7-day averages collected from (Our, 2021) for selected countries.

vector, the number of past inputs and outputs may become too high if nonlinear effects are needed. The nonlinear scaling reduces the number of input and output signals needed for the modelling of nonlinear systems. In the default dynamic LE model,

$$Y_{out}(k)+a_1Y_{out}(k-1)=\sum_{j=1,j\neq out}^{m} b_jU_j(k-n_j)+e(t), \quad (7)$$

where $Y_{out}(k)$, $Y_{out}(k-1)$ and $U_j(k-n_j)$ are the scaled values of the variables and $e(t)$ is a noise term. The scaling function f_{out} is not changing between time steps $k-1$ and k. The delays n_j of the inputs can depend on the operating conditions and need to be calculated if variable step numerical integration is used.

A wide variety of situations can be represented with composite local models which are based on the same model equations (7) and case specific scaling functions f_j and f_{out}. Overlapping local models can be combined with fuzzy set systems by using case specific solutions which are first scaled back to the real values.

4 Epidemiological modelling

Data collection and working practices were under development during the first months of the COVID-19 pandemic. In this research, the modelling was started with the variable selection and analysis in the varying operating conditions.

4.1 Variable selection

New cases. The modelling uses daily new confirmed COVID-19 cases and deaths as output variables. The study was done for the same countries as in (Juuso, 2021). Six countries were Finland, India, Italy, Sweden, United Kingdom and United States.

The rolling 7-day averages were used for the research study since they operated smoothly for the confirmed deaths as well. The cases were analyzed per million people to improve the sensitivity of the analysis for small countries. Situations vary strongly between countries and periods of time.

Hospital patients, hospital admissions and intensive care patents could be analyzed with the same methodologies, but the datasets are much more limited.

The datasets include also test activities, vaccination and stringency information.

Operating conditions. The number of tests was everywhere very low in the beginning of the pandemic in the spring of 2020. The confirmed new cases remained low for the first 240 days (Juuso, 2021) for all the countries compared in this research (Figure 1). Since similar differences were not seen in new confirmed deaths (Figure 2), the severe cases were detected already in the beginning. A high number of undetected cases is in agreement with the high values of excess mortality.

COVID-19 was activated country by country. In this research, a high number of cases per million people appeared in the sequence Italy, the UK, Sweden and the USA. Finland had a low number of cases and India hardly any. During the summer of 2020, both new cases and deaths went down, although the number of tests was increased. An exception was the USA where both remained high, which was true situation since the share of positive tests was high as well. India had all the time very few cases, but the test activity was very low and the share of positive tests was high meaning that the situation was similar with the spring of 2020 in other countries.

In the autumn of 2020, the number of cases and deaths started to rise again in the same sequence as previously. The number cases was rising fast and the number of deaths only slightly.

Vaccinations started in December 2020 and increased steadily, fastest in the UK and USA (Table 1) where the number of new cases dropped fast. The decrease continued although the number of tests was increased.

Table 1. Cumulative COVID-19 tests and vaccinations per 100 people by April 12, 2021 (Our, 2021).

	Tests	One dose	Two doses
Finland	76.1	20.93	1.79
India	18.5	6.82	0.97
Italy	88.2	15.6	6.65
Sweden	77.2	15.18	6.26
United Kingdom	189.0	47.28	11.22
United States	120.0	39.28	24.89

The society can take actions to prevent the pandemic to expand. The stringency index presented in (Hale et al., 2021) is a composite measure based on nine response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest). This index was highest in India and lowest in Finland.

The reproduction rate represented as the average number of new infections caused by a single infected individual. If the rate is greater than 1, the infection is able to spread in the population. If it is below 1, the number of cases occurring in the population will gradually decrease to zero. The estimate rates (Arroyo-Marioli et al., 2021) were very high in the spring of 2020 since the detected COVID-19 cases were almost all serious cases.

4.2 Data analysis

The nonlinear scaling approach aims to simplify the modelling work. The normalisation keeps the directions of the effects but would leave the analysis of nonlinear effects to the modelling. For all six countries, the data analysis was taken from the research on temporal analysis (Juuso, 2021). The country specific lines were extracted from the full dataset with all the countries by using the DataExplorer. Selected time periods were used in the analysis.

Within each country, the risk levels are represented by using nonlinear scaling. The scaling functions are defined by five corner points by generalized norms whose orders are obtained from the data. The parameters of the scaling functions are country and time period specific (Figures 3 and 4). The differences between the early and later stages of the pandemic are very clear. For the death cases, the differences are hardly visible.

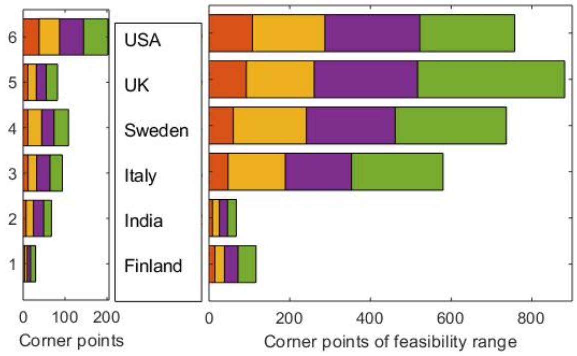

Figure 3. Parameters $\min(x_j)$, $(c_l)_j$. c_j, $(c_h)_j$ and $\max(x_j)$ for the daily new confirmed cases: first 240 days (left) and all data (right) , all based on 7-day rolling average of confirmed cases per million people (Juuso, 2021).

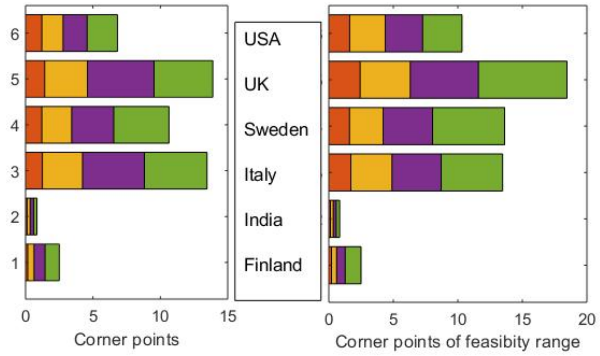

Figure 4. Parameters $\min(x_j)$, $(c_l)_j$. c_j, $(c_h)_j$ and $\max(x_j)$. for the daily new confirmed deaths: first 240 days (left) and all data (right), all based on 7-day rolling average of confirmed cases per million people (Juuso, 2021).

The confirmed cases and deaths were analysed per million people but still the relative values are still much higher in the UK, the USA and Italy than in Finland.

4.3 Feasibility results

The default dynamic LE model (7) is studied for all the analyzed data. The cases were the same as in the temporal analysis presented in (Juuso, 2021). The first phase is the nonlinear scaling of the variables. In this research, the variables are the daily new confirmed COVID-19 cases (Figure 1) and deaths (Figure 2).

The input variables could be the vaccination level, stringency index and active level of infections. Vaccinations were just coming at the end of the studied period. The

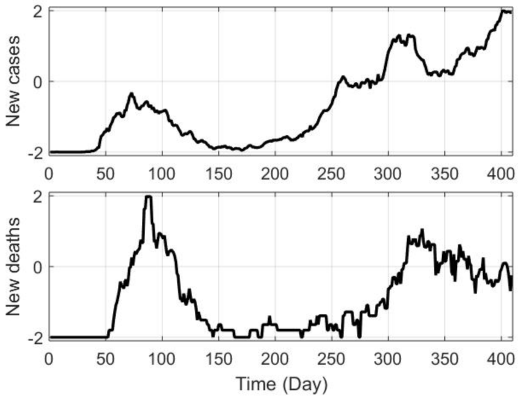

Figure 5. Scaled daily COVID-19 cases and deaths in Finland.

stringency index was not yet on a practical level. There were not measurements for the active level of infections.

This section is a feasibility study of the linear interactions of the scaled data in different countries having different operating conditions.

Finland. During the first days of the pandemic, the levels of the daily new cases were much lower than in Autumn and Winter. In spite of that, the levels of deaths were already at the same level as later. For Finland, this is seen in Figure 5 and 6. The difference in operating conditions is clear in the parameters of the scaling functions (Figures 3 and 4).

Linear interactions operate well for the changes in the scaled values. The time delay from the changes of confirmed cases to the confirmed deaths is about two weeks. The steep slope at the beginning of the cases could be caused by the delayed detection of them. The analysis reveals details which are hardly visible in Figures 1 and 2, including a drastic change in deaths seen in Figures 5 and 6). On December 2020, there was a maximum area for both new cases and deaths. After that, the new infections were increasing but the new deaths went down (Figure 5).

Italy. Italy was first country in Europe to meet the pandemic which may have cased the very slope seen in the linear model (Figure 7). There was also a very short time delay between the detected cases and deaths. Later the time delay went to the same two weeks which was in Finland. For Italy, the scaling functions were much steeper than for Finland (Figures 3 and 4). There was a maximum area slightly earlier than in Finland and the infection levels remained constant. The new deaths went down.

Sweden. Sweden had considerably softer actions for controlling the pandemic. According to the conformed deaths, the first period of the pandemic lasted about 50 days longer than in Finland and Italy (Figure 8). There were also a high level of infections during the whole Summer. The scaling functions were comparable with the Italian ones. Time delays between confirmed cases and deaths

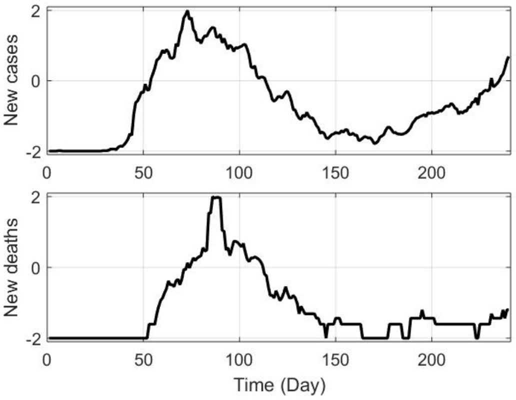

Figure 6. Scaled daily COVID-19 cases and deaths in Finland during the first 240 days.

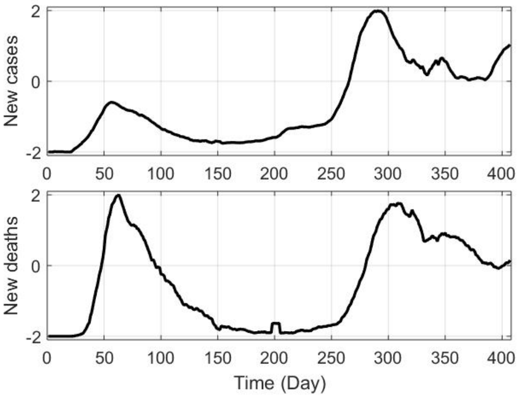

Figure 7. Scaled daily COVID-19 cases and deaths in Italy.

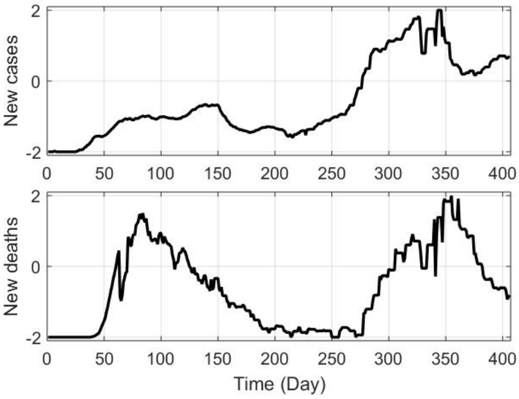

Figure 8. Scaled daily COVID-19 cases and deaths in Sweden.

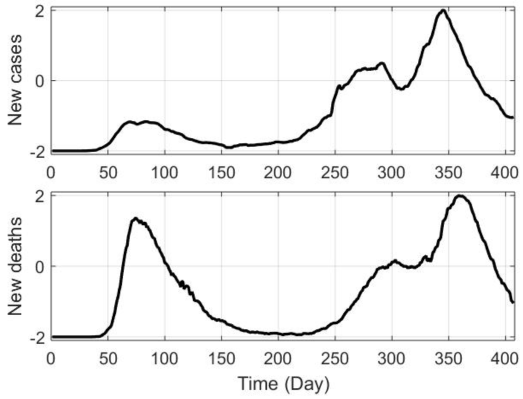

Figure 9. Scaled daily COVID-19 cases and deaths in the United Kingdom.

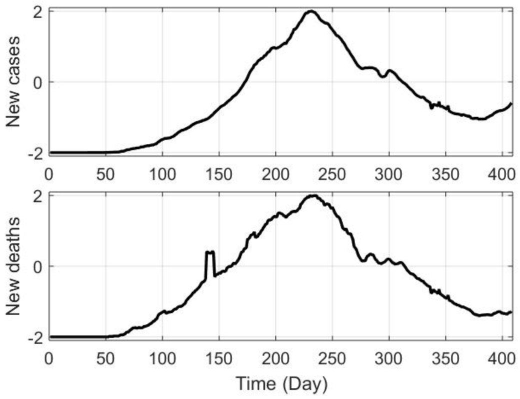

Figure 11. Scaled daily COVID-19 cases and deaths in India.

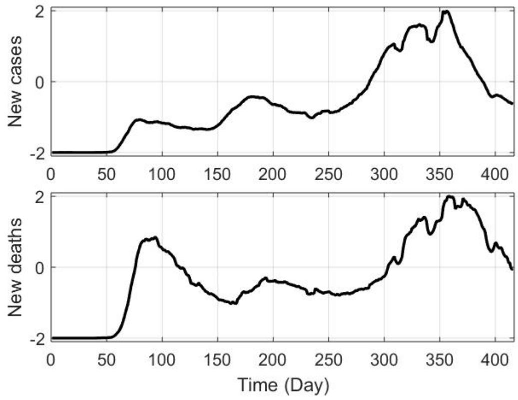

Figure 10. Scaled daily COVID-19 cases and deaths in the United States.

were not easy to find since the level of infections was high continuously.

United Kingdom. The scaling functions are the steepest for the United Kingdom (Figures 3 and 4). The first period of the pandemic went very closely in the same way as in Finland and Italy. In the autumn of 2020, a short maximum area was slightly before the similar period in Italy. This was followed by a short maximum period and a very steep decrease for both the new cases and deaths (Figure 9. The decrease is considerably faster than in the first pandemic period. This improvement is linked with the widening vaccination (Table 1).

USA. Scaling functions are quite steep already in the first period. The first period of the pandemic started in the same way as in Finland, Italy and the UK (Figure 10). Summer 2020 was different: the recovery was slow, even slower than in Sweden. The level of new infections and deaths were all the time higher than in other countries. Three peaks were detected in November-December 2020 could be from different infection transmission chains. After them, the level of new infections go down almost as fast as in the UK. Also, here this improvement is linked with the widening vaccination (Table 1). The time difference between the last peaks of the infections and the deaths is again close to two weeks.

India. India was in the starting phase: the parameters of the scaling functions have very low values (Figures 3 and 4 and the time difference between the peaks of the infections and the deaths are small (Figure 11).

5 Discussions

The nonlinear phenomena are well presented with the combination of the nonlinear scaling and linear interactions for different countries in varying operating conditions. The same approach operates for both the confirmed cases and deaths. Different countries and specific periods, including periods where the data quality is not sufficient, are analyzed in the same scale $[-2, 2]$.

The effects of the vaccinations at the end of the analyzed period are clear, especially the UK and the USA but also Finland, Italy and Sweden where the vaccinations were just started.

Limitations are seen if there are too many interacting things, e.g. several infection transmission chains which are in different stages. Situations of this kind of are active in the summer of 2020 in Sweden and the USA. Three infection chains in November-December 2020 are interacting. The actual inputs would be needed in those situations.

6 Conclusions and future studies

The combination of the nonlinear scaling and linear interactions operates well for different countries in varying operating conditions. This aggregated material was used for analyzing some countries. The analysis can be done with this approach for any country which has data in the overall dataset.

The analysis can be done in similarly for different subsets. Specific scaling functions can be used in local analysis and for groups of people to increase the sensitivity of the temporal analysis. The data material already includes hospital patients and patients in intensive care. The progress in people vaccinations provides more material for comparisons. Also, different variants can be taken in the analysis.

References

Our world in data. `https://ourworldindata.org/`, 2021. Accessed: 2021-07-10.

F. Arroyo-Marioli, F. Bullano, S. Kucinskas, and C. Rondón-Moreno. Tracking r of COVID-19: A new real-time estimation using the kalman filter. *PLoS ONE*, 16(1):1–16, 2021.

R. Babuška and H. Verbruggen. Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control*, 27 (1):73–85, 2003.

G. E. P. Box and K. B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B*, 13(1):1–45, 1951.

CSSE. COVID-19 data repository by the center for systems science and engineering (CSSE) at Johns Hopkins University. `https://github.com/CSSEGISandData/COVID-19`, 2021. Accessed: 2021-07-10.

E. Dong, H. Du, and L. Gardner. An interactive web-based dashboard to track covid-19 in real time. *Lancet Inf Dis.*, 20(5): 533–534, 2020. doi:10.1016/S1473-3099(20)30120-1.

R. W. Gerlach, B. R. Kowalski, and H. O. A. Wold. Partial least squares modelling with latent variables. *Anal. Chim. Acta*, 112(4):417–421, 1979.

Thomas Hale, Noam Angrist, Rafael AU Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. A global panel database of pandemic policies (oxford covid-19 government response tracker. *Nature Human Behaviour*, 5:529–538, 2021. doi:10.1038/s41562-021-01079-8.

T. Jefferson, E. A. Spencer, A. Plüddemann, N. Roberts, and C. Heneghan. Transmission dynamics of COVID-19: An open evidence review. `https://www.cebm.net/evidence-synthesis/transmission-dynamics-of-covid-19/`, 2021. Accessed: 2021-07-10.

I. T. Jolliffe. *Principal Component Analysis*. Springer, New York, 2 edition, 2002. 487 pp.

E. Juuso and S. Lahdelma. 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, pages 1358–1372, 2010. URL `www.scopus.com`.

E. K. Juuso. Integration of intelligent systems in development of smart adaptive systems. *International Journal of Approximate Reasoning*, 35(3):307–337, 2004. doi:10.1016/j.ijar.2003.08.008.

E. K. Juuso. Tuning of large-scale linguistic equation (LE) models with genetic algorithms. In M. Kolehmainen, editor, *Revised selected papers of the International Conference on Adaptive and Natural Computing Algorithms - ICANNGA 2009, Kuopio, Finland, Lecture Notes in Computer Science*, volume LNCS 5495, pages 161–170. Springer-Verlag, Heidelberg, 2009. doi:10.1007/978-3-642-04921-7_17.

E. K. Juuso. Recursive tuning of intelligent controllers of solar collector fields in changing operating conditions. In S. Bittani, A. Cenedese, and S. Zampieri, editors, *Proceedings of the 18th World Congress The International Federation of Automatic Control, Milano (Italy) August 28 - September 2, 2011*, pages 12282–12288. IFAC, 2011. doi:10.3182/20110828-6-IT-1002.03621.

E. K. Juuso. Intelligent methods in modelling and simulation of complex systems. *Simulation Notes Europe SNE*, 24(1):1–10, 2014. doi:10.11128/sne.24.on.102221.

E. K. Juuso. Expertise and uncertainty processing with nonlinear scaling and fuzzy systems for automation. *Open Engineering*, 10(1):712–720, 2020. doi:10.1515/eng-2020-0080.

E. K. Juuso. Intelligent temporal analysis of coronavirus statistical data. *Open Engineering*, 2021. doi:10.1515/eng-2020-0285.

E. K. Juuso and K. Leiviskä. Adaptive expert systems for metallurgical processes. In S.-L. Jämsä-Jounela and A. J. Niemi, editors, *Expert Systems in Mineral and Metal Processing, IFAC Workshop, Espoo, Finland, August 26-28, 1991, IFAC Workshop Series, 1992, Number 2*, pages 119–124, Oxford, UK, 1992. Pergamon.

S. Lahdelma and E. Juuso. Signal processing and feature extraction by using real order derivatives and generalised norms. Part 1: Methodology. *International Journal of Condition Monitoring*, 1(2):46–53, 2011. doi:10.1784/204764211798303805.

L. Ljung. *System Identification - Theory for the User*. Prentice Hall, Upper Saddle River, N.J., 2nd edition, 1999.

L. Ljung. Perspectives on system identification. In M. J. Chung and P. Misra, editors, *Plenary papers, milestone reports & selected survey papers, 17th IFAC World Congress, Seoul, Korea, July 6-11, 2008*, pages 47–59. IFAC, 2008. http://www.ifac-papersonline.net/.

Hannah Ritchie, Esteban Ortiz-Ospina, Diana Beltekian, Edouard Mathieu, Joe Hasell, Bobbie Macdonald, Charlie Giattino, Cameron Appel, Lucas Rodés-Guirao, and Max Roser. Coronavirus pandemic (COVID-19). *Our World in Data*, 2020. https://ourworldindata.org/coronavirus.