Tech Science Press

A Novel Heuristic Algorithm for the Modeling and Risk Assessment of the COVID-19 Pandemic Phenomenon

Panagiotis G. Asteris[1,*], Maria G. Douvika[1], Chrysoula A. Karamani[1], Athanasia D. Skentou[1], Katerina Chlichlia[2], Liborio Cavaleri[3], Tryfon Daras[4], Danial J. Armaghani[5] and Theoklis E. Zaoutis[6]

[1]Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Athens, 15122, Greece
[2]Laboratory of Molecular Immunology, Department of Molecular Biology and Genetics,
Democritus University of Thrace, Alexandroupolis, Greece
[3]Department of Civil, Environmental, Aerospace and Materials Engineering,
University of Palermo, Palermo, Italy
[4]School of Environmental Engineering, Technical University of Crete, Chania, Greece
[5]Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia
[6]Division of Infectious Diseases, The Children's Hospital of Philadelphia, Perelman School of Medicine,
University of Pennsylvania, PHL, USA
*Corresponding Author: Panagiotis G. Asteris. Email: asteris@aspete.gr; panagiotisasteris@gmail.com
Received: 31 July 2020; Accepted: 17 August 2020

Abstract: The modeling and risk assessment of a pandemic phenomenon such as COVID-19 is an important and complicated issue in epidemiology, and such an attempt is of great interest for public health decision-making. To this end, in the present study, based on a recent heuristic algorithm proposed by the authors, the time evolution of COVID-19 is investigated for six different countries/states, namely New York, California, USA, Iran, Sweden and UK. The number of COVID-19-related deaths is used to develop the proposed heuristic model as it is believed that the predicted number of daily deaths in each country/state includes information about the quality of the health system in each area, the age distribution of population, geographical and environmental factors as well as other conditions. Based on derived predicted epidemic curves, a new 3D-epidemic surface is proposed to assess the epidemic phenomenon at any time of its evolution. This research highlights the potential of the proposed model as a tool which can assist in the risk assessment of the COVID-19. Mapping its development through 3D-epidemic surface can assist in revealing its dynamic nature as well as differences and similarities among different districts.

Keywords: Algorithm; COVID-19; Gaussian-function; heuristic model; pandemic trend; prediction; SARS-CoV-2

1 Introduction

Taking into consideration the remarkable SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) viral spread causing the COVID-19 (coronavirus disease) pandemic [1,2] with more

than 5.5 million infected cases worldwide and 350.000 deaths as of May 26, 2020, mathematical models serve as helpful tools for authorities to make public health decisions, thereby ensuring optimal use of resources to reduce the morbidity and mortality associated with COVID-19 [3]. In order to control the pandemic viral spread, both mitigation and suppression of new infections have emerged as the predominant public health strategies [4]. Previous studies based on estimates of the fatality ratio showed a strong age gradient in risk of death [5]. The high risk of the COVID-19 pandemic phenomenon as well as the more available time of researchers due to restrictive received measures have had as a result, for the first time, that a plethora of researchers coming from the computational engineering area, focus their research methods on the pandemic's trend prediction [6–12]. The recently proposed forecasting models are of great interest in epidemiology promising more reliable and robust predictions not only for COVID-19 but for other "closely related families of viruses" of COVID-19. Thus, the modeling and risk assessment of a pandemic phenomenon is an important and complicated issue in epidemiology, and such an attempt is of great interest for public health decision-making.

In the present work, we have analyzed how to model the outbreak's spread taking as a parameter the number of deaths, and how the model changes over time as more information on the number of deaths becomes available. We have developed our model relying on mortality data from official sources that are in general more reliable than the reported confirmed cases based on diagnosis-testing. Based on the number of reported daily deaths from COVID-19, we developed a model of forecasting [13] that was applied to different parts of the world. In the present study, based on a recent heuristic algorithm proposed by the authors, the time evolution of COVID-19 is investigated for six different countries/states, namely the New York state, California, USA, Iran, Sweden and UK. The number of COVID-19-related deaths was used as it is believed that the predicted number of daily deaths in each state includes information about the quality of the health system in each area, the age distribution of population, geographical and environmental factors as well as other conditions. Based on derived predicted epidemic curves a 3D-epidemic surface is proposed to assess the epidemic phenomenon at any time of its evolution. This research highlights the potential of the proposed model as a tool which can assist in the risk assessment of the COVID-19.

In the light of the above, the manuscript is organized into 4 sections, including as its first section the introduction presented above. In Section 2, the proposed heuristic algorithm for the modeling of the COVID-19 trend is presented in detail, providing the basic assumptions and the mathematical details of Gaussian functions. Section 3 provides the various results and the discussion related to different countries/states under investigation. Here it should be noted that six different areas, namely New York, California, USA, Iran, Sweden and UK have been taken under consideration and the same heuristic algorithm has successfully been applied to all of them. Analysis of the findings with respect to key parameters is discussed. Finally, insightful conclusions and future recommendations are drawn in the final section.

2 Short Literature Review on COVID-19 Computational Models

Artificial intelligence and machine learning (ML) approaches e.g., artificial neural networks (ANNs), and genetic programming (GP) have been found to be feasible in predicting outbreak, trend or potential effect of COVID-19 in the near future [14–18]. Salgotra et al. [19] developed many gene expression programming (GEP) models to predict the potential effect of COVID-19 in 15 most effected countries i.e., USA, Turkey, Brazil, Iran, Germany, Canada, MEXICO, UK, Russia, Spain, Italy, France, China, South Africa, and Singapore. For prediction purposes, they

used two data categories of the mentioned countries including confirmed cases (CC) and death cases (DC) between January 2020 and May 2020. According to their results, it was expected that the maximum rises in CC and DC will be happed in Brazil and USA, respectively. In addition, they found a very high rate of increase (new cases) in USA, UK, Russia, Brazil and Mexico during June 2020. In another study, a time series analyses and forecasting of COVID-19 in India and its future behaviour, was performed by Salgotra et al. [20] using the GEP predictive technique. They considered and used data of CC and DC of the three major states in India i.e., Maharashtra, Gujarat and Delhi. Actually, for each state, they developed a mathematical GEP equation which is able to predict trends of CC and DC for future of the specific state. In addition, they did a same procedure for the entire India, which is the second most populous country in the world. They concluded that GEP-based predictive models/equations are highly reliable and can be treated as benchmark for time series predictions. In another similar project, Pinter et al. [21] conduced a research based on hybrid ML technique to estimate COVID-19 pandemic in Hungary. They proposed two models namely adaptive network-based fuzzy inference system (ANFIS) and multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) to estimate time series of infected individuals and mortality rate and after comparison, they selected the MLP-ICA predictive model because of its lower system error during prediction and validation stages. According to their conclusion, by late May, 2020, the outbreak and the total morality of COVID-19 would drop substantially. An advance hybrid intelligence system namely ISACL-MFNN, which integrates an improved interior search algorithm-based on chaotic learning strategy into a multi-layer feed-forward neural network, was developed in the study conducted by Rizk-Allah et al. [22] to predict the CS of COVID-19 in three countries i.e., Spain, Italy and USA. Then, to show capability of the proposed model, they compared its performance with the other techniques such as particle swarm optimization-MFNN and genetic algorithm-MFNN, and successfully indicated that the proposed ISACL-MFNN is able to provide higher performance compared to other techniques. In another investigation, Fanelli et al. [23] proposed a differential equation technique to evaluate the exponential growth of the COVID-19 in France, Italy and China based on data related to a period from 22/01/2020 to 15/03/2020. A comparative study of the ML techniques was carried out by Ardabili et al. [24] to predict COVID-19 outbreak in five countries i.e., USA, Italy, Iran, China, and Germany. These models include MLP and ANFIS predictive models. After constructing the predictive models and evaluating them using COVID-19 data of the mentioned countries, they found that both ML models are considered as an effective tool to model/predict the COVID-19 outbreak.

3 Methods

3.1 Assumptions and Data Sources

During the study of the development of the COVID-19 pandemic, the daily number of confirmed deaths due to COVID-19 for each location have been recorded and analyzed further. The selection of daily deaths was based on the authors' assumption that mortality rates provide more accurate and reliable data compared to the recordings of the number of daily infected individuals. The daily mortality rate is suggestive of additional information about the unique characteristics of each setting, which influence the pandemic transmission trend in each place. Such characteristics include:

- The climate and environmental conditions in each location
- The quality of the healthcare systems in each location
- The experience and expertise of the medical staff and healthcare workers

- The age distribution of the population
- The pandemic mitigation measures applied in each location

The main assumption during the design of this algorithm was the observation that the mortality rate, in particular the death numbers in the respective populations, follow a normal distribution. Although daily recording might not be the case for optimal normal distribution, it is important to note that the selection of death recordings every two or three days almost always leads to an optimal normal distribution. Following this assumption, the simulation of the pandemic spread was investigated for a variety of mortality rates in different settings, and the setting giving the most accurate results and predictions was selected in developing the model.

In the process of developing a new prediction model, it is common that scientists pay attention to the computational model; yet a reliable database is of high importance, and in order to achieve a reliable forecast, researchers should give the appropriate attention to the database used for the development, training, and validation of the model. In the light of the above, the final overall database was based on two individual databases. Data for the states/countries were obtained from the database Worldometer [25] and for the United States from the COVID Tracking project [26]. In our prediction we did not accommodate underreporting of cases or deaths, which is common in many parts of the world with considerable influence on the prediction results. Recent analysis shows that the official global COVID-19 death toll is much higher (60%) than officially reported [27].
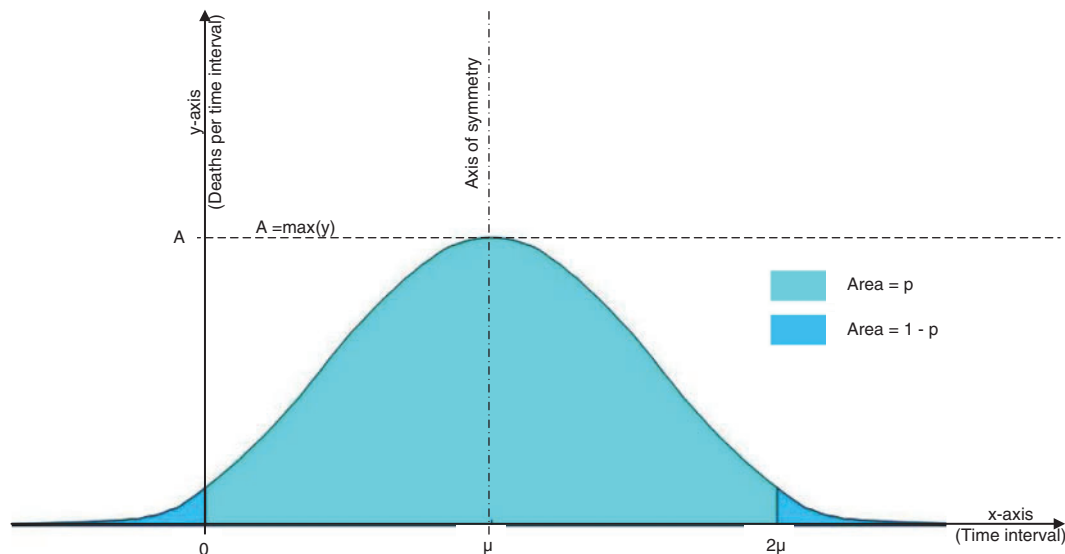

Figure 1: Shape, parameters and terminology of Gaussian function

3.2 Proposed Heuristic Algorithm

By analyzing the official data from China, including daily COVID-19 infections and deaths, it is clear that they can be expressed with meaningful accuracy using a suitable Gaussian curve (or, equivalently, a proper normal distribution density function). In addition, by studying the evolution of the pandemic and the course of the restrictions in this country, and taking into account that many European and other world countries have taken similarly strict restrictive mitigation

measures, we assumed that the development of COVID-19 pandemic would have similarities to its development in China. In other words, we propose that the number of new incidents or deaths will be expressed using a proper normal distribution (Fig. 1).

A Gaussian function is a function of the form:

$$y = f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

whose graph is a symmetrical bell-shaped curve centered at the position $x = \mu$. A is the height of the peak and the variance σ^2 controls its width. On both sides of the peak, the tails of the curve quickly fall of and approach the x-axis (asymptote). Our algorithm aims to determine in each state or setting the optimal normal curve for daily deaths by calculating the parameters A, μ, σ^2; i.e., by fitting the "best" possible normal curve. The optimality of the normal curve is given with reference to well-known statistical indices.

More precisely, the main steps of the algorithm are (through a triple loop):

(1) (for A/first inner loop) We start from a given value of A, and with **Step 1**, we continue up to a certain value (desired accuracy) depending on the maximum value of our available data (deaths),

(2) (for μ/second inner loop) We start from a value of $\mu = 10$ and we continue, with **Step 1** (day), up to a value of $\mu = 60$ (we observed for example that in the case of China, the phenomenon lasted for about 60 days with an average [peak day of deaths] at about the 30th day).

(3) (for p/third inner loop) $A = \frac{1}{\sigma\sqrt{2\pi}}$ so peaking σ^2 can be calculated. The algorithm then uses a probability value p, starting from $p = 0.85$, and with step 0.00001, continues up to 0.99999 (it is known that $P = p \times 100\%$ of the data under a normal distribution curve lie inside the interval $[\mu + z_{q/2}\sigma, \mu - z_{q/2}\sigma]$, where $q = 1 - p$ and $z_{q/2} = \Phi^{-1}(q/2)$, Φ the cumulative distribution function of the standard normal distribution $N(0, 1)$ and Φ^{-1} its inverse function. This interval is used to fit the actual data, using a proper transformation.

The algorithm application creates a large number of proper normal distributions by calculating the three parameters (theoretical/experimental values) each time. Finally, these values are compared with the empirical values (actual numbers of deaths) and the "best" possible curve is being selected using a number of indices, i.e., the algorithm searches for the optimal curve characteristics using the available data up to the forecast day.

In the following pseudo-code (Algorithm 1), the algorithmic implementation of the method is demonstrated.

3.3 Performance Assessment

The reliability and accuracy of the best fit Gaussian curves developed for each one prediction were evaluated using Pearson's coefficient of determination R^2, the root mean square error (RMSE) and the mean absolute percentage error (MAPE). RMSE presents information on short-term efficiency, which is a benchmark of the difference in predicated values compared to the experimental values. A lower RMSE indicates a more accurate evaluation. The Pearson's coefficient of determination R^2 measures the variance that is interpreted by the model. R^2 values ranges from 0 to 1, with the model having the healthiest predictive ability when it is near to 1 producing little analysis when it is near to 0.

Algorithm 1: Proposed heuristic algorithm' pseudo code for the finding of best Gaussian function that fits data (daily deaths due to COVID-19)

Input: Daily deaths
Output: Best fitting Gaussian curve of daily deaths
Definition of min and max values of Gaussian function parameters (A, μ, p) based on data
Index i=0
for $A \leftarrow$ min A:1:max A **do**
    for $\leftarrow$ min μ:1: max μ **do**
        for $\leftarrow$ min p:0.00001:$\leftarrow$ max p **do**
            i = i+1
            define i-th Gaussian function based on Eq. (1)
            define root mean square error (RMSE) of i-th Gaussian function
        end
    end
end
Return: Best fitting curve of daily deaths is the Gaussian function with the minimum RMSE

The aforementioned statistical parameters have been calculated by the following expressions [28–35].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i}\right| \tag{3}$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\right) \tag{4}$$

where x_i are the actual/experimental values and y_i the predicted/theoretical values.

3.4 Methodology

The present section outlines the methodology used to investigate the spread of COVID-19 in a country, state, city, or region. As an example, the methodology is presented here as it was conducted and applied in the investigation of the spread of the epidemic in China. Given that the epidemic in China preceded the spread of the epidemic to other countries, this allows us to apply the proposed algorithm at the beginning of the phenomenon; in the intermediate phase, which is usually characterized by a strong disease dynamic; and at its peak, where the phenomenon begins to fade or recede.

The main principles of the proposed methodology are as follows:

- In each step of the study of the phenomenon, the optimal normal distribution is calculated using the proposed algorithm and based on data available at the moment of calculation.
- The first assessment must be made 14 days after the first death recorded. The period of two weeks is considered necessary to reliably characterize the beginning of the phenomenon.

- At each time step, following the 14-day period from the first death record, the optimal data simulation curve is calculated with the use of the proposed algorithm. Fig. 2 shows the optimal curves for the country of China, that best simulate the data (number of deaths) for three different days (6, 12, and 18 February 2020).
- In addition to the above estimates it is possible to reliably predict the expected number of deaths for the time period up to 10 days following the time of the prediction (Fig. 3). The algorithm provides simultaneous estimates for its higher and lower limits. Based on a comprehensive study in the ten aforementioned countries, states, regions, and cities, and the results presented below, these limits, as well as the difference between the predicted and actual deaths, were confirmed for all states.
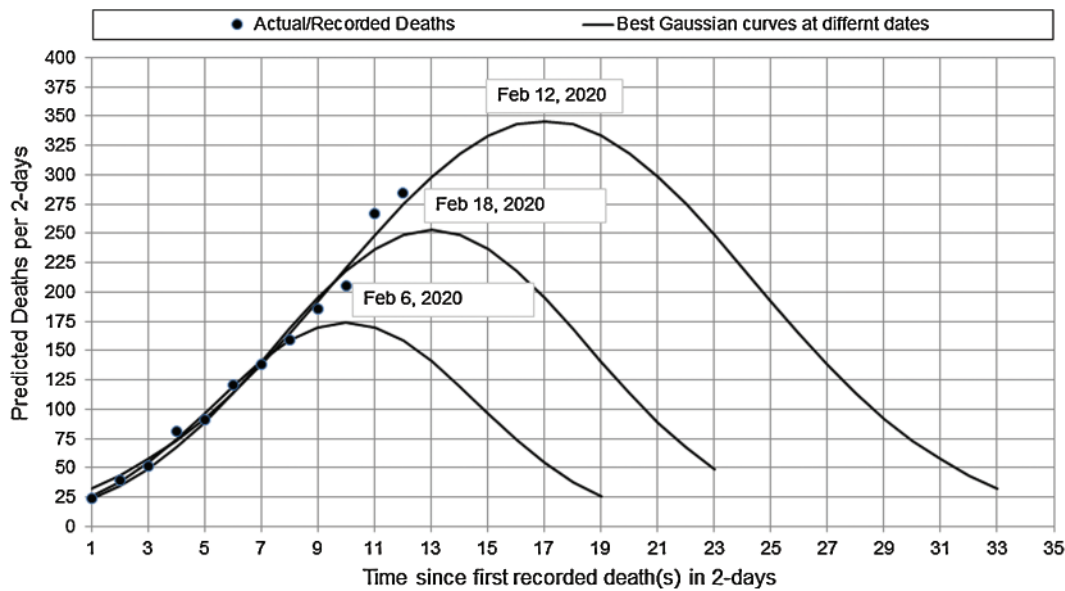

Figure 2: Prediction of the best fit curves of the actual deaths at three different dates of prediction (Country of China)

4 Results and Discussion

In order to implement the proposed algorithm, a computer program has been developed at the Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Athens, Greece. Utilizing this software through implementation of the heuristic algorithm, the development of the epidemic was investigated in six different geographical locations: the states of California and New York, and the countries United States, Iran, Sweden, and the United Kingdom. The software was used to predict deaths for each country from the first day deaths were recorded until available data as of May 4, 2020.

Fig. 4 presents the predicted deaths for the next eight days as well as the corresponding epidemic curves over time for USA. These curves show the evolution of the phenomenon as time passes. The predicted deaths (red dots) always follow the predicted curve (red line) with a deviation lower than ±30% (green area).
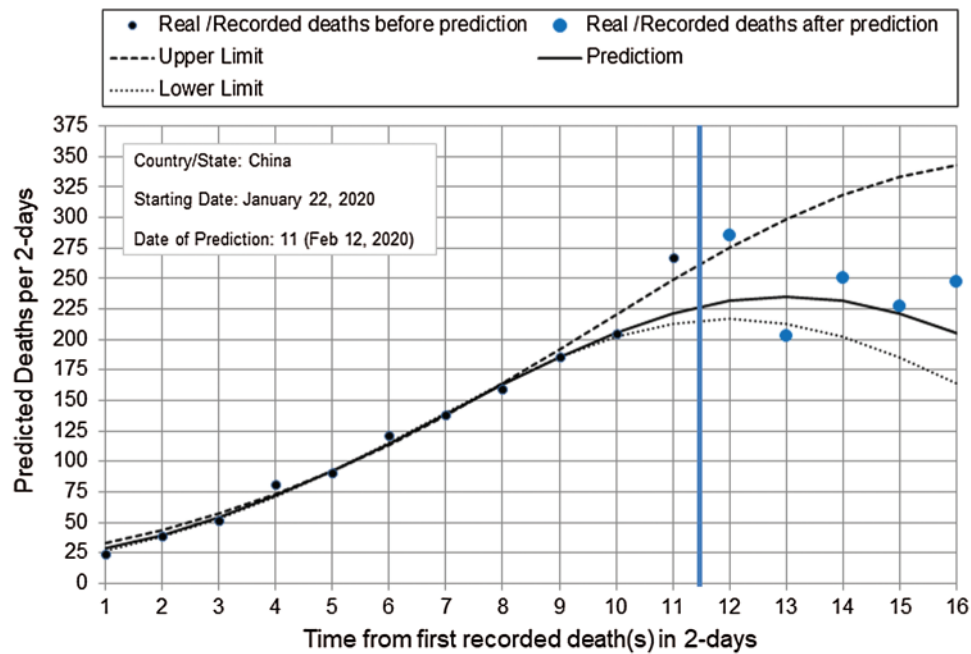
Figure 3: Prediction of number of deaths, in 2-day intervals, for the next ten days starting February 12, 2020, for the country of China. Black dots represent actual data until the day in which the algorithm made the prediction. Blue dots represent actual data after the day in which the algorithm made the prediction

Based on the predicted epidemic curves (Fig. 4) a 3D epidemic surface for the pandemic trend evolution over time is proposed. Fig. 5 represents the 3D epidemic surface for USA, that is the predicted number of deaths over time over a period for more than 2 months. In accordance with the epidemic curves presented above, this 3D epidemic surface strongly reveals the dynamic nature of the pandemic phenomenon with its three distinct phases. Furthermore, using the proposed heuristic algorithm, useful parameters about the pandemic phenomenon such as peak time, deaths at peak time and total deaths have been predicted (Tab. 1).

As mentioned, using our proposed algorithm, we examined in 6 different states/countries how our predictions of the model's fitting parameters change in real time as time progresses. The following epidemic mortality curves (Fig. 6) show the predicted total number of deaths over a period of 2 months, comparing in each diagram 2 states/countries, and in the last one all 6 locations together. Every epidemic curve reveals a strongly dynamic behavior of pandemic phenomenon characterized by three distinct phases (i) the first phase with strongly dynamic behavior (first three weeks of the phenomenon), (ii) the second phase where the phenomenon is characterized by an oscillation and (iii) the third phase where a balance takes place.

Key variables vary substantially among countries as well as among states in a large country such as US. The predictive model for a large country such as US aggregates heterogeneous sub-epidemics in local areas. Distinct differences are evident among California and New York, probably due to demographic factors and/or different climate and environmental conditions at the time period the phenomenon takes place. Namely, the average highest temperature in March was 20°C for California and 10°C for New York. New York is the most densely populated state in USA.
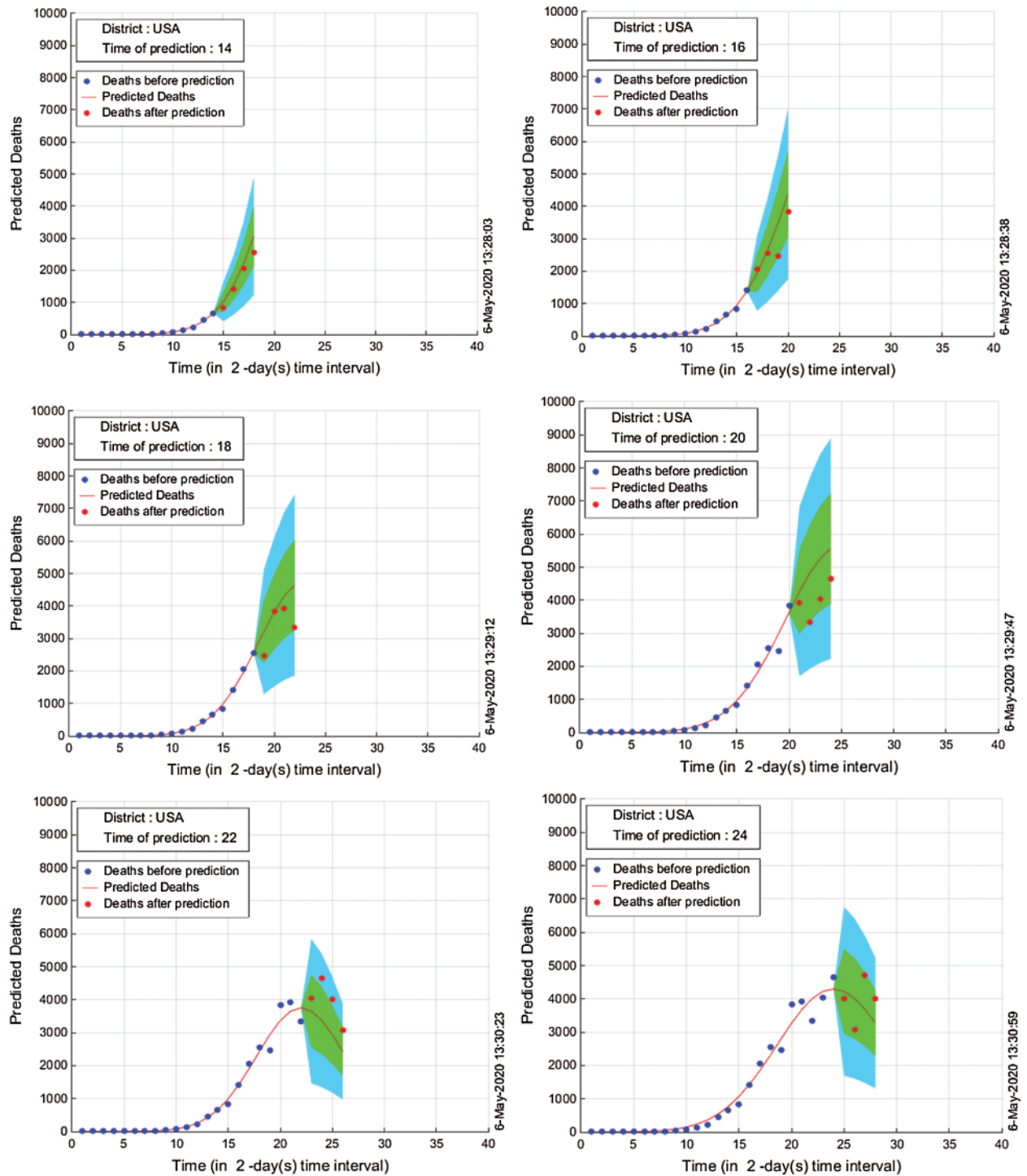
Figure 4: Verification of number of death predictions, in 2-day time intervals, for the next eight days, for USA. Blue dots represent the actual recorded deaths until the day of prediction while red dots represent actual data after the day on which the algorithm made the prediction. Green area represents deviation between predicted and actual deaths lower than ±30%, while the light blue area deviation smaller than ±60%
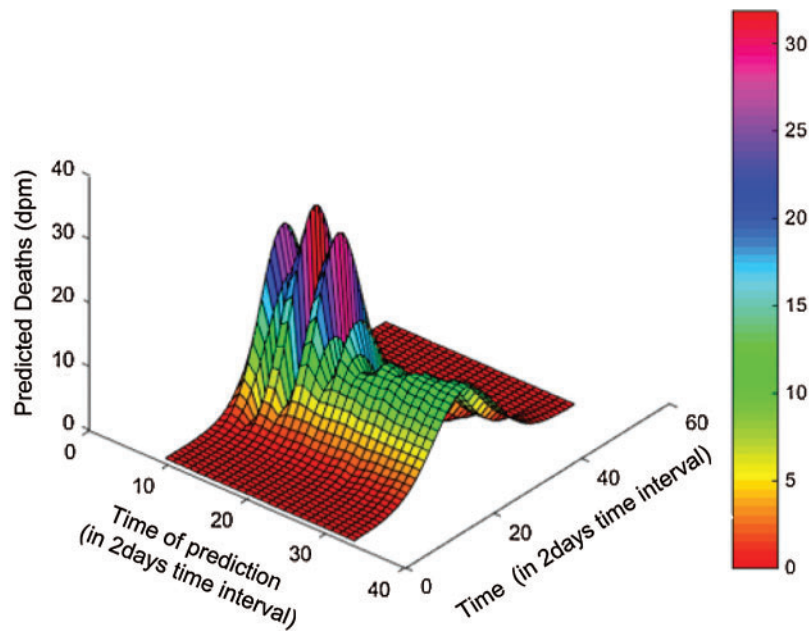
Figure 5: Epidemic surface for USA based on available data as of 4 May 2020 (dpm: deaths per million)

Table 1: Predicted parameters based on 4 May 2020 available data

Nr	Country/State	Population	Population density	Peak time	Deaths at peak time
		(million)	Number of people per square mile	(days)	(dpm)
1	California	39.51	251.3	60	2.08
2	Iran	81.8	30	46	1.64
3	USA	328.2	136	54	6.90
4	Sweden	10.23	57.5	44	8.99
5	UK	66.65	94	44	13.65
6	New York	8.399	421	30	110.13

Note: dpm: deaths per million

It is remarkable that Sweden and UK share similar dynamic mortality curves (Fig. 6), which could be attributed to the similarity of public health approaches; both Sweden and UK out of few countries have opted against a "lockdown" to contain the spread of coronavirus. California and Iran also show similarities in the course of the mortality curve (Fig. 6); both countries share the same latitude (about 35°N).

The proposed algorithm based on the data (deaths per time interval) determines the epidemic curve and surface, as well as the integrated parameters, such as 'peak time' (time period from the day the first death was traced until the day where maximum number of daily deaths were recorded) and 'number of deaths at peak time' (Tab. 1). Both are valuable parameters that could

help us classify the epidemic in different locations towards an understanding of the COVID-19 pandemic phenomenon. Moreover, these parameters could be useful for estimating and monitoring the severity of the phenomenon for health authorities to take action. Interestingly, New York State reached peak time in 30 days, while California in 60 days.
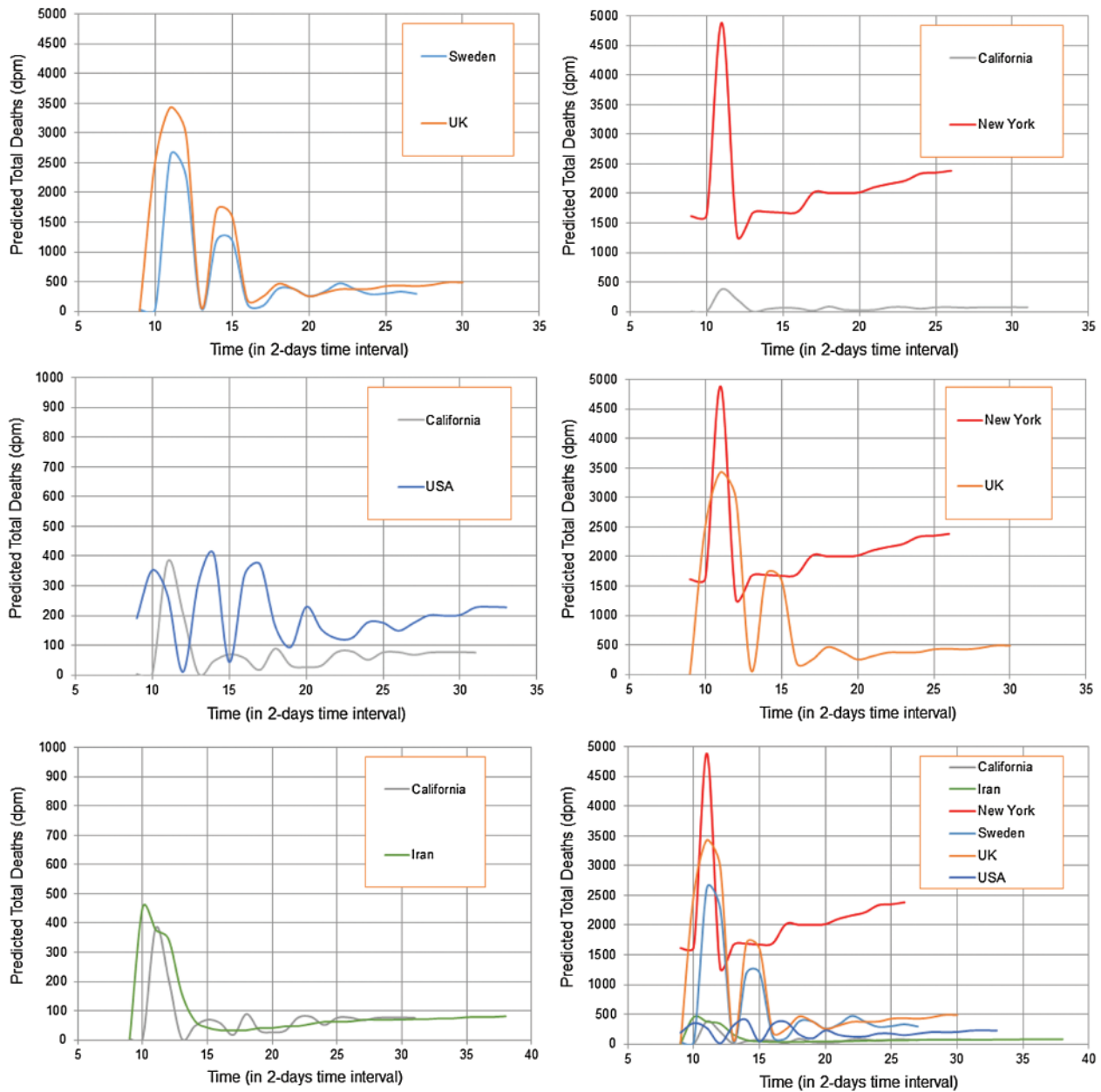

Figure 6: Dynamic behavior of the pandemic phenomenon (epidemic curves with temporal changes in number of deaths per million people) based on available data as of 4 May 2020 (dpm: deaths per million)

Careful attention should be made by worldwide organizations (WHO, CDC) to ensure reliable recording for mortality data. Accurate reporting is of particular importance to ensure the reliability and validity of the predictions.

The findings presented herein are preliminary results proposing two valuable parameters (peak time and deaths at peak time) for the COVID-19 epidemic. The authors have started, based on the proposed algorithm, in-depth investigations on the epidemic differences and similarities of the pandemic phenomenon among ten states in the USA as well as among different cities in Italy in order to reveal in more detail the nature of the COVID-19 dynamic phenomenon.

5 Conclusions

A model of forecasting for short-term prediction of COVID-19 mortality was applied to different countries/states in the world based on the number of reported deaths from COVID-19. The same model examined and estimated COVID-19 related mortality with great accuracy, even after the peak was reached. The proposed 'peak time' and 'deaths at peak time' could classify the epidemic among different countries/states as well as further explain similarities and differences among different locations, helping us to understand the COVID-19 pandemic phenomenon. Interestingly, our multidisciplinary approach, although not based on classical epidemic infection curves, leads to results that show epidemiologic relevance and can help epidemiologists classify and predict the course of the pandemic phenomenon, and at the same time provide a useful tool for public health authorities in decision-making and operational planning.

The proposed algorithm is also expected to make a substantial contribution to engineering problems, where it is frequent that the parameters of a multitude of engineering problems follow a normal distribution. Authors also believe that since data/parameters of other closely related families of viruses causing mortality are expected to have a normal distribution, the proposed algorithm will be applicable to those cases as well.

References

1. Fauci, A. S., Lane, H. C., Redfield, R. R. (2020). COVID-19—navigating the uncharted. *New England Journal of Medicine, 382(13),* 1268–1269. DOI 10.1056/NEJMe2002387.
2. Lv, M., Luo, X., Estill, J., Liu, Y., Ren, M. et al. (2020). Coronavirus disease (COVID-19): a scoping review. *Eurosurveillance, 25(15),* 2000125. DOI 10.2807/1560-7917.ES.2020.25.15.2000125.
3. Jewell, N. P., Lewnard, J. A., Jewell, B. L. (2020). Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *JAMA, 323(19),* 1893–1894. DOI 10.1001/jama.2020.6585.
4. Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A. et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature, 584(7820),* 257–261. DOI 10.1038/s41586-020-2405-7.
5. Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C. et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infectious Diseases, 20(6),* 669–677. DOI 10.1016/S1473-3099(20)30243-7.

6. Salgotra, R., Gandomi, M., Gandomi, A. H. (2020). Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming. *Chaos, Solitons & Fractals, 138,* 109945. DOI 10.1016/j.chaos.2020.109945.

7. Papastefanopoulos, V., Linardatos, P., Kotsiantis, S. (2020). COVID-19: a comparison of time series methods to forecast percentage of active cases per population. *Applied Sciences, 10(11),* 3880. DOI 10.3390/app10113880.

8. Younes, A. B., Hasan, Z. (2020). COVID-19: modeling, prediction, and control. *Applied Sciences, 10(11),* 3666. DOI 10.3390/app10113666.

9. Soufi, G. J., Hekmatnia, A., Nasrollahzadeh, M., Shafiei, N., Sajjadi, M. et al. (2020). SARS-CoV-2 (COVID-19): new discoveries and current challenges. *Applied Sciences, 10(10),* 3641. DOI 10.3390/app10103641.

10. Almeshal, A. M., Almazrouee, A. I., Alenizi, M. R., Alhajeri, S. N. (2020). Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models. *Applied Sciences, 10(10),* 3402. DOI 10.3390/app10103402.

11. Torrealba-Rodriguez, O., Conde-Gutiérrez, R. A., Hernández-Javier, A. L. (2020). Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons & Fractals, 138,* 109946. DOI 10.1016/j.chaos.2020.109946.

12. Barmparis, G. D., Tsironis, G. P. (2020). Estimating the infection horizon of COVID-19 in eight countries with a data-driven approach. *Chaos, Solitons & Fractals, 135,* 109842. DOI 10.1016/j.chaos.2020.109842.

13. Asteris, P. G., Douvika, M., Karamani, C., Skentou, A., Daras, T. et al. (2020): A novel heuristic global algorithm to predict the COVID-19 pandemic trend. *MedRxiv,* DOI 10.1101/2020.04.16.20068445.

14. Velásquez, R. M. A., Lara, J. V. M. (2020). Forecast and evaluation of COVID-19 spreading in USA with reduced-space Gaussian process regression. *Chaos, Solitons & Fractals, 126*, 109924.

15. Ghanbari, B. (2020). On forecasting the spread of the COVID-19 in Iran: the second wave. *Chaos, Solitons & Fractals, 140*, 110176.

16. Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., Millham, R. C. (2020). Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. *International Journal of Environmental Research and Public Health, 17(15),* 5330. DOI 10.3390/ijerph17155330.

17. Yang, Z., Zeng, Z., Wang, K., Wong, S. S., Liang, W. et al. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease, 12(3),* 165–174. DOI 10.21037/jtd.2020.02.64.

18. Allam, Z., Jones, D. S. (2020). On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management. In *Healthcare, 8(1),* 46. Multidisciplinary Digital Publishing Institute.

19. Salgotra, R., Gandomi, M., Gandomi, A. H. (2020a). Evolutionary modelling of the COVID-19 pandemic in fifteen most affected countries. *Chaos, Solitons & Fractals, 140*, 110118.

20. Salgotra, R., Gandomi, M., Gandomi, A. H. (2020b). Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming. *Chaos, Solitons & Fractals, 138*, 109945.

21. Pinter, G., Felde, I., Mosavi, A., Ghamisi, P., Gloaguen, R. (2020). COVID-19 pandemic prediction for hungary; a hybrid machine learning approach. *Mathematics, 8(6),* 890. DOI 10.3390/math8060890.

22. Rizk-Allah, R. M., Hassanien, A. E. (2020). COVID-19 forecasting based on an improved interior search algorithm and multi-layer feed forward neural network. arXiv preprint arXiv: 2004.05960.

23. Fanelli, D., Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals, 134,* 109761. DOI 10.1016/j.chaos.2020.109761.

24. Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R. et al. (2020). COVID-19 outbreak prediction with machine learning. *MedRxiv,* DOI 10.1101/2020.04.17.20070094.

25. https://www.worldometers.info/about/.

26. https://covidtracking.com/about-data.

27. Burn-Murdoch, J., Romei, V., Giles, C. (2020). Global coronavirus death toll could be 60% higher than reported. *Financial Times*.

28. Alavi, A. H., Gandomi, A. H. (2012). Energy-based numerical models for assessment of soil liquefaction. *Geoscience Frontiers, 3(4),* 541–555. DOI 10.1016/j.gsf.2011.12.008.

29. Armaghani, D. J., Hatzigeorgiou, G. D., Karamani, C., Skentou, A., Zoumpoulaki, I. et al. (2019). Soft computing-based techniques for concrete beams shear strength. *Procedia Structural Integrity, 17,* 924–933. DOI 10.1016/j.prostr.2019.08.123.

30. Asteris, P. G., Apostolopoulou, M., Skentou, A. D., Moropoulou, A. (2019). Application of artificial neural networks for the prediction of the compressive strength of cement-based mortars. *Computers and Concrete, 24(4),* 329–345.

31. Asteris, P. G., Armaghani, D. J., Hatzigeorgiou, G. D., Karayannis, C. G., Pilakoutas, K. (2019). Predicting the shear strength of reinforced concrete beams using artificial neural networks. *Computers and Concrete, 24(5),* 469–488. DOI 10.12989/cac.2019.24.5.469.

32. Asteris, P. G., Mokos, V. G. (2019). Concrete compressive strength using artificial neural networks. *Neural Computing and Applications, 32(15),* 11807–11826. DOI 10.1007/s00521-019-04663-2.

33. Apostolopoulou, M., Asteris, P. G., Armaghani, D. J., Douvika, M. G., Lourenço, P. B. et al. (2020). Mapping and holistic design of natural hydraulic lime mortars. *Cement and Concrete Research, 136,* 106167. DOI 10.1016/j.cemconres.2020.106167.

34. Armaghani, D. J., Asteris, P. G. (2020). A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. *Neural Computing and Applications, 93(1),* 45. DOI 10.1007/s00521-020-05244-4.

35. Ly, H. B., Pham, B. T., Le, L. M., Le, T. T., Le, V. M. et al. (2020). Estimation of axial load-carrying capacity of concrete-filled steel tubes using surrogate models. *Neural Computing and Applications, 20(20–20),* 861. DOI 10.1007/s00521-020-05214-w.