Machine Learning Approaches for Temporal and Spatio-Temporal Covid-19 Forecasting: A Brief Review and a Contribution

Ana Clara Gomes da Silva, Clarisse Lins de Lima, Cecilia Cordeiro da Silva,
Giselle Machado Magalhães Moreno, Eduardo Luiz Silva,
Gabriel Souza Marques, Lucas Job Brito de Araújo, Luiz Antônio
Albuquerque Júnior, Samuel Barbosa Jatobá de Souza,
Maíra Araújo de Santana, Juliana Carneiro Gomes,
Valter Augusto de Freitas Barbosa, Anwar Musah, Patty Kostkova,
Abel Guilhermino da Silva Filho, and Wellington P. dos Santos

1 Introduction

In the last month of 2019, in the city of Wuhan, China, a local outbreak of an illness was found that was cursing with symptoms such as cough, fever, sore throat, shortness of breath, fatigue, pneumonia evolving into severe acute respiratory syndrome, possibly being fatal. It was soon discovered that the disease was caused by a new coronavirus, SARS-CoV2 (severe acute respiratory syndrome coronavirus), and that the contagiousness and course of the disease would make

A. C. G. da Silva · W. P. dos Santos (✉)
Department of Biomedical Engineering, Federal University of Pernambuco, Recife, Pernambuco, Brazil
e-mail: wellington.santos@ufpe.br

V. A. de Freitas Barbosa
Academic Unit of Serra Talhada, Rural Federal University of Pernambuco, Serra Talhada, Brazil

C. L. de Lima · M. A. de Santana · J. C. Gomes
Graduate Program in Computer Engineering, Polytechnique School of the University of Pernambuco, Recife, Brazil

C. C. da Silva · E. L. Silva · G. S. Marques · L. J. B. de Araújo · L. A. A. Júnior
S. B. J. de Souza · A. G. da Silva Filho
Center for Informatics, Federal University of Pernambuco, CIn-UFPE, Recife, Brazil

G. M. M. Moreno · P. Kostkova
Department of Atmospheric Sciences, IAG, University of São Paulo, USP, São Paulo, Brazil

A. Musah
Institute for Risk and Disaster Reduction, University College London, UCL-IRDR, London, UK

it a threat to the world's different health systems. The disease, named COVID-19, quickly spread across four continents, until on March 11, 2020 it was declared a pandemic by the World Health Organization [20]. In little more than a year after the pandemic was recognized, covid-19 has now reached more than 130,000,000 people, adding up to more than 2,800,000 fatalities [98]. These numbers continue to grow. The biggest public health crisis in decades had begun. If on one hand COVID-19 exposed the epidemiological fragility of an extremely connected world, using tourist and commercial routes to reach the most diverse populations; on the other hand, the global connection proved its value once again, with scientists from various fields of knowledge and from the most varied countries establishing collaborations in the urgent effort to know, detail, prevent, detect, and contain the virus. Thus, a large amount of research has begun on epidemiological and pathophysiological aspects, drug development, virus detection tests, vaccines, and case prediction and control. Science has advanced by leaps and bounds and made the distance between the emergence of a new disease, the identification of the causative agent, the sequencing of its genetic material, and the appearance of the first viable vaccines seem shorter in just 1 year. Despite these achievements, one year after the WHO recognized the pandemic, the disease continues to spread, presenting an exuberance of possible clinical manifestations. The virus has new and even more transmissible variants (REF). The most viable form of control since the beginning of the pandemic continues to be: case identification, tracking and contact isolation. The ability to identify the presence of the pathogen plays an important role both in preventing the spread of the disease and in adequately combating it. Delay in diagnosis can delay proper patient care, hindering recovery, and especially allowing undiagnosed infected people to circulate in society, spreading the virus. The most well accepted test for diagnosing COVID-19 is RT-PCR (Reverse Transcription Polymerase Chain Reaction); however, the procedures for this test take several hours [24] and the result can take days to be available. In addition, there is the possibility of virus presence and transmission even if the RT-PCR test is negative, depending on the time of contamination at which the test was performed. Understanding more about the behavior of the virus in populations (identifying risk groups, more vulnerable social groups) or about its spatial and temporal spread in a region was, since the beginning of the pandemic, a factor that reduced the impact of the virus. And this has allowed greater assertiveness in the measures of isolation, protection, and vaccination, besides being determinant for economic, social, and administrative decisions of governments that have the intention to contain the pandemic by COVID-19. In this context, a relatively new area of Public Health, digital epidemiology, has gained space and recognition, providing effective monitoring of confirmed cases, accumulations, and excess deaths. Moreover, the possibility of using machine learning to make temporal and spatial predictions about the occurrence of COVID-19 has definitely brought artificial intelligence into the healthcare field. This chapter is dedicated to exploring some of the major studies that have been done on the use of forecasting by compartmental, statistical, machine learning, and hybrid approaches.

This chapter is organized as follows: in Sect. 2, we present the theoretical basis and a review of compartment forecasting models; in Sect. 3, we detail the forecasting

approaches based on statistical learning and present the basis of the main machine learning methods applied to Covid-19 forecast, as well as state-of-the-art works selected taking into account academic relevance, i.e. the number of citations and the impact factor of journals and books; finally, in Sect. 4 we present our final considerations and general conclusions.

2 Forecasting by Statistical Learning and Compartment Models

With the outbreak of the 2019 coronavirus disease many researchers have become interested in mathematically modeling this new disease. Many have done these studies using compartmental models based on differential equations. These models can be described by two types of equations, ordinary differential equations (ODEs) and partial differential equations (PDEs). The techniques for solving each of these models and methods for doing numerical simulations are different.

The following are some studies that have used mathematical modeling to understand how disease dynamics work and even make predictions using computational techniques associated with these models.

Among the ODE-based compartmental models the researchers Sarkar et al. [85] developed a 6-compartment model that extends the classical SEIR to predict Covid-19 dynamics, where a sensitivity analysis was conducted to recognize the most influential parameters with respect to the infected population. For this purpose, the partial rank correlation coefficient (PRCC) technique was used for all input parameters with respect to variable I(infected or symptomatic individuals). And the numerical implementation was done in the FORTRAN program with the method of least squares (MMQ) to adjust the diary cases of the disease.

The researchers Suba et al. [89] developed a model based on ODEs and also used implementation by means of the method of least squares. In this work, seven models were developed, and to find the parameters of the model, excel spreadsheet and MMQ and plotted graphs in MATLAB were used. This study did the sensitivity analysis using real data from Tamil Nadu. Good results with simple methods, but the system is sensitive to the change of the basic reproduction number R_0, which changes the whole system automatically.

Some other studies follow the line of numerical implementation with MATLAB. This same software was used in the study developed by Zhong et al. [105], to perform the numerical calculation of the created differential equation system. And real data was used to predict the number of infected. This study brought predictions of the epidemic in different scenarios and with different levels of anti-epidemic measure and medical care represented by beta rate and gamma rate, with unreliable data through objective analysis. But this study has a prediction limited by the data and their reliability, because data before January 18, 2020 should be used with caution. Mandal et al. [59] also used MATLAB software to solve the system of

differential equations that describes the proposed SEIQR model. The method used was the fourth-order Runge–Kutta (RK4). In this study, a theoretical analysis and numerical simulation are performed, as well as a stability analysis and estimation of R_0. The prediction made is sensitive to some parametric conditions, and since human behavior is uncertain there are changes in the parametric space corroborating to the change in the graphs of the COVID-19 cases. Therefore, the prediction made is short term. MATLAB was also used by Jiang [37]. Initially this work used the simulation repository built into the Netlogo software to create a SIR model to simulate virus transmission. The simulation took place in a closed environment (Small World) and assumed that there were no vital dynamics, i.e., no one died or was born naturally. To optimize the parameters of the proposed model, the MATLAB function fmincon was used. To find the numerical solution of the ODE system and adjust the curves, MATLAB's ode45 function was used. By using this function the values obtained were quite consistent with the real data as well as the simulation curves. This model was done for USA and for Hubei, China. For USA a model without vital dynamics was used, due to lack of data. The parameters definitely change with time in the real situation. The data from asymptomatic individuals is late, which makes it difficult to establish a SEIR-based model for fitting and prediction. As for Hubei, the prediction does not match the real situation. And finally, none of the models divides infected people into isolated and non-isolated infected individuals, or whether they received effective treatment. Massonis et al. [60] did a multi-state review using SIR and SEIR models described by systems of ODEs in which it evaluates structural identifiability, i.e., ability to provide insights into their unknown parameters, and observability (unmeasured states). A total of 255 articles were evaluated, 98 with SIR models and 157 with SEIR models. And a list of 36 model structures was made. The ability to provide reliable information was evaluated, and theoretical concepts of structural identifiability and reliability control were used for this. STRIKE-GOLDD, an open source toolbox and GenSSI2 MATLAB were used as analysis tools, and for some models the Observability Test code in Maple, Identifiability Analysis in Mathematica, SIAN in Maple, and others were used. Most models found in the literature have identifiable parameters. Often allowing for variability in an unknown parameter improves the observability and/or the identifiability of the model. This work has contributed to providing a detailed analysis of the structural identifiability and observability of a large set of compartmentalized COVID-19 models presented in the recent literature. To model and make prediction of COVID-19 evolution in Brazil, Bastos and Cajueiro [11] proposed two models SIRD and SIRASD described by ODEs. And to find the numerical solution of the ODE system and fit the curves, ode45 function, also from MATLAB, was used. And although this method controls the error by assuming fourth-order precision, it uses a precise fifth order formula to perform the steps. As a starting condition we used data from the Brazilian Institute of Geography and Statistics (IBGE). And the data used were from the Brazilian Ministry of Health (February 25 to March 30, 2020). For the estimation procedure, we minimized the loss functions using the method "$optimize.least_squares$" also from the scipy Python 34 library using the Cauchy loss with scaling parameter. It

is notable that although the SIRASD model predicts that the number of infected is higher than the SIRD model estimates, it also manages to predict a lower peak for those infected with symptoms, which are those who require medical attention. This model is advantageous for short-term prediction for Brazil. The methodology of this study was able to estimate the asymptomatic individuals, who may not be entirely present in the data. But because the study was done at the beginning of the pandemic, there was little data and there were cases of underreporting of the actual number of infected people. In addition, this study did not take demographic effects into account, and it was assumed that there was no reinfection. The SIRASD model proved sensitive to the initial condition of asymptomatic individuals. Because the number of tests is small to map the entire population, it is necessary to work with assumptions.

A modeling of the spread of coronavirus taking into account the cases of undetected infections in China was done by Ivorra et al. [34]. In this paper, a deterministic SEIHRD model was made, which has low computational complexity and possibility of using ODE theory to analyze and interpret properly. This model is solved numerically via fourth-order Runge–Kutta (RK4) with 4 h time interval to approximate the solution of the system. Both Runge–Kutta and the WASF-GA algorithm have been implemented in Java. It is advantageous to use a deterministic model when you have little data, but the methodology used aimed precisely at solving this limitation. And a robust approach for overfitting the model parameters with respect to the reported data was created. However, the results are unsatisfactory because the estimation was done at the early stage of the epidemic.

Ambikapathy and Krishnamurthy [6] developed and validated a mathematical model to assess the impact of various scenarios on COVID-19 transmission in India. A compartmentalized ODE model incorporating the actual cases from 14 countries, China, Italy, Germany, France, USA, UK, Sweden, Netherlands, Austria, Canada, Australia, Malaysia, Singapore, and India, was proposed. The model was applied to predict transmission in India and the highest exposure situations, such as transit stations and shopping malls, were evaluated. It was validated using the infections reported in the adopted period and was used to predict future infected cases in the above countries, considering a 65-day period (IndiaSim implementation). Different intervention strategies were used with blocking periods of 4, 14, 21, 42, and 60 days. The model developed can capture the infection dynamics in each country to a considerable extent and predict future cases. The use of an ODE system to describe the models is advantageous because it is possible to apply controls to the model and find results. Nevertheless, the model suffers from numerical errors because at the beginning of the disease the S-compartment has a high value and the I and R-compartments have very low values. In addition, the model proposed in India assumes no spread of the disease in the community until the first week of March 2020, and the dynamic prediction interval is limited (110 days). The model will need to be updated. Also in order to do a predictive analysis of COVID-19 in China, Italy, and France, developed a SIRD model to predict the position of the epidemic peak of the disease. For stochastic evolution, the Python-Scipy package was used. For Italy, the prediction with nonlinear fit strategy for the endemic peak is robust.

With simulations it was shown that the recovery rate is the same for China as for Italy, but the infection and mortality rates seem to be different. This model showed that cultural factors influence the infection rate, varying from one country to another. The model has the limitation of data sensitivity, so it changes from one country to another. And when making numerical solution adjustments, it was found that the data reported for the outbreak in France is still too preliminary to justify a significant adjustment of this kind. The researchers [48] introduced a SIRD model described by a system of ODEs to analyze the behavior of COVID-19 disease in the USA, Germany, UK, and Russia and solved using numerical methods, and the data agree well with the model. The model predicts the peak of the epidemic in each country and compares the results obtained. Germany's prediction was the optimistic one. The authors Khajanchi et al. [42] proposed another paper in which two mathematical models were developed to describe the dynamics of the virus in China described by systems of ODEs and curves constructed for the number of infected, recovered, and dead. The optimal values of the model parameters, which accurately describe the statistical data, were found. World Health Organization (WHO) data was used to obtain the model parameters, obtaining good agreement between the statistical data and the model curves. Thus, it is shown that there was a broad fit of the proposed mathematical model. This indicates a high adequacy of the mathematical model for coronavirus infection. Hamzah et al. [31] have developed a framework to manage and track COVID-19 data called CoronaTracker. This framework is based on a SEIR predictive compartmental model to predict the outbreak of COVID-19 inside and outside China based on daily observations, analyzing the influence of news of people's behavior both positively and economically. John Hopkins University (UJH), World Health Organization (WHO), and Ding Xiang Yuan databases were used as data sources. The data collected in CoronaTracker is available on the data lakes platform. For numerical simulation, the Scipy implementation was used, and for numerical integration, odeint was used. The study showed that the spread of the outbreak is influenced by the social policy of each country. The developed platform has an easy interface in which citizens can register their feelings and express their opinions about news articles. CoronaTracker can assist the government and authorities to disseminate articles, provide updates on the situation, and advocate good personal hygiene. This study has the limitation that when using data from John Hopkins University (UJH), an initial number of exposed individuals was missing. A decision-making system for COVID-19 (CDMS) was created by Varotsos and Krapivin [93] for USA, Brazil, Russia, and Greece. For the model with deterministic components, a compartmental SPRD model was created, similar to the classic SIRD in the literature, described by ODEs and with parameters determined by the reported data. And for the model with stochastic compartments, the classes were represented with a stability indicator that characterizes the COVID-19 propagation trend. Numerical evaluations were done by the SARD block using stochastic reports on the state of disease effects. This study showed that temperature and humidity slowly affect the effects of the pandemic. The analysis of the spread of the disease and the loss of income due to the pandemic has different impacts for each country. The analysis of official data from Russia and Greece showed the results of the

pandemic. The risk of infection and mortality increases with increasing population density. What limits this study is that there is not enough data to make the study reliable. In practice, it was impossible to coordinate measures to contain the COVID-19 pandemic under conditions of high uncertainty. In the study of Sadun [81] a compartmentalized SEIR model was developed. In this study strategies are developed to try to estimate the reproduction number R_0, and come to the conclusion that there is no direct way to measure it. The estimated value of R_0 depends on the length of the latency period for three versions of the classical SEIR model. The estimates of the reproduction number that have been published should be viewed with skepticism, one needs to understand the latency of COVID-19. However, there is no direct way to measure R_0, so what one can do is measure the time scale of the exponential growth of the pandemic and try to estimate R from it. The SEIPAHRF model was created by Ndaïrou et al. [67] to understand the transmission dynamics of COVID-19 in Wuhan. This model introduced a modification of the classical SEIR model by introducing asymptomatic (A), hospitalized (H), and fatality (F) infectious class. To study the basic reproduction number, a generation matrix was used in a sensitivity analysis. The local stability of the model was also studied. In this study, the theoretical findings and numerical results fit well with the actual results and reflect reality in Wuhan, China. This model can be used to study the reality in other countries whose outbreaks are increasing. However, the limited data at the beginning of the study, since it was early in the disease, was limited. Also to model and predict the dynamics of the COVID-19 pandemic in India, Sarkar et al. [85] created a 6-compartment model that extends the standard SEIR. And it divides the coronavirus-infected population from the susceptible individuals before the progression of clinical symptoms. It was also proven that quarantine decreases contact between uninfected and infected, and thus there is a reduction in the contact rate and can effectively reduce R_0. A sensitivity analysis was performed to recognize the most influential parameters with respect to the clinically infected population. And this sensitivity analysis was done by evaluating the technique of partial rank correlation coefficients (PRCC) for all input parameters in relation to variable I. The indices were evaluated at six time points: 30, 45, 60, 75, 90, and 100 days before steady state. The model variable was selected for sensitivity analysis I (infected or symptomatic individuals), generating six more influential parameters out of nine. And the actual daily COVID-19 data are fitted using least squares method (MMQ), which locally minimizes the sum of squares of errors. The numerical implementation was done in FORTRAN program. This model provides an important tool for assessing the consequences of possible policies, incorporating social distancing and blocking. Unfortunately, because of the short time scale, demographic effects are not considered. The Abou-Ismail [1] researchers focused on explaining and mathematically simplifying three models: SIR, SEIR, and SUQC (susceptible, unquarantined, quarantined, and confirmed). The goal was to understand the nature of the pandemic and to measure the impacts of social distancing through mathematical models. Making use of a system analysis of ODEs that describe the disease.

To analyze mathematically and do a numerical study the authors Viguerie et al. [95] created a new framework for understanding compartmental models by means of equilibrium equations similar to those found in Continuum Mechanics (Lotka–Volterra type). The model is SEIRD and made use of differential equation models to derive and analyze R_0. For models based on ODEs it has the concept of the basic reproduction number R_0 well defined, but the extension to a model based on PDEs is not clear due to the influence of diffusion. Therefore, in this work the EDO version of the EDP model was derived and its efficiency was evaluated with numerical tests. For the numerical tests either implicit second-order Backward Euler (BDF2) or implicit first-order Backward Euler was used. Picard linearization was performed at each time step. And the iterative Generalized Minimum Residual method (GMRES) with Jacobi preconditioning was used to solve all linear systems. PDE models are advantageous in that they allow a continuous space description of the relevant dynamics, allowing the dynamics to be described in time and space at all scales. Since models described by EDOs are limited for describing spatial information, implicit models are effective in describing the temporal dynamics of the system. In this model deaths other than by COVID-19 and births are not considered. The study developed by the researchers Khoshnaw et al. [47] used MATLAB's System Biology Tool (SBedit) package to compute the class dynamics of the model, and thus obtained a better understanding and identification of the key critical model parameters. And thus it was possible to understand the impacts of transmission rate and contact for New York. However, having several different models implies that one needs to create or identify the critical elements of each of the models. Furthermore, the model cannot simply be extrapolated to conditions in another country. Its parameters must be estimated from the new conditions. Making use of the same MATLAB package to obtain numerical solutions and calculate local sensitivity, Khoshnaw et al. [46] developed a model. Sensitivity analysis was done with the dynamics of the biological system modeled with law of mass action. This study also concluded that the most effective factors for the spread of coronavirus are: (1) the rate of person-to-person transmission, (2) the rate of quarantined exposure, and (3) the rate of transition from exposed individuals to individuals infected. MATLAB was used to numerically solve the compartmental model described by nonlinear differential equations proposed by Ahmed et al. [3]. And for the logistic model, the fitVirus function was used. The union of mathematical models and computer simulations is an effective tool that provides us with more understanding and good numerical predictions of the model states. However, in this study it is noticed that the number of people exposed to quarantine becomes stable after 40 days but the number of recovered people increases rapidly and becomes stable slowly. Having many approaches to identifying the estimates and understanding the disease makes the issue murky. However, this study has brought the identification of critical parameters of the model, helping to understand the overall issue more effectively and broadly. Using a code in MATLAB, Shao et al. [87] performed the numerical simulations. In this study, two time-delayed dynamic models were used to track COVID-19. The time-delayed dynamic coronavirus pneumonia model (TDD-NCP) introduced the delay process into the differential equations to describe the

latent period of the epidemic and can be used to predict the trend of coronavirus outbreak. Whereas the Fudan-Chinese Center for Disease Control and Prevention (CCDC) model was established to determine the kernel functions in the TDD-NCP model by the public data of CDCC, this model is suggested to use the time delay model to adjust the real data. The advantage of the Fudan-Chinese model is that it can track the initial date of the epidemic, when provided the $I(t_0)$. Moreover, this model can reconstruct parameters such as the growth rate and the "isolation rage," and predict the cumulative number of confirmed cases in some cities in China. However, because this work was done in early March, there was still little knowledge about the disease and little data on confirmed cases. Rajagopal et al. [75] have developed a SEIRD model with integer and fractional differential equations to describe coronavirus in Italy. The fractional model is of the Caputo type, the most popular and most widely used for real problems. To find the optimal parameters, the model parameters are estimated. The number of infected, the number of deaths, and the associated mean square error (RMSE) are also considered. The fractional model gives more realistic predictions and has fewer modeling errors. And with that, the proposed model agrees with the actual data from Italy better than the classical model. A SCEAQHR model for predicting cases in Cameroon has been proposed by Nabi et al. [65]. This model integrates a new class for individuals who have made imperfect quarantine and disregarded blocking policies. The model parameters were estimated with real-time data, followed by a projection of the disease evolution. The model is described by Caputo fractional differential equations, and the existence and uniqueness of the solutions are presented. The optimization algorithm is based on the reliable-region-reflective (TRR) algorithm, which is the evolution of the Levenberg–Marquardt algorithm. The numerical implementation is done using the lsqcurve fit function of MATLAB. The Partial Rank Correlation Coefficient (PRCC) method was used to quantify the dominant mechanisms. The optimization is robust to solving nonlinear least squares problems.

The researchers Roda et al. [78] used the Akaike Information Criterion (AIC) to select the model. And performed an analysis of the predictions of the SIR and SEIR models. The SIR model outperformed the SEIR model in representing the information contained in the confirmed case data. The calibration of the model was done using the Monte Carlo Markov Chain algorithm, and the calibration was done with data from January 21 to February 04 from Wuhan city in China. The authors state that data before January 23 is unreliable and there is a lack of data. There is no identifiability because a group of model parameters cannot be determined solely from the data provided during model calibration. This impacts the reliability of the model.

Din et al. [22] brought out a new three-compartment model (PIQ) described by EDPs for COVID-19 transmission. To study the stability, the Atangana, Baleanu, and Caputo (ABC) model with arbitrary order was used. Banach's fixed point theorem and Guo–Krasnoselskii were used to prove the existence of the model. And the numerical simulations were done using the Adams–Bashforth (AB) method with fractional differentiation. Using this method is a sophisticated and powerful tool for

investigating nonlinear problems. The model proves mathematically that it is well defined.

Through a system of ordinary differential equations, the disease is contextualized through social parameters to understand how the spread works and how it is possible to control the epidemics that affect society and thereby create preventive measures. Examples of this type of model are the modified SEIR models proposed by Yang et al. [101] as well as the SEIR (Susceptible, Exposed, Infectious, Recovered) model with age-structured quarantine class with the two types of control measures used to analyze the effects of policy control for the coronavirus epidemic in Brazil [15], and the SEIRQ (Susceptible, Exposed, Infectious, Recovered, Quarantine) model with age structure, proposed by Gondim and Machado [29]. This model aims to analyze optimal quarantine strategies in order to help in decision-making through health managers.

Regarding statistical epidemiological models, Sarkar et al. [85] propose a mathematical model to monitor the dynamics of six compartments: Susceptible (S), Asymptomatic (A), Recovered (R), Infected (I), Isolated Infected (Iq), and Quarantined Susceptible (Sq), collectively expressed SARIIqSq. The authors applied their proposal to real data on the COVID-19 pandemic in India. Starting from the date of first COVID-19 case reported in India, the authors have simulated the SARIIqSq model for 260 days for each states and for whole India to study the dynamics of the SARS-CoV-2 disease. They statistically confirmed that a reduction in the contact rate between uninfected and infected individuals by quarantined the susceptible individuals can effectively reduce the basic reproduction number. They also demonstrate that the elimination of ongoing SARS-CoV-2 pandemic is possible by combining the restrictive social distancing and contact tracing. However, the authors also emphasize the uncertainty of accessible authentic data, specially concerning to the accurate baseline number of infected individuals due to subnotifications, which may guide to equivocal outcomes and inappropriate predictions by orders of size.

Ndaïrou et al. [67] propose a novel epidemiological compartment model that takes into account the super-spreading phenomenon of some individuals. They consider a fatality compartment, related to death due to the virus infection. The constant total population size N is subdivided into eight epidemiological classes: Susceptible class (S), Exposed class (E), Symptomatic and Infectious class (I), Super-Spreaders class (P), Infectious but Asymptomatic class (A), Hospitalized (H), Recovery class (R), and Fatality class (F). This model reached a reasonably good approximation of the reality of the Wuhan outbreak, predicting a diminishing on the daily number of confirmed cases of the disease. The model also fits well the real data of daily confirmed deaths. The model can be considered useful for other realities than Wuhan, China, since the amount of hospitalized individuals is relevant as an estimate of the Intensive Care Units (ICU) needed.

Khajanchi and Sarkar [43] developed a new compartmental model to explain the transmission dynamics of Covid-19. They calibrated their model with daily Covid-19 data for four Indian states: Jharkhand, Gujarat, Andhra Pradesh, and Chandigarh. They studied the feasible equilibria of the proposed model and their stability with

respect to the basic reproduction number R_0. The disease-free equilibrium becomes stable and the endemic equilibrium becomes unstable when the recovery rate of infected individuals increases, but if the disease transmission rate remains higher, then the endemic equilibrium always remains stable. The proposed model obtained $R_0 > 1$ for all studied Indian states, suggesting a significant outbreak. The model is able to provide short-time Covid-19 forecasting as well.

Samui et al. [84] proposed a deterministic ordinary differential equation model able to represent the overall dynamics of SARS-CoV-2. They stratified the total human population into four compartments: susceptible individuals (uninfected), asymptomatic individuals (pauci-symptomatic or clinically undetected), reported symptomatic infected individuals (symptomatic infectious individuals are reported by the public health service), and unreported symptomatic infected individuals (clinically ill but not reported) to formulate the SAIU (susceptible or uninfected (S), asymptomatic (A), reported symptomatic infectious (I), unreported symptomatic infectious (U)) model. This model assumes that infected individuals informed will no longer be associated with infections, as they are isolated or transferred to Intensive Care Units (ICU). Thus, only infectious individuals belonging to I(t) or U(t) spread or transmit the diseases. The authors designed the SAIU model to study the transmission dynamics of COVID-19 based on the accessible data for India during the time period January 30, 2020 to April 30, 2020. Based on the estimated data, the SAIU model predicts the outbreak of COVID-19 and computes the basic reproduction number R_0. The authors assessed the sensitivity indices of the basic reproductive number R_0, given that R_0 expresses the initial disease transmission and the sensitivity indices describe the relative importance of various parameters in coronavirus transmission. The SAIU model showed the persistence of diseases for $R_0 > 1$. The endemic equilibrium point $E*$, for this study, was locally asymptotically stable for $R_0 > 1$.

Khajanchi et al. [44] extended the classical deterministic Susceptible–exposed–infectious–removed (SEIR) compartmental model refined by introducing contact tracing-hospitalization strategies to study the epidemiological properties of Covid-19. They calibrated their mathematical model using data of confirmed cases in India and estimated the basic reproduction number for the disease transmission. The authors have their calibrated epidemic model for the short term prediction in the four provinces and the Republic of India. The simulation of the calibrated model was able to capture the increasing growth patterns for three different provinces, namely Delhi, Maharashtra, West Bengal and the Republic of India, whereas in case of the province Kerala, the model fitting is not good compared to other states and overall India. Model simulation and prediction suggest that Covid-19 has a potential to exhibit oscillatory but controllable dynamics in the near future by maintaining social distancing and effectiveness of home isolation and hospitalization. The proposed model forecasts that isolation or hospitalization of the symptomatic population, under stringent hygiene safeguards and social distancing, is considerably effective. Finally, Khajanchi et al. [44] give evidences that the size and duration of an epidemic can be considerably affected by timely implementation of the hospitalization or isolation program.

The classic mathematical models of epidemiological prediction are quite useful, but deterministic, demonstrating only the average behavior of the epidemic, which makes it difficult to quantify uncertainty. Wang et al. [97] proposed an analysis of the spatial structure and dynamics of the spread of Covid-19, providing a spatio-temporal prediction of the Covid-19 outbreak in the USA. Kapoor et al. [39] investigated large-scale spatio-temporal prediction using neural network graphs and human mobility data in US counties. Through this method and space-time information, the model learns the epidemiological dynamics. Tomar and Gupta [92] proposed a space-time approach to control and monitor Covid-19 using LSTM (Long Short-Term Memory) neural networks and adjusting curves to predict chaos. Ren et al. [76] used Ecological Niche Models (ENM) to gather epidemiological and socioeconomic data, aiming to accurately predict the risk areas for Covid-19 infection. Yesilkanat [102] made a study with space-time approach for 190 countries in the world and compared it with the number of real cases of the disease using the Random Forest method. Also using a space-time approach, Pourghasemi et al. [70] did a risk mapping, change detection and trend analysis of the Covid-19 spread in Iran using regression and machine learning. Roy et al. [79] developed a short-term prediction model for the new Coronavirus using canonical ARIMA (Autoregressive Integrated Moving Average) and disease risk analysis done using weighted overlap analysis in geographic information systems.

3   Forecasting by Machine Learning and Hybrid Approaches

Several efforts to aid Covid-19 screening and monitoring can be perused in the works of Dong et al. [25]. In this work, Dong et al. [25] created an online interactive panel to visualize Covid-19 infected cases and deaths in real time, providing researchers, health authorities, and the general public a tool to track cases as the disease progresses. Due to the rapid development of the coronavirus, the need to classify infected patients and analyze which individuals were more vulnerable to the disease also grew. Therefore, Xie et al. [100] proposed a model of clinical prediction for patient mortality based on multivariable logistic regression, to improve the use of limited healthcare resources and calculate the patient's survival rate. Furthermore, in order to aid the diagnosis, Feng et al. [26] developed the online calculator S-COVID-19-P based on Lasso regression, for early identification of suspected Covid-19 pneumonia in the admission of adult patients with fever. Jin et al. [38] proposed a system based on deep learning for the rapid diagnosis of Covid-19 with precision comparable to experienced radiologists, and can accurately classify pneumonia, CAP (Community-Acquired Pneumonia), influenza A and B, and Covid-19. They used LASSO to find the 12 most discriminating characteristics in the distinction between Covid-19 and other pneumonias. Gomes et al. [28] proposed a system to support the diagnosis of Covid-19 by analyzing chest X-ray images, capable of differentiating Covid-19 from bacterial and viral pneumonias using texture-based image representation and classification by Random Forests.

Different from other more complex Covid-19 X-ray feature extraction approaches [7, 8, 12, 19, 33, 35, 45, 53, 54, 63, 66, 96], Gomes et al. [28] avoided deep learning based solutions and adopted texture and shape features to provide the users a low-cost computational web-based computational environment able to deal with several simultaneous users without overcharging network resources.

In order to find a new way to perform early, efficient, and accurate control and screening of suspected individuals, Meng et al. [62] created the Covid-19 Diagnostic Aid APP to calculate the probability of infection through simple and easy laboratory test results. Screening a large number of suspicious people could optimize the diagnostic process and save medical resources. Barbosa et al. [10] considered the fact that, in many regions of the world, RNA testing is not always available due to the scarcity of inputs, created HegIA, an intelligent system based on Bayes Networks and Random Forests to aid at the diagnosis of Covid-19 based on blood tests from 24 blood tests. The performance is close to RT-PCR (Reverse Transcription Polymerase Chain Reaction) for symptomatic individuals, though coronavirus RNA is not searched [10]. HegIA is a fully functional system, available for free use, to provide low-cost rapid testing.

Several works have used Evolutionary Computing and Swarm Intelligence Methods to automatically adjust compartmental models [61, 71, 73, 83]. Putra and Khozin Mu'tamar [71] automatically estimated parameters in the Susceptible, Infected, Recovered (SIR) model using the Particle Swarm Optimization (PSO) algorithm. Their results suggest that the proposed method is able to tune SIR models precisely compared to other analytical approaches. Similarly, Mbuvha and Marwala [61] calibrated a SIR model to South Africa's Covid-19 reported cases taking into account several scenarios of the reproduction number R_0 for reporting infections and healthcare resource estimations. They assumed that the reported confirmed cases represent between 0.2% and 1% of the total infected population. The authors also assumed that SIR model parameters are fixed albeit at multiple ranges. However, they detected the uncertainty around SIR parameters and propose a Bayesian treatment using Markov Chain Monte Carlo techniques in the near future.

Qi et al. [73] investigated the influence of daily temperature (AT) and relative humidity (ARH) on the occurrence of Covid-19 in 31 Chinese provinces, mainly in Hubei. The authors collected daily counts of laboratory-confirmed cases in all provinces in China from the official reports of the National Health Commission of People's Republic of China from December 1, 2019 to February 11, 2020 for Hubei province and from January 20, 2020 to February 11, 2020 for other provinces. Tibet was not included in the following model since only one case was reported during the 23-day cited period. The meteorological data, including daily average temperature (AT) and daily average relative humidity (ARH) of each provincial capital, were retrieved from Weather Underground. Although this study suggests that both daily temperature and relative humidity influenced the occurrence of COVID-19 in Hubei province and in some other provinces, the association between COVID-19 and AT and ARH across the provinces was not considered consistent. The authors found

spatial heterogeneity of COVID-19 incidence, as well as its relationship with daily AT and ARH, among provinces in Mainland China.

Salgotra et al. [83] propose prediction models based on genetic programming (GP) for confirmed cases and death cases across the three most affected states in India: Maharashtra, Gujarat, and Delhi. The authors also applied the model to forecast Covid-19 cases in whole India. The proposed prediction models are presented using explicit formula. The authors studied the impotence of prediction variables as well. Statistical parameters and metrics have been used to evaluate and validate the evolved models. Genetic evolutionary programming models have proven to be highly reliable for Covid-19 cases in India.

Rahimi et al. [74] present a systematic review on Computational Intelligence algorithms for Covid-19 forecasting. They searched on Web of Science (WoS) and Scopus for publications in accordance with the following keywords: forecasting, prediction, Covid-19, and coronavirus. The authors selected 920 technical research articles presenting just algorithmic descriptions, review articles, conference papers, case studies, and able to provide managerial insights, published until October 10, 2020. The authors focused on papers indexed by the Web of Science. Rahimi et al. [74] categorized the main forecasting works according to the following classification regarding the algorithms:

– Simple Moving Average [16] as defined by Maleki and Arellano-Valle [55], Maleki and Nematollahi [58], Zarrin et al. [103], Maleki et al. [56], and Hajrajabi and Maleki [30];
– Auto-Regressive Integrated Moving Average (ARIMA) [5, 50, 64, 80, 88];
– Two-piece distributions based on the scale [57];
– Logistic functions: S-shaped functions to model epidemiological curves [17, 52, 72];
– Regression Methods [4, 36, 77, 90, 94];
– Canonical neural networks [27, 64, 91];
– Deep learning methods based on Convolutional Neural Networks (CNNs) [13, 51, 86];
– Deep learning methods based on Long-Short Term Memory (LSTM) neural networks [9, 18];
– Genetic programming [82, 83]);
– Classical and modified compartment models: SIR, SEIR, and SIRD [2, 14, 41, 69].

Tamang et al. [91] used artificial neural network-based curve fitting techniques to predict and forecast Covid-19 infected and death cases in India, USA, France, and United Kingdom, considering the progressive trends of China and South Korea. The authors considered three cases to analyze the Covid-19 outbreak: (1) forecasting as per the present trend of rising cases of different countries; (2) one-week forecasting following up with the improvement trends as per China and South Korea; and (3) forecasting if followed up the progressive trends as per China and South Korea before a week. According to the authors, to reduce infection rates and achieve leveling of trends in epidemiological curves, these countries will require fewer

days according to the forecast with the trend in China and more days with steady progress are seen with the South Korea's trend. In addition, it can also be concluded that, with the trend of China, countries with a greater number of cases could be better in fewer days with possibly stricter measures of social isolation, detachment, and confinement. Considering that South Korea's trend is toward slower and more constant control, which could be more effective in the initial stage with lower reported cases. All conclusions were made in accordance with the predictions obtained with the application of the multilayer perceptron artificial neural network technique. Although the case data used in the study are based on reliable sources, the predictions are in accordance with the conditions and techniques applied. Consequently, their experimental results suggest that artificial neural networks are able to forecast the future cases of COVID 19 outbreak of practically any country at low error rates.

Huang et al. [32] propose a new model of CNN deep neural network with multiple inputs to predict the cumulative number of confirmed cases of Covid-19. The cumulative number of confirmed cases on the following day is predicted according to the total number of confirmed cases from the previous 5 days, total new confirmed cases, total cured cases, total new cured cases, total deaths, and total new deaths. Datasets from seven Chinese cities in the provinces of Hubei, Guangdong, and Zhejiang were used with confirmed serious cases for the training and forecasting of the models. Data on confirmed cases of COVID-19 from January 23, 2020 to March 2, 2020, and from January 23, 2020 to March 2, 2020, were obtained from the media outlet Surging News Network and from the World Health Organization, respectively. The two evaluation indexes of the mean absolute error (MAE) and root mean square error (RMSE) were used. According to the authors, the proposed algorithm can quickly use small datasets to establish models with high predictive precision. This is a considerable advantage of this model over other models with similar characteristics. Through the proposed algorithm, a prediction model was established for the number of confirmed cases of COVID-19. Verification and comparison were conducted between different deep learning algorithms. The accuracy and reliability of the deep learning algorithm have been verified by predicting the future trend of Covid-19. In addition, experiments for several cities with more serious confirmed cases in China indicated that the prediction model in this study had the lowest error rate among its tested equivalents. As future work, the authors envisage using deep learning networks with a mixed structure, seeking to build more accurate models, which can be applied to more countries.

Distante et al. [23] modeled spreading of Covid-19 using Chinese data and used the model to predict epidemic curve in each Italian region, allowing to gain better information on the new daily cases peaks with the predicted epidemiological curve. According to the authors, the forecast portion of the curve allows to have a better prediction of active cases with the SEIR model, by computing the position of the peaks of active cases for each Italian region. Interestingly, the process of training on Chinese data and using the knowledge to forecast Italian spreading of Covid-19 has resulted in good forecasting results, considering the mean average precision between official Italian data and the forecast. SEIR models may fit better than

other compartment models since they are based on the complete curve dynamic. Therefore, the proposed approach is valid since the predictive model learns from the dynamics of Covid-19 in China and exploits its knowledge to predict future daily cases in Italy.

Wieczorek et al. [99] proposes a predictive model based on a deep 7-layer neural network trained by the NAdam method to predict the number of infected cases. The authors used a dataset provided by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University on their github page. This dataset is composed of the following sources: (a) World Health Organization (WHO); (b) European Center for Disease Prevention and Control (ECDC); (c) DXY.cn. Pneumonia. 2020; (d) COVID Tracking Project; (e) National Health Commission of the People's Republic of China (NHC); (f) China CDC (CCDC); (g) Washington State Department of Health; (h) other smaller, regional US health departments. The predictive model was able to predict new cases with very high efficiency, above 99% in some geographic regions. However, the authors noticed that analysts should take into account several factors able to influence the epidemiological curves: behavior of the population in a given region, behavior of governments of given countries as well as access to knowledge and medical equipment. The neural network-based predictor employs a unified architecture. According to their experimental results, the authors do not need to change the architecture in dependence with each region or country. Accuracy for most of regions is around 87.70%. However, the authors believe that dedicated architectures should be used to contemplate differences among countries, like population and government behaviors.

Kırbaş et al. [49] modeled confirmed COVID-19 cases of Denmark, Belgium, Germany, France, United Kingdom, Finland, Switzerland, and Turkey using Auto-Regressive Integrated Moving Average (ARIMA), Nonlinear Autoregression Neural Network (NARNN) and Long-Short Term Memory (LSTM) approaches. They tested six model performance metrics: MSE, PSNR, RMSE, NRMSE, MAPE, and SMAPE. Cumulative confirmed case data of eight different European countries were used for modeling: Denmark, Belgium, Germany, France, United Kingdom, Finland, Switzerland, and Turkey. The datasets were acquired from the European Center for Disease Prevention and Control. Data were taken from the day the first case was seen, and the number of data for each country varies. The data covers 67, 90, 97, 100, 94, 90, 68, and 55 days, respectively, and ends on 3 May 2020. The data from cumulative confirmed cases in some European countries are modeled using three different approaches. According to the results, it was determined that LSTM approach has much higher success compared to ARIMA and NARNN. The lowest number of cases was observed in Finland during the epidemic, while the highest rate of increase was observed in the United Kingdom. According to the 2-week prospective estimation study, in many countries, the total case increase rate is expected to decrease slightly. Since the work was carried out entirely by considering statistical data and methodologies, the effects of social distancing and other similar measures, compliance with hygiene rules or lockdown were ignored. However, according to the results on real data, the authors considered the predictions satisfactory.

Pal et al. [68] have proposed to use the local data trend with a shallow Long Short-Term Memory (LSTM) based neural network combined with a fuzzy rule based system to predict long term risk of a country. The country-specific neural networks are optimized using Bayesian optimization. The authors used the dataset (https://github.com/datasets/covid-19) that included date, country, the number of confirmed cases, the number of recovered cases, and the total number of deaths. This data was combined with weather data (https://darksky.net/): humidity, dew, ozone, perception, maximum temperature, minimum temperature, and UV for analyzing the effect of weather. The authors considered mean and standard deviation over different cities of a country. The data spanned the duration 22-01-2020 to 02-08-2020. The authors propose to use country-specific optimized networks for accurate prediction, since this approach seems suitable for small and uncertain dataset. Combining the overall optimized LSTMs, they noticed that a shallow networks perform better compared to deep neural networks. The authors also noticed that the weather data does not affect the forecasting accuracy.

Zeroual et al. [104] performed a comparative study of five deep learning methods to forecast the number of new cases and recovered cases: simple Recurrent Neural Network (RNN), Long short-term memory (LSTM), Bidirectional LSTM (BiLSTM), Gated recurrent units (GRUs), and Variational AutoEncoder (VAE). These methods were applied for global forecasting of Covid-19 cases based on a small volume of data. This study is based on daily confirmed and recovered cases collected from six countries namely Italy, Spain, France, China, USA, and Australia. The values of parameters of deep learning models are selected such that the loss function is minimized during the training. The authors adopted the Adam optimizer. In the testing stage, the previously constructed models with the selected parameters are used to forecast the number of COVID cases. The accuracy of the model was verified by comparing the measured data with real data via different statistical indicators including RMSE, MAE, MAPE, and RMSLE (Root Mean Squared Log Error). The research was based on daily figures of confirmed and recovered cases collected from six highly impacted countries namely Italy, Spain, Italy, China, the USA, and Australia. The considered datasets are gathered from the starting of COVID-19 for the respective countries, i.e. 22 January 2020, till June 17th, 2020. These datasets are made publically by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (https://github.com/CSSEGISandData). Results demonstrate that the Variational AutoEncoder achieved the best forecasting performance in comparison to the other models.

Kapoor et al. [40] propose a novel spatio-temporal forecasting approach for Covid-19 case prediction based on Graph Neural Networks and mobility data. Differently from time series forecasting models, the proposed model learns from a single large-scale spatio-temporal graph, where nodes represent the region-level human mobility, spatial edges represent the human mobility based inter-region connectivity, and temporal edges represent node features through time. The authors applied their method to the US county level COVID-19 dataset. They perceived that the spatial and temporal information leveraged by the graph neural network allows the model to learn considerably complex dynamics. It is noticed a 6%

reduction of RMSLE and an absolute Pearson Correlation improvement from 0.9978 to 0.998 in comparison with the state-of-the-art models. According to the authors, the combination of graph-based deep learning approaches can be very useful to aid to understand the spread and evolution of Covid-19.

de Lima et al. [21] proposed a real-time surveillance, forecast, and spatial visualization of Covid-19, named COVID-SGIS. As a case study, the forecasting system was applied to monitor Brazil. The system captures routinely reported Covid-19 information from 27 federative units from the Brazil.io database. It uses Covid-19 confirmed case data notified through Brazil's National Notification System, SINAN, from March to May 2020. Time series ARIMA models were integrated to forecast the cumulative number of Covid-19 cases and deaths. These include 6-days forecasts as graphical outputs for each federal state in Brazil, separately, with its corresponding 95% confidence interval. The worst and the best scenarios are both presented. The overall percentage error between the forecasted values and the actual values varied between 2.56% and 6.50%. For the days when the forecasts fell outside the forecast interval, the percentage errors in relation to the worst case scenario were below 5%. Considering the good results obtained with the proposed tool, the authors claimed that the proposed method for dynamic forecasting may be used to guide social policies and plan direct interventions in a cost-effective, concise, and robust manner.

4   Conclusions

COVID-19 is a disease that was discovered and soon assumed pandemic status as it spread to several countries around the world. It drew attention for its ease of transmission and for exposing the vulnerabilities of health systems around the world. The individuals who were infected and their families were left with the pain and suffering and the after-effects of the disease. Although there are vaccines, there is still no proven effective drug against the disease, so following safety protocols and social isolation are indispensable. In addition to hygiene practices such as the use of masks and hand-washing, the use of models to understand the behavior of the disease and even to predict it helps to shed light on the next steps to be taken in this pandemic.

The representation of disease through mathematical models facilitates monitoring and can help analyze and understand disease dynamics through key characteristics. Through a system of equations, it is possible to model a disease and contribute to a quantitative understanding. These characteristics become useful information on how the spread of the disease works. They also make it possible to understand how to build time prediction and thus help to create measures to control and prevent COVID-19. However, for this type of modeling, some assumptions are made, such as assuming that disease transmission occurs homogeneously, or selecting only one among several climatic factors. Therefore, this limits the model's ability to predict. And if more features were added, the model would lose robustness.

With the large amount of data available and thanks to speed and storage technologies, Artificial Intelligence is increasingly strong and present in several areas. Then, the use of machine learning techniques grew in order to obtain insights from this data. These models are applied in several areas, such as economics, for the performance of a stock in the stock market, in banking, in e-commerce, determining whether a customer will like the product or not, in health, as is done in the present work with digital epidemiology. But models with this kind of approach are black boxes, that is, they are not intelligible to experts, because their goal is to correctly map inputs to outputs. Another type of approach using Artificial Intelligence is the one that uses hybrid models, i.e. it combines machine learning models with statistical models. By doing this they combine the advantages of each of these types of models in order to obtain a more robust prediction model. Another type of hybrid model is one that combines compartmental models and machine learning. With this approach the model does not have as good a prediction quality as machine learning based models; however, it can aid in human understanding of epidemiological aspects as phenomena, while machine learning based models can return accurate predictions, thus combining intelligent systems for accurate human learning emergent predictions. The use of all these approaches is very important to support us in temporal and spatio-temporal prediction of cases and deaths. For these solutions can shed light on strategies to assist decision-making by health managers.

Finally, COVID-19 brings with it all the challenges of a new disease with only 1 year of existence, in facing this unknown, science makes use of all its arsenal. At this time when there is no extensive background to teach how the disease behaves, daily experience determines adjustments and creation of clinical protocols. Predicting the temporal and spatial behavior of COVID-19 through machine learning becomes a valuable tool to guide strategies, policies, and hope.

References

1. Abou-Ismail, A. (2020). Compartmental models of the COVID-19 pandemic for physicians and physician-scientists. *SN Comprehensive Clinical Medicine*, 2, 852–858.
2. Ahmar, A. S., & Del Val, E. B. (2020). SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain. *Science of The Total Environment*, *729*, 138883.
3. Ahmed, A., Salam, B., Mohammad, M., Akgul, A., & Khoshnaw, S. (2020). Analysis coronavirus disease (covid-19) model using numerical approaches and logistic model. *Aims Bioengineering*, *7*(3), 130–146.
4. 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.
5. Alzahrani, S. I., Aljamaan, I. A., & Al-Fakih, E. A. (2020). Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of Infection and Public Health*, *13*(7), 914–919.
6. Ambikapathy, B., & Krishnamurthy, K. (2020). Mathematical modelling to assess the impact of lockdown on covid-19 transmission in India: Model development and validation. *JMIR Public Health and Surveillance*, *6*(2), e19368.

7. Apostolopoulos, I., Aznaouridis, S., & Tzani, M. (2020). Extracting possibly representative covid-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. Preprint. arXiv:2004.00338.

8. Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, 43(2), 635–640.

9. Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & Kalhori, S. R. N. (2020). Predicting COVID-19 incidence through analysis of google trends data in Iran: data mining and deep learning pilot study. *JMIR Public Health and Surveillance*, 6(2), e18828.

10. Barbosa, V. A. d. F., Gomes, J. C., de Santana, M. A., Jeniffer, E. d. A., de Souza, R. G., de, Souza, R. E., & dos Santos, W. P. (2021). Heg.IA: An intelligent system to support diagnosis of Covid-19 based on blood tests. *Research on Biomedical Engineering*, 2021, 1–18.

11. Bastos, S. B., & Cajueiro, D. O. (2020). Modeling and forecasting the early evolution of the covid-19 pandemic in Brazil. *Scientific Reports*, *10*(1), 1–10.

12. Basu, S., Mitra, S., & Saha, N. (2020). Deep learning for screening covid-19 using chest X-ray images. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 2521–2527).

13. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828.

14. Capasso, V., & Serio, G. (1978). A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences*, *42*(1–2), 43–61.

15. Castilho, C., Gondim, J. A. M., Marchesin, M., & Sabeti, M. (2020). Assessing the Efficiency of Different Control Strategies for the Coronavirus (COVID-19) Epidemic. Preprint. arXiv:2004.03539. Retrieved from http://arxiv.org/abs/2004.03539

16. Chaudhry, R. M., Hanif, A., Chaudhary, M., & Minhas, S. (2020). Coronavirus Disease 2019 (COVID-19): Forecast of an emerging urgency in Pakistan. *Cureus*, *12*(5).

17. Chen, D.-G., Chen, X., & Chen, J. K. (2020). Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Global Health Research and Policy*, *5*, 1–7.

18. Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, *135*, 109864.

19. Civit-Masot, J., Luna-Perejón, F., Domínguez Morales, M., & Civit, A. (2020). Deep learning system for covid-19 diagnosis aid using X-ray pulmonary images. *Applied Sciences*, *10*(13), 4640.

20. Coronavirus disease (covid-19) pandemic [Computer software manual]. (2020). Retrieved from www.who.int/emergencies/diseases/novel-coronavirus-2019. Last accessed: 22 April 2020.

21. de Lima, C. L., da Silva, C. C., da Silva, A. C. G., Silva, E. L., Marques, G. S., de Araújo, L. J. B., . . . da Silva-Filho, A. G. (2020). COVID-SGIS: A smart tool for dynamic monitoring and temporal forecasting of Covid-19. *Frontiers in Public Health*, *8*, 761.

22. Din, A., Shah, K., Seadawy, A., Alrabaiah, H., & Baleanu, D. (2020). On a new conceptual mathematical model dealing the current novel coronavirus-19 infectious disease. *Results in Physics*, *19*, 103510.

23. Distante, C., Pereira, I. G., Goncalves, L. M. G., Piscitelli, P., & Miani, A. (2020). Forecasting Covid-19 outbreak progression in Italian regions: A model based on neural network training from Chinese data. MedRxiv.

24. Döhla, M., Boesecke, C., Schulte, B., Diegmann, C., Sib, E., Richter, E., . . . et al. (2020). Rapid point-of-care testing for SARS-CoV-2 in a community screening setting shows low sensitivity. *Public Health*, *182*, 170–172.

25. Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533–534. Retrieved from https://doi.org/10.1016/S1473-3099(20)30120-1

26. Feng, C., Huang, Z., Wang, L., Chen, X., Zhai, Y., Zhu, F., . . . Li, T. (2020). A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. *medRxiv*, 1–68. Retrieved from https://ssrn.com/abstract=3551355

27. Fong, S. J., Li, G., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2020). Finding an accurate early forecasting model from small dataset: A case of 2019-nCoV novel coronavirus outbreak. Preprint. arXiv:2003.10776, *2020*.

28. Gomes, J. C., de Freitas Barbosa, V. A., de Santana, M. A., Bandeira, J., Valenca, M. J. S., de Souza, R. E., . . . dos Santos, W. P. (2020). Ikonos: An intelligent tool to support diagnosis of covid-19 by texture analysis of X-ray images. *Research on Biomedical Engineering*, *2020*, 1–14.

29. Gondim, J. A. M., & Machado, L. (2020). Optimal quarantine strategies for the COVID-19 pandemic in a population with a discrete age structure. Preprint. arXiv: 2005.09786.

30. Hajrajabi, A., & Maleki, M. (2019). Nonlinear semiparametric autoregressive model with finite mixtures of scale mixtures of skew normal innovations. *Journal of Applied Statistics*, *2019*, 2010–2029.

31. Hamzah, F. B., Lau, C., Nazri, H., Ligot, D. V., Lee, G., Tan, C. L., . . . et al. (2020). Coronatracker: worldwide covid-19 outbreak data analysis and prediction. *Bull World Health Organ*, *1*(32), 1–32.

32. Huang, C.-J., Chen, Y.-H., Ma, Y., & Kuo, P.-H. (2020). Multiple-input deep convolutional neural network model for Covid-19 forecasting in China. *MedRxiv*, *2020*, 1–16.

33. Ismael, A. M., & Şengür, A. (2021). Deep learning approaches for covid-19 detection based on chest X-ray images. *Expert Systems with Applications*, *164*, 114054.

34. Ivorra, B., Ferrández, M. R., Vela-Pérez, M., & Ramos, A. (2020). Mathematical modeling of the spread of the coronavirus disease 2019 (covid-19) taking into account the undetected infections. The case of China. *Communications in Nonlinear Science and Numerical Simulation*, *88*, 105303.

35. Jain, G., Mittal, D., Thakur, D., & Mittal, M. K. (2020). A deep learning approach to detect covid-19 coronavirus with X-ray images. *Biocybernetics and Biomedical Engineering*, *40*(4), 1391–1405.

36. Ji, D., Zhang, D., Xu, J., Chen, Z., Yang, T., Zhao, P., . . . Qin, E. (2020). Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clinical Infectious Diseases*, *71*(6), 1393–1399.

37. Jiang, N., Liu, Y., Yang, B., Li, Z., Si, D., Ma, P., . . . & Yu, Q.(2020). Analysis of the factors associated with negative conversion of severe acute respiratory syndrome coronavirus 2 rna of coronavirus disease 2019. *Open Access Macedonian Journal of Medical Sciences*, *8*(1), 436–442.

38. Jin, C., Chen, W., Cao, Y., Xu, Z., Zhang, X., Deng, L., . . . Feng, J. (2020). Development and evaluation of an AI system for COVID-19 diagnosis. *medRxiv*. Retrieved from http://medrxiv.org/content/early/2020/03/27/2020.03.20.20039834.abstract

39. Kapoor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais, M., & O'Banion, S. (2020a). Examining COVID-19 forecasting using spatio-temporal graph neural networks. ArXiv preprint. Retrieved from http://arxiv.org/abs/2007.03113

40. Kapoor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais, M., & O'Banion, S. (2020b). Examining covid-19 forecasting using spatio-temporal graph neural networks. Preprint. arXiv:2007.03113, *2020*.

41. Kermack, W. O., & McKendrick, A. G. (1932). Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *138*(834), 55–83.

42. Khajanchi, S., Bera, S., & Roy, T. K. (2021). Mathematical analysis of the global dynamics of a HTLV-I infection model, considering the role of cytotoxic t-lymphocytes. *Mathematics and Computers in Simulation*, *180*, 354–378.

43. Khajanchi, S., & Sarkar, K. (2020). Forecasting the daily and cumulative number of cases for the covid-19 pandemic in India. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(7), 071101.

44. Khajanchi, S., Sarkar, K., Mondal, J., & Perc, M. (2020). Dynamics of the covid-19 pandemic in India. Preprint. arXiv:2005.06286.

45. Khan, A. I., Shah, J. L., & Bhat, M. M. (2020). Coronet: A deep neural network for detection and diagnosis of covid-19 from chest X-ray images. *Computer Methods and Programs in Biomedicine*, *196*, 105581.

46. Khoshnaw, S. H., Salih, R. H., & Sulaimany, S. (2020). Mathematical modelling for coronavirus disease (covid-19) in predicting future behaviours and sensitivity analysis. *Mathematical Modelling of Natural Phenomena*, *15*, 33.

47. Khoshnaw, S. H., Shahzad, M., Ali, M., & Sultan, F. (2020). A quantitative and qualitative analysis of the covid-19 pandemic model. *Chaos, Solitons & Fractals*, *138*, 109932.

48. Khrapov, P., & Loginova, A. (2020). Comparative analysis of the mathematical models of the dynamics of the coronavirus covid-19 epidemic development in the different countries. *International Journal of Open Information Technologies*, *8*(5), 17–22.

49. Kırbaş, İ., Sözen, A., Tuncer, A. D., & Kazancıoğlu, F. Ş. (2020). Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons & Fractals*, *138*, 110015.

50. Kufel, T. (2020). ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(2), 181–204.

51. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

52. Li, Q., Feng, W., & Quan, Y.-H. (2020). Trend and forecasting of the COVID-19 outbreak in China. *Journal of Infection*, *80*(4), 469–496.

53. Luz, E., Silva, P. L., Silva, R., & Moreira, G. (2020). Towards an efficient deep learning model for covid-19 patterns detection in X-ray images. Preprint. arXiv:2004.05717.

54. Maghdid, H. S., Asaad, A. T., Ghafoor, K. Z., Sadiq, A. S., & Khan, M. K. (2020). Diagnosing covid-19 pneumonia from x-ray and ct images using deep learning and transfer learning algorithms. Preprint. arXiv:2004.00038.

55. Maleki, M., & Arellano-Valle, R. B. (2017). Maximum a-posteriori estimation of autoregressive processes based on finite mixtures of scale-mixtures of skew-normal distributions. *Journal of Statistical Computation and Simulation*, *87*(6), 1061–1083.

56. Maleki, M., Arellano-Valle, R. B., Dey, D. K., Mahmoudi, M. R., & Jalali, S. M. J. (2017). A Bayesian approach to robust skewed autoregressive processes. *Calcutta Statistical Association Bulletin*, *69*(2), 165–182.

57. Maleki, M., Mahmoudi, M. R., Wraith, D., & Pho, K.-H. (2020). Time series modelling to forecast the confirmed and recovered cases of COVID-19. *Travel Medicine and Infectious Disease*, *37*, 101742.

58. Maleki, M., & Nematollahi, A. (2017). Autoregressive models with mixture of scale mixtures of gaussian innovations. *Iranian Journal of Science and Technology, Transactions A: Science*, *41*(4), 1099–1107.

59. Mandal, M., Jana, S., Nandi, S. K., Khatua, A., Adak, S., & Kar, T. (2020). A model based study on the dynamics of covid-19: Prediction and control. *Chaos, Solitons & Fractals*, *136*, 109889.

60. Massonis, G., Banga, J. R., & Villaverde, A. F. (2020). Structural identifiability and observability of compartmental models of the covid-19 pandemic. *Annual Reviews in Control*. Volume 51, 2021, Pages 441–459

61. Mbuvha, R. R., & Marwala, T. (2020). On data-driven management of the Covid-19 outbreak in South Africa. *medRxiv*, *2020*.

62. Meng, Z., Wang, M., Song, H., Guo, S., Zhou, Y., Li, W., . . . Ying, B. (2020). Development and utilization of an intelligent application for aiding COVID-19 diagnosis. *medRxiv* (37), Volume 2020, Pages 1–21.

63. Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., & Soufi, G. J. (2020). Deep-covid: Predicting covid-19 from chest X-ray images using deep transfer learning. *Medical Image Analysis*, *65*, 101794.

64. Moftakhar, L., Mozhgan, S., & Safe, M. S. (2020). Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. *Iranian Journal of Public Health*, *2020*.

65. Nabi, K. N., Abboubakar, H., & Kumar, P. (2020). Forecasting of covid-19 pandemic: from integer derivatives to fractional derivatives. *Chaos, Solitons & Fractals*, *141*, 110283.

66. Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic detection of coronavirus disease (covid-19) using X-ray images and deep convolutional neural networks. Preprint. arXiv:2003.10849.

67. Ndaïrou, F., Area, I., Nieto, J. J., & Torres, D. F. (2020). Mathematical modeling of covid-19 transmission dynamics with a case study of Wuhan. *Chaos, Solitons & Fractals*, *135*, 109846.

68. Pal, R., Sekh, A. A., Kar, S., & Prasad, D. K. (2020). Neural network based country wise risk prediction of COVID-19. *Applied Sciences*, *10*(18), 6448.

69. Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. Preprint. arXiv:2002.06563, *2020*.

70. Pourghasemi, H. R., Pouyan, S., Heidari, B., Farajzadeh, Z., Fallah Shamsi, S. R., Babaei, S., . . . Sadeghian, F. (2020). Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *International Journal of Infectious Diseases*, *98*, 90–108. Retrieved from https://doi.org/10.1016/j.ijid.2020.06.058

71. Putra, S., & Khozin Mu'tamar, Z. (2019). Estimation of parameters in the SIR epidemic model using particle swarm optimization. *American Journal of Mathematical and Computer Modelling*, *4*(4), 83–93.

72. Qeadan, F., Honda, T., Gren, L. H., Dailey-Provost, J., Benson, L. S., VanDerslice, J. A., . . . Shoaf, K. (2020). Naive forecast for COVID-19 in Utah based on the South Korea and Italy models-the fluctuation between two extremes. *International Journal of Environmental Research and Public Health*, *17*(8), 2750.

73. Qi, H., Xiao, S., Shi, R., Ward, M. P., Chen, Y., Tu, W., . . . Zhang, Z. (2020). COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis. *Science of the Total Environment*, *728*, 138778.

74. Rahimi, I., Chen, F., & Gandomi, A. H. (2021). A review on COVID-19 forecasting models. *Neural Computing and Applications*, *2020*, 1–11.

75. Rajagopal, K., Hasanzadeh, N., Parastesh, F., Hamarash, I. I., Jafari, S., & Hussain, I. (2020). A fractional-order model for the novel coronavirus (covid-19) outbreak. *Nonlinear Dynamics*, *101*(1), 711–718.

76. Ren, H., Zhao, L., Zhang, A., Song, L., Liao, Y., Lu, W., & Cui, C. (2020). Early forecasting of the potential risk zones of COVID-19 in China's megacities. *Science of the Total Environment*, *729*, 138995. Retrieved from https://doi.org/10.1016/j.scitotenv.2020.138995

77. Ribeiro, M. H. D. M., da Silva, R. G. da, Mariani, V. C., & dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals*, *135*, 109853.

78. Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the covid-19 epidemic? *Infectious Disease Modelling*, *5*, 271–281.

79. Roy, S., Bhunia, G. S., & Shit, P. K. (2020a). Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Modeling Earth Systems and Environment*, *2019*(0123456789). Retrieved from https://doi.org/10.1007/s40808-020-00890-y

80. Roy, S., Bhunia, G. S., & Shit, P. K. (2020b). Spatial prediction of covid-19 epidemic using arima techniques in India. *Modeling Earth Systems and Environment*, *2020*, 1–7.

81. Sadun, L. (2020). Effects of latency on estimates of the covid-19 replication number. *Bulletin of Mathematical Biology*, *82*(9), 1–14.

82. 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.

83. 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.

84. Samui, P., Mondal, J., & Khajanchi, S. (2020). A mathematical model for covid-19 transmission dynamics with a case study of India. *Chaos, Solitons & Fractals*, *140*, 110173.

85. Sarkar, K., Khajanchi, S., & Nieto, J. J. (2020). Modeling and forecasting the covid-19 pandemic in India. *Chaos, Solitons & Fractals*, *139*, 110049.

86. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117.

87. Shao, N., Zhong, M., Yan, Y., Pan, H., Cheng, J., & Chen, W. (2020). Dynamic models for coronavirus disease 2019 and data analysis. *Mathematical Methods in the Applied Sciences*, *43*(7), 4943–4949.

88. Singh, S., Parmar, K. S., Kumar, J., & Makkhan, S. J. S. (2020). Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos, Solitons & Fractals*, *135*, 109866.

89. Suba, M., Shanmugapriya, R., Balamuralitharan, S., & Joseph, G. A. (n.d.). Current mathematical models and numerical simulation of sir model for coronavirus disease-2019 (covid-19). *European Journal of Molecular & Clinical Medicine*, *7*(05), 2020.

90. Sujath, R., Chatterjee, J. M., & Hassanien, A. E. (2020). A machine learning forecasting model for COVID-19 pandemic in India. *Stochastic Environmental Research and Risk Assessment*, *34*, 959–972.

91. Tamang, S., Singh, P., & Datta, B. (2020). Forecasting of Covid-19 cases based on prediction using artificial neural network curve fitting technique. *Global Journal of Environmental Science and Management*, *6*(Special Issue (Covid-19)), 53–64.

92. Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of the Total Environment*, *728*, 138762. Retrieved from https://doi.org/10.1016/j.scitotenv.2020.138762

93. Varotsos, C. A., & Krapivin, V. F. (2020). A new model for the spread of covid-19 and the improvement of safety. *Safety Science*, *132*, 104962.

94. 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*, *136*, 109924.

95. Viguerie, A., Veneziani, A., Lorenzo, G., Baroli, D., Aretz-Nellesen, N., Patton, A., . . . Auricchio, F. (2020). Diffusion–reaction compartmental models formulated in a continuum mechanics framework: application to covid-19, mathematical analysis, and numerical study. *Computational Mechanics*, *66*(5), 1131–1152.

96. Wang, L., Lin, Z. Q., & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. *Scientific Reports*, *10*(1), 1–12.

97. Wang, L., Wang, G., Gao, L., Li, X., Yu, S., Kim, M., . . . Gu, Z. (2020). Spatiotemporal dynamics, nowcasting and forecasting of COVID-19 in the United States. ArXiv, 1–26. Retrieved from http://arxiv.org/abs/2004.14103

98. WHO. (2021). WHO Coronavirus (COVID-19) Dashboard [Computer software manual]. Retrieved from https://covid19.who.int/. Last accessed: 06 April 2021.

99. Wieczorek, M., Siłka, J., & Woźniak, M. (2020). Neural network powered COVID-19 spread forecasting model. *Chaos, Solitons & Fractals*, *140*, 110203.

100. Xie, J., Hungerford, D., Chen, H., Abrams, S. T., Li, S., Wang, G., . . . Toh, C.-H. (2020). Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. *The Lancet*, 2020, 1–29.

101. Yang, Z., Zeng, Z., Wang, K., Wong, S.-S., Liang, W., Zanin, M., . . . 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.

102. Yesilkanat, C. M. (2020). Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. *Chaos, Solitons and Fractals*, *140*, 110210.
103. Zarrin, P., Maleki, M., Khodadai, Z., & Arellano-Valle, R. B. (2019). Time series models based on the unrestricted skew-normal process. *Journal of Statistical Computation and Simulation*, *89*(1), 38–51.
104. Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, *140*, 110121.
105. Zhong, L., Mu, L., Li, J., Wang, J., Yin, Z., & Liu, D. (2020). Early prediction of the 2019 novel coronavirus outbreak in the mainland China based on simple mathematical model. *IEEE Access*, *8*, 51761–51769.