COVID-19 Outbreak Prediction with Machine Learning

Sina F. Ardabili 1, Amir Mosavi 23*, Pedram Ghamisi 4, Filip Ferdinand 2, Annamaria R. Varkonyi-Koczy 2, Uwe Reuter 3, Timon Rabczuk 5, Peter M. Atkinson 6

1 Thuringian Institute of Sustainability and Climate Protection, 07743 Jena, Germany; s.ardabili@ieee.org
2 Department of Mathematics and Informatics, J. Selye University, 94501 Komarno, Slovakia; filipf@ujs.sk
3 Faculty of Civil Engineering, Technische Universität Dresden, 01069 Dresden, Germany.
4 Machine Learning Group, Exploration Division, Helmholtz Institute Freiberg for Resource Technology, Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany; p.ghamisi@hzdr.de
5 Institute of Structural Mechanics (ISM), Bauhaus-Universität Weimar, 99423 Weimar, Germany
6 Institute of Structural Mechanics (ISM), Bauhaus-Universität Weimar, 99423 Weimar, Germany
* Correspondence: a.mosavi@brookes.ac.uk

Received: date; Accepted: date; Published: date

Abstract: Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Among the standard models for COVID-19 global pandemic prediction, simple epidemiological and statistical models have received more attention by authorities, and they are popular in the media. Due to a high level of uncertainty and lack of essential data, standard models have shown low accuracy for long-term prediction. Although the literature includes several attempts to address this issue, the essential generalization and robustness abilities of existing models needs to be improved. This paper presents a comparative analysis of machine learning and soft computing models to predict the COVID-19 outbreak. Among a wide range of machine learning models investigated, two models showed promising results (i.e., multi-layered perceptron, MLP, and adaptive network-based fuzzy inference system, ANFIS). Based on the results reported here, and due to the highly complex nature of the COVID-19 outbreak and variation in its behavior from nation-to-nation, this study suggests machine learning as an effective tool to model the outbreak.

Keywords: COVID-19; Coronavirus disease; Coronavirus; SARS-CoV-2; model; prediction; machine learning

1. Introduction

Access to accurate outbreak prediction models is essential to obtain insights into the likely spread and consequences of infectious diseases. Governments and other legislative bodies rely on insights from prediction models to suggest new policies and to assess the effectiveness of the enforced policies [1]. The novel Coronavirus disease (COVID-19) has been reported to infect more than 2 million people, with more than 132,000 confirmed deaths worldwide. The recent global COVID-19 pandemic has exhibited a nonlinear and complex nature [2]. In addition, the outbreak has differences with other recent outbreaks, which brings into question the ability of standard models to deliver accurate results [3]. Besides the numerous known and unknown variables involved in the spread, the complexity of population-scale behavior in various geopolitical areas and differences in containment strategies had dramatically increased model uncertainty [4]. Consequently, standard epidemiological models face new challenges to deliver more reliable results. To overcome this challenge, many novel models have emerged which introduce several assumptions to modeling (e.g., adding social distancing in the form of curfews, quarantines, etc.) [5-7].

To elaborate on the effectiveness of enforcing such assumptions understanding standard dynamic epidemiological (e.g., susceptible-infected-recovered, SIR) models is essential [8]. The modeling strategy is formed around the assumption of transmitting the infectious disease through contacts, considering three different classes of well-mixed populations; susceptible to infection (class
\( S \), infected (class \( I \)), and the removed population (class \( R \) is devoted to those who have recovered, developed immunity, been isolated or passed away). It is further assumed that the class \( I \) transmits the infection to class \( S \) where the number of probable transmissions is proportional to the total number of contacts [9-11]. The number of individuals in the class \( S \) progresses as a time-series, often computed using a basic differential equation as follows:

\[
\frac{dS}{dt} = -\alpha SI
\]

where \( I \) is the infected population, and \( S \) is the susceptible population both as fractions. \( \alpha \) represents the daily reproduction rate of the differential equation, regulating the number of susceptible infectious contacts. The value of \( S \) in the time-series produced by the differential equation gradually declines. Initially, it is assumed that at the early stage of the outbreak \( S \approx 1 \) while the number of individuals in class \( I \) is negligible. Thus, the increment \( \frac{dI}{dt} \) becomes linear and the class \( I \) eventually can be computed as follows:

\[
\frac{dI}{dt} = \alpha SI - \beta I
\]

where \( \beta \) regulates the daily rate of new infections by quantifying the number of infected individuals competent in the transmission. Furthermore, the class \( R \), representing individuals excluded from the spread of infection, is computed as follows:

\[
\frac{dR}{dt} = \beta I
\]

Under the unconstrained conditions of the excluded group, Eq. 3, the outbreak exponential growth can be computed as follows:

\[
I(t) \approx I_0 \exp((\alpha - \beta))
\]

The outbreaks of a wide range of infectious diseases have been modeled using Eq. 4. However, for the COVID-19 outbreak prediction, due to the strict measures enforced by authorities, the susceptibility to infection has been manipulated dramatically. For example, in China, Italy, France, Hungary and Spain the SIR model cannot present promising results, as individuals committed voluntarily to quarantine and limited their social interaction. However, for countries where containment measures were delayed (e.g., United States) the model has shown relative accuracy [12]. Figure 1 shows the inaccuracy of conventional models applied to the outbreak in Italy by comparing the actual number of confirmed infections and epidemiological model predictions. The SEIR models through considering the significant incubation period during which individuals have been infected showed progress in improving the model accuracy for Varicella and Zika outbreak [13,14]. SEIR models assume that the incubation period is a random variable and similarly to the SIR model, there would be a disease-free-equilibrium [15,16]. It is worth mentioning that SEIR model will not work well where the parameters are non-stationary through time [17]. A key cause of non-stationarity is where the social mixing (which determines the contact network) changes through time. Social mixing determines the reproductive number \( R_0 \) which is the number of susceptible individuals that an infected person will infect. Where \( R_0 \) is less than 1 the epidemic will die out. Where it is greater than 1 it will spread. \( R_0 \) for COVID-19 prior to lockdown was estimated as a massive 4 presenting a pandemic. It is expected that lockdown measures should bring \( R_0 \) down to less than 1. The key reason why SEIR models are difficult to fit for COVID-19 is non-stationarity of mixing, caused by nudging (step-by-step) intervention measures.

One can calculate that standard epidemiological models can be effective and reliable only if (a) the social interactions are stationary through time (i.e., no changes in interventions or control measures), or (b) there exists a great deal of knowledge of class \( R \) with which to compute Eq. 3. Often to acquire information on class \( R \), several novel models included data from social media or call data records (CDR), which showed promising results [18-25]. However, observation of the behavior of
COVID-19 in several countries demonstrates a high degree of uncertainty and complexity [26]. Thus, for epidemiological models to be able to deliver reliable results, they must be adapted to the local situation with an insight into susceptibility to infection [27]. This imposes a huge limit on the generalization ability and robustness of conventional models. Advancing accurate models with a great generalization ability to be scalable to model both the regional and global pandemic is, thus, essential [28].

A further drawback of conventional epidemiological models is the short lead-time. To evaluate the performance of the models, the median success of the outbreak prediction presents useful information. The median prediction factor can be calculated as follows:

$$f = \frac{Prediction}{True \ value}$$ (5)

As the lead-time increases, the accuracy of the model declines. For instance, for the COVID-19 outbreak in Italy, the accuracy of the model for more than 5-days-in-the-future reduces from $f = 1$ for the first five days to $f = 0.86$ for day 6 [12].

![Figure 1. Italy’s COVID-19 outbreak: the actual number of confirmed infections vs. epidemiological model.](image)

Due to the complexity and the large-scale nature of the problem in developing epidemiological models, machine learning (ML) has recently gained attention for building outbreak prediction models. ML approaches aim at developing models with higher generalization ability and greater prediction reliability for longer lead-times [29-33].

Although ML methods were used in modeling former pandemics (e.g., Ebola, Cholera, swine fever, H1N1 influenza, dengue fever, Zika, oyster norovirus [8,34-43]), there is a gap in the literature for peer-reviewed papers dedicated to COVID-19. Table 1 represents notable ML methods used for outbreak prediction. These ML methods are limited to the basic methods of random forest, neural networks, Bayesian networks, Naïve Bayes, genetic programming and classification and regression tree (CART). Although ML has long been established as a standard tool for modeling natural disasters and weather forecasting [44,45], its application in modeling outbreak is still in the early stages. More sophisticated ML methods (e.g., hybrids, ensembles) are yet to be explored. Consequently, the contribution of this paper is to explore the application of ML for modeling the COVID-19 pandemic. This paper aims to investigate the generalization ability of the proposed ML models and the accuracy of the proposed models for different lead-times.
The rest of this paper is organized as follows. Section two describes the methods and materials. The results are given in section three. Sections four and five present the discussion and the conclusions, respectively.

2. Materials and Methods

Data were collected from https://www.worldometers.info/coronavirus/country for five countries, including Italy, Germany, Iran, USA, and China on total cases over 30 days. Figure 2 presents the total case number (cumulative statistic) for the considered countries. Currently, to contain the outbreak, the governments have implemented various measures to reduce transmission through inhibiting people’s movements and social activities. Although for advancing the epidemiological models information on changes in social distancing is essential, for modeling with machine learning no assumption is required. As can be seen in Figure 2, the growth rate in China is greater than that for Italy, Iran, Germany and the USA in the early weeks of the disease.

![Figure 2. Cumulative number of cases for five countries during thirty days](https://www.worldometers.info/coronavirus/country)
The next step is to find the best model for the estimation of the time-series data. Logistic, Linear, Logarithmic, Quadratic, Cubic, Compound, Power and exponential equations (Table 2) are employed to develop the desired model.

Table 2. Models for mathematical forecasting

<table>
<thead>
<tr>
<th>Model description</th>
<th>Model name</th>
<th>Equation number</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R = \frac{A}{1 + \exp((4\mu)((L-x)/A)+2)} )</td>
<td>Logistic</td>
<td>(6)</td>
</tr>
<tr>
<td>( R = Ax-B )</td>
<td>Linear</td>
<td>(7)</td>
</tr>
<tr>
<td>( R = A + B \log(x) )</td>
<td>Logarithmic</td>
<td>(8)</td>
</tr>
<tr>
<td>( R = A + Bx + Cx^2 )</td>
<td>Quadratic</td>
<td>(9)</td>
</tr>
<tr>
<td>( R = A + Bx + Cx^2 + Dx^3 )</td>
<td>Cubic</td>
<td>(10)</td>
</tr>
<tr>
<td>( R = ABx )</td>
<td>Compound</td>
<td>(11)</td>
</tr>
<tr>
<td>( R = Ax^b )</td>
<td>Power</td>
<td>(12)</td>
</tr>
<tr>
<td>( R = A \exp(Bx) )</td>
<td>Exponential</td>
<td>(13)</td>
</tr>
</tbody>
</table>

A, B, C, \( \mu \), and \( L \) are parameters (constants) that characterize the above-mentioned functions. These constants need to be estimated to develop an accurate estimation model. One of the goals of this study was to model time-series data based on the logistic microbial growth model. For this purpose, the modified equation of logistic regression was used to estimate and predict the prevalence (i.e., \( I/P \) population at a given time point) of disease as a function of time. Estimation of the parameters was performed using evolutionary algorithms like GA, particle swarm optimizer, and the grey wolf optimizer. These algorithms are discussed in the following.

Evolutionary algorithms

Evolutionary algorithms (EA) are powerful tools for solving optimization problems through intelligent methods. These algorithms are often inspired by natural processes to search for all possible answers as an optimization problem [46-48]. In the present study, the frequently used algorithms, (i.e., genetic algorithm (GA), particle swarm optimizer (PSO) and grey wolf optimizer (GWO)) are employed to estimate the parameters by solving a cost function.

Genetic Algorithm (GA)

GAs are considered a subset of "computational models" inspired by the concept of evolution [49]. These algorithms use "Potential Solutions" or "Candidate Solutions" or "Possible Hypotheses" for a specific problem in a "chromosome-like" data structure. GA maintains vital information stored in these chromosome data structures by applying "Recombination Operators" to chromosome-like data structures [50-53]. In many cases, GAs are employed as "Function Optimizer" algorithms, which are algorithms used to optimize "Objective Functions." Of course, the range of applications that use the GA to solve problems is very wide [52,54]. The implementation of the GA usually begins with the production of a population of chromosomes generated randomly and bound up and down by the variables of the problem. In the next step, the generated data structures (chromosomes) are evaluated, and chromosomes that can better display the optimal solution of the problem are more likely to be used to produce new chromosomes. The degree of "goodness" of an answer is usually measured by the population of the current candidate's answers [55-59]. The main algorithm of a GA process is demonstrated in Figure 3.
Problem

Modeling of the problem

Formation of the initial population

Evaluation of the Population

Choosing the parents

Offspring

New population

Replacing the new population size

Selection

Response

Figure 3. GA algorithm

In the present study, GA [59] was employed for estimation of the parameters of Eq. 6 to 13. The population number was selected to be 300 and the maximum generation (as iteration number) was determined to be 500 according to different trial and error processes to reduce the cost function value. The cost function was defined as the mean square error between the target and estimated values according to Eq. 14:

\[ MSE = \sqrt{\frac{(Es-T)^2}{N}} \]  

(14)

where, \( Es \) refer to estimated values, \( T \) refers to the target values and \( N \) refers to the number of data.

Particle Swarm Optimization (PSO)

In 1995, Kennedy and Eberhart [60] introduced the PSO as an uncertain search method for optimization purposes. The algorithm was inspired by the mass movement of birds looking for food. A group of birds accidentally looked for food in a space. There is only one piece of food in the search space. Each solution in PSO is called a particle, which is equivalent to a bird in the bird’s mass movement algorithm. Each particle has a value that is calculated by a competency function which increases as the particle in the search space approaches the target (food in the bird’s movement model). Each particle also has a velocity that guides the motion of the particle. Each particle continues to move in the problem space by tracking the optimal particles in the current state [60-62]. The PSO method is rooted in Reynolds’ work, which is an early simulation of the social behavior of birds. The mass of particles in nature represents collective intelligence. Consider the collective movement of fish in water or birds during migration. All members move in perfect harmony with each other, hunt together if they are to be hunted, and escape from the clutches of a predator by moving another prey if they are to be preyed upon [63-65]. Particle properties in this algorithm include [65-67]:

- Each particle independently looks for the optimal point.
- Each particle moves at the same speed at each step.
- Each particle remembers its best position in the space.
- The particles work together to inform each other of the places they are looking for.
- Each particle is in contact with its neighboring particles.
- Every particle is aware of the particles that are in the neighborhood.
- Every particle is known as one of the best particles in its neighborhood.
The PSO implementation steps can be summarized as: the first step establishes and evaluates the primary population. The second step determines the best personal memories and the best collective memories. The third step updates the speed and position. If the conditions for stopping are not met, the cycle will go to the second step.

The PSO algorithm is a population-based algorithm [68,69]. This property makes it less likely to be trapped in a local minimum. This algorithm operates according to possible rules, not definite rules. Therefore, PSO is a random optimization algorithm that can search for unspecified and complex areas. This makes PSO more flexible and durable than conventional methods. PSO deals with non-differential target functions because the PSO uses the information result (performance index or target function to guide the search in the problem area). The quality of the proposed route response does not depend on the initial population. Starting from anywhere in the search space, the algorithm ultimately converges on the optimal answer. PSO has great flexibility to control the balance between the local and overall search space. This unique PSO property overcomes the problem of improper convergence and increases the search capacity. All of these features make PSO different from the GA and other innovative algorithms [61,65,67].

In the present study, PSO was employed for estimation of the parameters of Eq. 6 to 13. The population number was selected to be 1000 and the iteration number was determined to be 500 according to different trial and error processes to reduce the cost function value. The cost function was defined as the mean square error between the target and estimated values according to Eq. 14.

**Grey Wolf Optimizer (GWO)**

One recently developed smart optimization algorithm that has attracted the attention of many researchers is the grey wolf algorithm. Like most other intelligent algorithms, GWO is inspired by nature. The main idea of the grey wolf algorithm is based on the leadership hierarchy in wolf groups and how they hunt [70]. In general, there are four categories of wolves among the herd of grey wolves, alpha, beta, delta and omega. Alpha wolves are at the top of the herd’s leadership pyramid, and the rest of the wolves take orders from the alpha group and follow them (usually there is only one wolf as an alpha wolf in each herd). Beta wolves are in the lower tier, but their superiority over Delta and omega wolves allows them to provide advice and help to alpha wolves. Beta wolves are responsible for regulating and orienting the herd based on alpha movement. Delta wolves, which are next in line for the power pyramid in the wolf herd, are usually made up of guards, elderly population, caregivers of damaged wolves, and so on. Omega wolves are also the weakest in the power hierarchy [70]. Eq. 15 to 18 are used to model the hunting tool:

\[
\vec{D} = |\vec{C}, \vec{X}_p(t) - \vec{X}(t)| \quad (15)
\]

\[
\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}, \vec{D} \quad (16)
\]

\[
\vec{A} = 2\vec{a}, \vec{r}_1 - \vec{a} \quad (17)
\]

\[
\vec{C} = 2\vec{r}_2 \quad (18)
\]

where \(t\) is represents repetition of the algorithm. \(\vec{A}\) and \(\vec{C}\) are vectors of the prey site and the \(\vec{X}\) vectors represent the locations of the grey wolves. \(\vec{a}\) is linearly reduced from 2 to 0 during the repetition. \(\vec{r}_1\) and \(\vec{r}_2\) are random vectors where each element can take on realizations in the range [0,1]. The GWO algorithm flowchart is shown in Figure 4.
In the present study, GWO [70] was employed for estimation of the parameters of Eq.1 to 8. The population number was selected to be 500 and the iteration number was determined to be 1000 according to different trial and error processes to reduce the cost function value. The cost function was defined as the mean square error between the target and estimated values according to Eq. 14.

**Machine learning (ML)**

ML is regarded as a subset of AI. Using ML techniques, the computer learns to use patterns or “training samples” in data (processed information) to predict or make intelligent decisions without overt planning [71,72]. In other words, ML is the scientific study of algorithms and statistical models used by computer systems that use patterns and inference to perform tasks instead of using explicit instructions [73,74].

Time-series are data sequences collected over a period of time [75], which can be used as inputs to ML algorithms. This type of data reflects the changes that a phenomenon has undergone over time. Let \( X_t \) be a time-series vector, in which \( x_t \) is the outbreak at time point \( t \) and \( T \) is the set of all equidistant time points. To train ML methods effectively, we defined two scenarios, listed in Table 3.

**Table 3. Input and output variables for training ML methods by time-series data**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Inputs</th>
<th>Input number</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>( X_{t-1}, X_{t-7}, X_{t-14}, \text{and } X_{t-21} )</td>
<td>Four inputs</td>
<td>( x_t ) (outbreak)</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>( X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, \text{and } X_{t-5} )</td>
<td>Five inputs</td>
<td>( x_t ) (outbreak)</td>
</tr>
</tbody>
</table>
As can be seen in Table 3, scenario 1 employs data for three weeks to predict the outbreak on day $t$ and scenario 2 employs outbreak data for five days to predict the outbreak for day $t$. Both of these scenarios were employed for fitting the ML methods. In the present research, two frequently used ML methods, the multi-layered perceptron (MLP) and adaptive network-based fuzzy inference system (ANFIS) are employed for the prediction of the outbreak in the five countries.

Multi-layered perceptron (MLP)

ANN is an idea inspired by the biological nervous system, which processes information like the brain. The key element of this idea is the new structure of the information processing system [76-78]. The system is made up of several highly interconnected processing elements called neurons that work together to solve a problem [78,79]. ANNs, like humans, learn by example. The neural network is set up during a learning process to perform specific tasks, such as identifying patterns and categorizing information. In biological systems, learning is regulated by the synaptic connections between nerves. This method is also used in neural networks [80]. By processing experimental data, ANNs transfer knowledge or a law behind the data to the network structure, which is called learning. Basically, learning ability is the most important feature of such a smart system. A learning system is more flexible and easier to plan, so it can better respond to new issues and changes in processes [81].

In ANNs, with the help of programming knowledge, a data structure is designed that can act like a neuron. This data structure is called a node [82,83]. In this structure, the network between these nodes is trained by applying an educational algorithm to it. In this memory or neural network, the nodes have two active states (on or off) and one inactive state (off or 0), and each edge (synapse or connection between nodes) has a weight. Positive weights stimulate or activate the next inactive node, and negative weights inactivate or inhibit the next connected node (if active) [78,84]. In the ANN architecture, for the neural cell $c$, the input $b_y$ enters the cell from the previous cell $p$. $w_{pc}$ is the weight of the input $b_y$ with respect to cell $c$ and $a_c$ is the sum of the multiplications of the inputs and their weights [85]:

$$a_c = \sum w_{pc} b_{pc}$$

A non-linear function $\Theta_c$ is applied to $a_c$. Accordingly, $b_c$ can be calculated as Eq. 20 [85]:

$$b_c = \Theta_c(a_c)$$

Similarly, $w_{cn}$ is the weight of the $b_{on}$ which is the output of $c$ to $n$. $W$ is the collection of all the weights of the neural network in a set. For input $x$ and output $y$, $h_o(x)$ is the output of the neural network. The main goal is to learn these weights for reducing the error values between $y$ and $h_o(x)$. That is, the goal is to minimize the cost function $Q(W)$, Eq. 21 [85]:

$$Q(W) = \frac{1}{2} \sum_{i=1}^{n} (y_i - o_i)^2$$

In the present research, one of the frequently used types of ANN called the MLP [76] was employed to predict the outbreak. MLP was trained using a dataset related to both scenarios (according to Table 2). For the training of the network, 8, 12, and 16 inner neurons were tried to achieve the best response. Results were evaluated by RMSE and correlation coefficient to reduce the cost function value. Figure 5 presents the architecture of the MLP.
Adaptive neuro fuzzy inference system (ANFIS)

An adaptive neuro fuzzy inference system is a type of ANN based on the Takagi-Sugeno fuzzy system [86]. This approach was developed in the early 1990s. Since this system integrates the concepts of neural networks and fuzzy logic, it can take advantage of both capabilities in a unified framework. This technique is one of the most frequently used and robust hybrid ML techniques. It is consistent with a set of fuzzy if-then rules that can be learned to approximate nonlinear functions [87,88]. Hence, ANFIS was proposed as a universal estimator. An important element of fuzzy systems is the fuzzy partition of the input space [89,90]. For input $k$, the fuzzy rules in the input space make a $k$ faces fuzzy cube. Achieving a flexible partition for nonlinear inversion is non-trivial. The idea of this model is to build a neural network whose outputs are a degree of the input that belongs to each class [91-93]. The membership functions (MFs) of this model can be nonlinear, multidimensional and, thus, different to conventional fuzzy systems [94-96]. In ANFIS, neural networks are used to increase the efficiency of fuzzy systems. The method used to design neural networks is to employ fuzzy systems or fuzzy-based structures. This model is a kind of division and conquest method. Instead of using one neural network for all the input and output data, several networks are created in this model:

- A fuzzy separator to cluster input-output data within multiple classes.
- A neural network for each class.
- Training neural networks with output input data in the corresponding classes.

Figure 6 presents a simple architecture for ANFIS.
In the present study, ANFIS is developed to tackle two scenarios described in table 3. Each input included by two MFs with the Tri. shape; Trap. shape and Gauss. shape MFs. The output MF type was selected to be linear with a hybrid optimiser type.

Evaluation criteria

Evaluation was conducted using the root mean square error (RMSE) and correlation coefficient. These statistics compare the target and output values and calculate a score as an index for the performance and accuracy of the developed methods [87,97]. Table 4 presents the evaluation criteria equations.

\[ \text{Correlation coefficient} = \frac{N \sum (AP - \bar{A} \bar{P})}{\sqrt{[N \sum A^2 - (\sum A)^2][N \sum P^2 - (\sum P)^2]}} \]

\[ \text{RMSE} = \sqrt{\frac{1}{N} \sum (A - P)^2} \]

Where, \( N \) is the number of data, \( P \) and \( A \) are, respectively, the predicted (output) and desired (target) values.

3. Results

Tables 5 to 12 present the results of the accuracy statistics for the logistic, linear, logarithmic, quadratic, cubic, compound, power and exponential equations, respectively. The coefficients of each equation were calculated by the three ML optimizers; GA, PSO and GWO. The table contains country name, model name, population size, number of iterations, processing time, RMSE and correlation coefficient.

<table>
<thead>
<tr>
<th>Country</th>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>82 s</td>
<td>1028.98</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>36 s</td>
<td>3358.1</td>
<td>0.997</td>
</tr>
<tr>
<td>Country</td>
<td>Model</td>
<td>Pop. size</td>
<td>iteration</td>
<td>Processing time</td>
<td>RMSE</td>
<td>Correlation coefficient</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------------------</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>92 s</td>
<td>7774.06</td>
<td>0.845</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>42 s</td>
<td>3645.76</td>
<td>0.844</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>16 s</td>
<td>3642.44</td>
<td>0.844</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>91 s</td>
<td>7188.95</td>
<td>0.981</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>39 s</td>
<td>6644.16</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>14 s</td>
<td>5039.48</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>96 s</td>
<td>3330.45</td>
<td>0.943</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>45 s</td>
<td>2072.71</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>18 s</td>
<td>1981.97</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>88 s</td>
<td>850.22</td>
<td>0.745</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>40 s</td>
<td>596.69</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>17 s</td>
<td>592.48</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>93 s</td>
<td>1118.77</td>
<td>0.758</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>47 s</td>
<td>964.46</td>
<td>0.759</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>20 s</td>
<td>951.63</td>
<td>0.759</td>
</tr>
</tbody>
</table>

**Table 6. Accuracy statistics for the linear model**

<table>
<thead>
<tr>
<th>Country</th>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>98 s</td>
<td>8325.33</td>
<td>0.634</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>51 s</td>
<td>8818.2</td>
<td>0.634</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>20 s</td>
<td>9296.59</td>
<td>0.634</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>96 s</td>
<td>40828.2</td>
<td>0.847</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>42 s</td>
<td>43835.37</td>
<td>0.847</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>17 s</td>
<td>42714.93</td>
<td>0.847</td>
</tr>
<tr>
<td>Iran</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>102 s</td>
<td>4929.97</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>59 s</td>
<td>8775.56</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>22 s</td>
<td>8995.52</td>
<td>0.756</td>
</tr>
<tr>
<td>USA</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>94 s</td>
<td>889.15</td>
<td>0.538</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>37 s</td>
<td>1130.33</td>
<td>0.538</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>15 s</td>
<td>1135.12</td>
<td>0.538</td>
</tr>
<tr>
<td>Germany</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>95 s</td>
<td>1552.22</td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>47 s</td>
<td>964.46</td>
<td>0.759</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>20 s</td>
<td>951.63</td>
<td>0.759</td>
</tr>
</tbody>
</table>

**Table 7. Accuracy statistics for the logarithmic model**
Table 8. Accuracy statistics for the quadratic model

<table>
<thead>
<tr>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>102 s</td>
<td>6710.01</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>54 s</td>
<td>5102.4</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>26 s</td>
<td>1272.1</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>100 s</td>
<td>7921.33</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>46 s</td>
<td>4328.71</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>20 s</td>
<td>3710.16</td>
</tr>
<tr>
<td>Iran</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>105 s</td>
<td>6771.74</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>62 s</td>
<td>822.09</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>24 s</td>
<td>310.02</td>
</tr>
<tr>
<td>USA</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>98 s</td>
<td>754.6</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>38 s</td>
<td>791.92</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>19 s</td>
<td>307.58</td>
</tr>
<tr>
<td>Germany</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>101 s</td>
<td>7577</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>49 s</td>
<td>752.95</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>26 s</td>
<td>472.62</td>
</tr>
</tbody>
</table>

Table 9. Accuracy statistics for the cubic model

<table>
<thead>
<tr>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>112 s</td>
<td>7973.11</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>61 s</td>
<td>4827.08</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>34 s</td>
<td>324.33</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>113 s</td>
<td>15697.84</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>59 s</td>
<td>3611.15</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>34 s</td>
<td>2429.45</td>
</tr>
<tr>
<td>Iran</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>120 s</td>
<td>5852.66</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>88 s</td>
<td>3809.76</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>39 s</td>
<td>250.2</td>
</tr>
<tr>
<td>USA</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>110 s</td>
<td>37766.56</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>49 s</td>
<td>678.36</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>25 s</td>
<td>118.24</td>
</tr>
<tr>
<td>Germany</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>116 s</td>
<td>1709.06</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>59 s</td>
<td>1812.78</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>29 s</td>
<td>196.8</td>
</tr>
</tbody>
</table>

Table 10. Accuracy statistics for the compound model

<table>
<thead>
<tr>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>92 s</td>
<td>8347.51</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>53 s</td>
<td>195705.52</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>22 s</td>
<td>12585.79</td>
</tr>
</tbody>
</table>
### Table 1. Accuracy statistics for the power model

<table>
<thead>
<tr>
<th>Country</th>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>72 s</td>
<td>7063.4</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>40 s</td>
<td>6150.52</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>13 s</td>
<td>3450.96</td>
<td>0.991</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>65 s</td>
<td>39669.92</td>
<td>0.976</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>39 s</td>
<td>19365.58</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>12 s</td>
<td>4078.99</td>
<td>0.989</td>
</tr>
<tr>
<td>Iran</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>83 s</td>
<td>2343032.5</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>65 s</td>
<td>92755.53</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>15 s</td>
<td>1031.6</td>
<td>0.991</td>
</tr>
<tr>
<td>USA</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>79 s</td>
<td>1030.01</td>
<td>0.779</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>24 s</td>
<td>1005.27</td>
<td>0.751</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>11 s</td>
<td>790.16</td>
<td>0.837</td>
</tr>
<tr>
<td>Germany</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>85 s</td>
<td>1475.39</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>69 s</td>
<td>1387.94</td>
<td>0.916</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>14 s</td>
<td>1341.91</td>
<td>0.875</td>
</tr>
</tbody>
</table>

### Table 12. Accuracy statistics for the exponential model

<table>
<thead>
<tr>
<th>Country</th>
<th>Model</th>
<th>Pop. size</th>
<th>iteration</th>
<th>Processing time</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>79 s</td>
<td>8163.1</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>48 s</td>
<td>52075925.37</td>
<td>0.839</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>18 s</td>
<td>12585.79</td>
<td>0.951</td>
</tr>
<tr>
<td>China</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>71 s</td>
<td>68991.73</td>
<td>0.866</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>45 s</td>
<td>80104.27</td>
<td>0.865</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>17 s</td>
<td>24987.34</td>
<td>0.895</td>
</tr>
<tr>
<td>Iran</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>89 s</td>
<td>1436025.84</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>70 s</td>
<td>3745673.26</td>
<td>0.744</td>
</tr>
<tr>
<td></td>
<td>GWO</td>
<td>500</td>
<td>1000</td>
<td>21 s</td>
<td>13635.01</td>
<td>0.864</td>
</tr>
<tr>
<td>USA</td>
<td>GA</td>
<td>300</td>
<td>500</td>
<td>84 s</td>
<td>457051.4</td>
<td>0.974</td>
</tr>
<tr>
<td></td>
<td>PSO</td>
<td>1000</td>
<td>500</td>
<td>30 s</td>
<td>982.37</td>
<td>0.932</td>
</tr>
</tbody>
</table>
According to Tables 5 to 12, GWO provided the highest accuracy (smallest RMSE and largest correlation coefficient) and smallest processing time compared to PSO and GA for fitting the logistic, linear, logarithmic, quadratic, cubic, power, compound, and exponential-based equations for all five countries. It can be suggested that GWO is a sustainable optimizer due to its acceptable processing time compared with PSO and GA. Therefore, GWO was selected as the best optimizer by providing the highest accuracy values compared with PSO and GA. In general, it can be claimed that GWO, by suggesting the best parameter values for the functions presented in Table 2, increases outbreak prediction accuracy for COVID-19 in comparison with PSO and GA. Therefore, the functions derived by GWO were selected as the best predictors for this research.

Tables 13 to 17 present the description and coefficients of the linear, logarithmic, quadratic, cubic, compound, power, exponential and logistic equations estimated by GWO. Tables 13 to 17 also present the RMSE and r-square values for each equation fitted to data for China, Italy, Iran, Germany and USA, respectively.

### Table 13. Model description for China fitted by GWO

<table>
<thead>
<tr>
<th>Model name</th>
<th>Description</th>
<th>RMSE</th>
<th>r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$R = 3036.4 \times x - 13509.84$</td>
<td>5039.48</td>
<td>0.964</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$R = -33948.15 + 27124.70 \times \log(x)$</td>
<td>42714.93</td>
<td>0.718</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$R = -5080.88 + 1455.98 \times x + 50.98 \times x^2$</td>
<td>3710.16</td>
<td>0.98</td>
</tr>
<tr>
<td>Cubic</td>
<td>$R = 3984.73 - 1790.2 \times x + 308.52 \times x^2 - 5.53 \times x^3$</td>
<td>2429.45</td>
<td>0.99</td>
</tr>
<tr>
<td>Compound</td>
<td>$R = 1601.03 \times x^{1.16}$</td>
<td>24987.34</td>
<td>0.801</td>
</tr>
<tr>
<td>Power</td>
<td>$R = 262.27 \times x^{1.69}$</td>
<td>4079.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Exponential</td>
<td>$R = 1601.03 \times \exp(0.15 \times x)$</td>
<td>24987.34</td>
<td>0.801</td>
</tr>
<tr>
<td>Logistic</td>
<td>$R = 85011.297/(1 + \exp(((4 \times 4483.304) \times (9.423 - x))/85011.297) + 2))$</td>
<td>2270.58</td>
<td>0.992</td>
</tr>
</tbody>
</table>

### Table 14. Model description for Italy fitted by GWO

<table>
<thead>
<tr>
<th>Model name</th>
<th>Description</th>
<th>RMSE</th>
<th>r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$R = 663.71 \times x - 5437.25$</td>
<td>3642.44</td>
<td>0.713</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$R = -7997.93 + 5162.83 \times \log(x)$</td>
<td>9296.59</td>
<td>0.402</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$R = 2998.21 - 917.93 \times x + 51.02 \times x^2$</td>
<td>1272.1</td>
<td>0.965</td>
</tr>
</tbody>
</table>
Table 15. Model description for Iran fitted by GWO

<table>
<thead>
<tr>
<th>Model name</th>
<th>Description</th>
<th>RMSE</th>
<th>r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$R = 656,068 \times x - 4527,69$</td>
<td>1981.97</td>
<td>0.891</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$R = -7921,009 + 5449,784 \times \log(x)$</td>
<td>8995.52</td>
<td>0.574</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$R = 310,48 - 251,09 \times x + 29,26 \times x^2$</td>
<td>310.027</td>
<td>0.997</td>
</tr>
<tr>
<td>Cubic</td>
<td>$R = 902,33 - 463,02 \times x + 46,07 \times x^2 - 0,36 \times x^3$</td>
<td>250.204</td>
<td>0.998</td>
</tr>
<tr>
<td>Compound</td>
<td>$R = 13,26 \times 1,33^x$</td>
<td>13635.014</td>
<td>0.748</td>
</tr>
<tr>
<td>Power</td>
<td>$R = 0,51 \times x^{3.09}$</td>
<td>1031.607</td>
<td>0.982</td>
</tr>
<tr>
<td>Exponential</td>
<td>$R = 13,26 \times \exp(0,28 \times x)$</td>
<td>13635.014</td>
<td>0.748</td>
</tr>
<tr>
<td>Logistic</td>
<td>$R = 21936,052/(1 + \exp(((4 \times 3962,88) \times (23,88 - x)/70731,08) + 2))$</td>
<td>392.88</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Table 16. Model description for Germany fitted by GWO

<table>
<thead>
<tr>
<th>Model name</th>
<th>Description</th>
<th>RMSE</th>
<th>r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$R = 128,421 \times x - 1130,294$</td>
<td>951.635</td>
<td>0.577</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$R = -1528,684 + 959,941 \times \log(x)$</td>
<td>1878.672</td>
<td>0.3</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$R = 911,113 - 254,342 \times x + 12,347 \times x^2$</td>
<td>472.624</td>
<td>0.895</td>
</tr>
<tr>
<td>Cubic</td>
<td>$R = -478,087 + 243,097 \times x - 27,118 \times x^2 + 0,848 \times x^3$</td>
<td>196.809</td>
<td>0.981</td>
</tr>
<tr>
<td>Compound</td>
<td>$R = 3,821 \times 1,263^x$</td>
<td>431.975</td>
<td>0.996</td>
</tr>
</tbody>
</table>
Table 17. Model description for USA fitted by GWO

<table>
<thead>
<tr>
<th>Model name</th>
<th>Description</th>
<th>RMSE</th>
<th>r-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$R = 76,833 \times x - 666,79$</td>
<td>592.486</td>
<td>0.557</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$R = -902,637 + 573,32 \times \log(x)$</td>
<td>1135.124</td>
<td>0.289</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$R = 584,76 - 157,831 \times x + 7,569 \times x^2$</td>
<td>307.585</td>
<td>0.88</td>
</tr>
<tr>
<td>Cubic</td>
<td>$R = -333,235 + 170,881 \times x - 18,509 \times x^2 + 0.56 \times x^3$</td>
<td>118.247</td>
<td>0.982</td>
</tr>
<tr>
<td>Compound</td>
<td>$R = 6,296 \times 1,214^x$</td>
<td>364.875</td>
<td>0.977</td>
</tr>
<tr>
<td>Power</td>
<td>$R = 1,707 \times x^{1.735}$</td>
<td>790.163</td>
<td>0.702</td>
</tr>
<tr>
<td>Exponential</td>
<td>$R = 6,296 \times \text{EXP}(0,194 \times x)$</td>
<td>364.875</td>
<td>0.977</td>
</tr>
<tr>
<td>Logistic</td>
<td>$R = 32604,552/(1 + \text{EXP}(((4 \times 2288,932) \times (30,303 - x)/32604,552) + 2))$</td>
<td>22.354</td>
<td>0.999</td>
</tr>
</tbody>
</table>

As is clear from Tables 13 to 17, in general, the logistic equation followed by the quadratic and cubic equations provided the smallest RMSE and the largest r-square values for the prediction of COVID-19 outbreak. The claim can also be considered from Figure 7 to 11, which presents the capability and trend of each model derived by GWO in the prediction of COVID-19 cases for China, Italy, Iran, Germany, and the USA, respectively.
Figure 7. Fitness graph for China fitted by GWO

Figure 8. Set of models for Italy fitted by GWO
Figure 9. Set of models for Iran fitted by GWO

Figure 10. Set of models for Germany fitted by GWO
Figures 7 to 11 illustrate the fit of the models investigated in this paper. The best fit for the prediction of COVID-19 cases was achieved for the logistic model followed by cubic and quadratic models for China (Figure 7), logistic followed by cubic models for Italy (Figure 8), cubic followed by logistic and quadratic models for Iran (Figure 9), the logistic model for Germany (Figure 10), and logistic model for the USA (Figure 11).

Machine learning results

This section presents the results for the training stage of ML methods. MLP and ANFIS were employed as single and hybrid ML methods, respectively. ML methods were trained using two datasets related to scenario 1 and scenario 2. Table 14 presents the results of the training phase.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MLP</th>
<th>ANFIS</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of neurons</td>
<td>r</td>
<td>RMSE</td>
<td>r</td>
<td>RMSE</td>
</tr>
<tr>
<td>Italy</td>
<td>8</td>
<td>0.999</td>
<td>190.81</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.999</td>
<td>194.84</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0.999</td>
<td>188.18</td>
<td>0.998</td>
</tr>
<tr>
<td>Average</td>
<td>0.999</td>
<td>191.27</td>
<td>0.946</td>
<td>1418.1</td>
</tr>
<tr>
<td>China</td>
<td>8</td>
<td>0.995</td>
<td>2287.55</td>
<td>0.996</td>
</tr>
</tbody>
</table>
According to Table 18, the dataset related to scenarios 1 and 2 have different performance values. Accordingly, for Italy, the MLP with 16 neurons provided the highest accuracy for scenario 1 and ANFIS with tri. MF provided the highest accuracy for scenario 2. By considering the average values of the RMSE and correlation coefficient, it can be concluded that scenario 1 is more suitable for modeling outbreak cases in Italy, as it provides a higher accuracy (the smallest RMSE and the largest correlation coefficient) than scenario 2.

For the dataset related to China, for both scenarios, MLP with 12 and 16 neurons, respectively for scenarios 1 and 2, provided the highest accuracy compared with the ANFIS model. By considering the average values of RMSE and correlation coefficient, it can be concluded that scenario 2 with a larger average correlation coefficient and smaller average RMSE than scenario 1 is more suitable for modeling the outbreak in China.

For the dataset of Iran, MLP with 12 neurons in the hidden layer for scenario 1 and ANFIS with Gaussian MF type for scenario 2 provided the best performance for the prediction of the outbreak. By considering the average values of the RMSE and correlation coefficient, it can be concluded that scenario 1 provided better performance than scenario 2. Also, in general, the MLP has higher prediction accuracy compared with the ANFIS method.

In Germany, MLP with 12 neurons in its hidden layer provided the highest accuracy (smallest RMSE and largest correlation coefficient). By considering the average values of the RMSE and correlation coefficient, it can be concluded that scenario 1 is more suitable for the prediction of the outbreak in Germany than scenario 2.

In the USA, the MLP with 8 and 12 neurons, respectively, for scenarios 1 and 2, provided higher accuracy (the smallest RMSE and the largest correlation coefficient values) than the ANFIS model. By considering the average values of the RMSE and correlation coefficient values, it can be concluded that scenario 1 is more suitable than scenario 2, and MLP is more suitable than ANFIS for outbreak prediction.

Figures 12 to 16 present the model fits for Italy, China, Iran, Germany, and the USA, respectively. By comparing Figure 12 to 16 with Figures 7 to 11, it can be concluded that the MLP and the logistic model fitted by GWO provided a better fit than the other models. In addition, the ML methods provided better performance compared with other models.
Figure 12. Set of models for Italy fitted by ML methods

Figure 13. Set of models for China fitted by ML methods
Figure 14. Set of models for Iran fitted by ML methods

Figure 15. Set of models for Germany fitted by ML methods
Comparing the fitted models

This section presents a comparison of the accuracy and performance of the selected models for the prediction of 30 days’ outbreak. Figure 17 to 21 shows the deviation from the target values for the selected models.

**Figure 16.** Set of models for USA fitted by ML methods

![Graph showing the deviation from target values for models related to Italy]
Figure 18. Deviation from target value for models related to China

Figure 19. Deviation from target value for models related to Iran

Figure 20. Deviation from target value for models related to Germany
As is clear from Figure 17 to 21, the smallest deviation from the target values is related to the MLP for scenario 1 followed by MLP for scenario 2. This indicates the highest performance of the MLP method for the prediction of the outbreak.

Figures 22 to 26 present the outbreak prediction for 75 days and Tables 19 to 23 present the outbreak prediction for 150 days.

Figure 21. Deviation from target value for models related to USA

Figure 22. The outbreak prediction for Italy through 75 days
Table 19. The outbreak prediction for Italy through 150 days

<table>
<thead>
<tr>
<th></th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 20th</td>
<td>3794.045</td>
<td>7837.054</td>
<td>-1280.93</td>
<td>5047.906</td>
<td>3225.523</td>
<td>3792.734</td>
<td>3796.738</td>
</tr>
<tr>
<td>Day 40th</td>
<td>58966.55</td>
<td>21111.37</td>
<td>273.235</td>
<td>47914.4</td>
<td>35898.08</td>
<td>58966.74</td>
<td>58964.96</td>
</tr>
<tr>
<td>Day 60th</td>
<td>70571.86</td>
<td>34385.68</td>
<td>1182.365</td>
<td>131597.7</td>
<td>146966.2</td>
<td>70571.66</td>
<td>70572.12</td>
</tr>
<tr>
<td>Day 80th</td>
<td>70729.28</td>
<td>47659.99</td>
<td>1827.402</td>
<td>256097.8</td>
<td>399523.4</td>
<td>70729.27</td>
<td>70729.15</td>
</tr>
<tr>
<td>Day 100th</td>
<td>70731.06</td>
<td>60934.31</td>
<td>2327.733</td>
<td>421414.7</td>
<td>867822</td>
<td>70731.09</td>
<td>70730.93</td>
</tr>
<tr>
<td>Day 120th</td>
<td>70731.08</td>
<td>74208.62</td>
<td>2736.532</td>
<td>627548.4</td>
<td>1635643</td>
<td>70731.14</td>
<td>70730.87</td>
</tr>
<tr>
<td>Day 140th</td>
<td>70731.08</td>
<td>87482.94</td>
<td>3082.167</td>
<td>874498.9</td>
<td>2795218</td>
<td>70731.19</td>
<td>70730.79</td>
</tr>
<tr>
<td>Day 150th</td>
<td>70731.08</td>
<td>94120.09</td>
<td>3236.862</td>
<td>1013280</td>
<td>3552851</td>
<td>70731.21</td>
<td>70730.75</td>
</tr>
</tbody>
</table>

Figure 23. The outbreak prediction for China through 75 days

Table 20. The outbreak prediction for China through 150 days

<table>
<thead>
<tr>
<th></th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
</table>

*Note: The data is available under a CC-BY 4.0 International license.*
<table>
<thead>
<tr>
<th>Day</th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>20th</td>
<td>47397.6</td>
<td>47218.47</td>
<td>1341.899</td>
<td>44431.48</td>
<td>41916.55</td>
<td>47397.6</td>
<td>47360.98</td>
</tr>
<tr>
<td>40th</td>
<td>84030.16</td>
<td>107946.8</td>
<td>9507.249</td>
<td>134729.1</td>
<td>135599.1</td>
<td>84030.17</td>
<td>84030.39</td>
</tr>
<tr>
<td>60th</td>
<td>84996.7</td>
<td>168675.1</td>
<td>14283.67</td>
<td>437680.2</td>
<td>438660.2</td>
<td>85011.08</td>
<td>85011.05</td>
</tr>
<tr>
<td>80th</td>
<td>85011.08</td>
<td>229403.4</td>
<td>17672.6</td>
<td>84030.39</td>
<td>84030.17</td>
<td>85011.22</td>
<td></td>
</tr>
<tr>
<td>100th</td>
<td>85011.29</td>
<td>290131.7</td>
<td>20301.26</td>
<td>650333.6</td>
<td>640132.8</td>
<td>85011.3</td>
<td>85011.13</td>
</tr>
<tr>
<td>120th</td>
<td>85011.3</td>
<td>350860</td>
<td>22449.02</td>
<td>903772.3</td>
<td>871733.6</td>
<td>85011.3</td>
<td>85011.13</td>
</tr>
<tr>
<td>140th</td>
<td>85011.3</td>
<td>411588.3</td>
<td>24264.94</td>
<td>1197996</td>
<td>1131815</td>
<td>85011.3</td>
<td>85011.05</td>
</tr>
<tr>
<td>150th</td>
<td>85011.3</td>
<td>441952.5</td>
<td>25077.68</td>
<td>1360403</td>
<td>1272113</td>
<td>85011.41</td>
<td>85011.01</td>
</tr>
</tbody>
</table>

**Figure 24.** The outbreak prediction for Iran through 75 days

**Table 21.** The outbreak prediction for Iran through 150 days

<table>
<thead>
<tr>
<th>Day</th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>20th</td>
<td>6898.344</td>
<td>8593.676</td>
<td>-830.677</td>
<td>6993.955</td>
<td>5494.377</td>
<td>6902.315</td>
<td>6875.585</td>
</tr>
<tr>
<td>40th</td>
<td>21455.58</td>
<td>21715.05</td>
<td>809.8719</td>
<td>37087.98</td>
<td>47060.48</td>
<td>21457.4</td>
<td>21456.65</td>
</tr>
<tr>
<td>60th</td>
<td>21931.01</td>
<td>34836.43</td>
<td>1769.531</td>
<td>90592.56</td>
<td>165300.1</td>
<td>21932.24</td>
<td>21930.68</td>
</tr>
</tbody>
</table>
Table 22. The outbreak prediction for Germany through 150 days

<table>
<thead>
<tr>
<th>Day</th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>20th</td>
<td>431.027</td>
<td>1438.128</td>
<td>-279.772</td>
<td>763.1467</td>
<td>400.0548</td>
<td>432.8991</td>
<td>431.8119</td>
</tr>
<tr>
<td>40th</td>
<td>35356.27</td>
<td>4006.551</td>
<td>9.199328</td>
<td>10492.96</td>
<td>1624.405</td>
<td>35355.14</td>
<td>35355.72</td>
</tr>
<tr>
<td>60th</td>
<td>55043.44</td>
<td>6574.974</td>
<td>178.2366</td>
<td>30100.56</td>
<td>3687.126</td>
<td>55036.14</td>
<td>55044.03</td>
</tr>
<tr>
<td>80th</td>
<td>55179.07</td>
<td>9143.397</td>
<td>298.1705</td>
<td>59585.93</td>
<td>6595.829</td>
<td>55179.05</td>
<td>55178.88</td>
</tr>
<tr>
<td>100th</td>
<td>55179.67</td>
<td>11711.82</td>
<td>391.1984</td>
<td>98949.09</td>
<td>10355.87</td>
<td>55179.9</td>
<td>55179.47</td>
</tr>
<tr>
<td>120th</td>
<td>55179.67</td>
<td>14280.24</td>
<td>467.2078</td>
<td>148190</td>
<td>14971.42</td>
<td>55179.92</td>
<td>55179.42</td>
</tr>
</tbody>
</table>

Figure 25. The outbreak prediction for Germany through 75 days
### Table 23. The outbreak prediction for the USA for 150 days

<table>
<thead>
<tr>
<th>Day</th>
<th>Logistic by GWO</th>
<th>Linear by GWO</th>
<th>Logarithmic by GWO</th>
<th>Quadratic by GWO</th>
<th>Power by GWO</th>
<th>MLP</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>20th</td>
<td>242.6091</td>
<td>869.8855</td>
<td>-156.73</td>
<td>456.0663</td>
<td>309.616</td>
<td>244.0038</td>
<td>243.6504</td>
</tr>
<tr>
<td>40th</td>
<td>21951.15</td>
<td>2406.562</td>
<td>15.85698</td>
<td>6383.264</td>
<td>1031.324</td>
<td>21942.25</td>
<td>21948.25</td>
</tr>
<tr>
<td>60th</td>
<td>32547.08</td>
<td>3943.238</td>
<td>116.8138</td>
<td>18366.35</td>
<td>2084.876</td>
<td>32552.6</td>
<td>32548.47</td>
</tr>
<tr>
<td>80th</td>
<td>32604.34</td>
<td>5479.914</td>
<td>188.4437</td>
<td>36405.33</td>
<td>3435.319</td>
<td>32606.19</td>
<td>32604.47</td>
</tr>
<tr>
<td>100th</td>
<td>32604.55</td>
<td>7016.591</td>
<td>244.0043</td>
<td>60500.21</td>
<td>5060.548</td>
<td>32606.63</td>
<td>32604.72</td>
</tr>
<tr>
<td>120th</td>
<td>32604.55</td>
<td>8553.267</td>
<td>289.4005</td>
<td>90650.97</td>
<td>6944.676</td>
<td>32606.7</td>
<td>32604.76</td>
</tr>
<tr>
<td>140th</td>
<td>32604.55</td>
<td>10889.94</td>
<td>327.7825</td>
<td>126857.6</td>
<td>9075.446</td>
<td>32606.78</td>
<td>32604.8</td>
</tr>
<tr>
<td>150th</td>
<td>32604.55</td>
<td>10858.28</td>
<td>344.9611</td>
<td>147231.9</td>
<td>10230.16</td>
<td>32606.81</td>
<td>32604.82</td>
</tr>
</tbody>
</table>

Figure 26. The outbreak prediction for the USA through 75 days
4. Discussion

The parameters of several simple mathematical models (i.e., logistic, linear, logarithmic, quadratic, cubic, compound, power and exponential) were fitted using GA, PSO, and GWO. The logistic model outperformed other methods and showed promising results on training for 30 days. Extrapolation of the prediction beyond the original observation range of 30 days should not be expected to be realistic considering the new statistics. The fitted models generally showed low accuracy and also weak generalization ability for the five countries. Although the prediction for China was promising, the model was insufficient for extrapolation, as expected. In turn, the logistic GWO outperformed the PSO and GA and the computational cost for GWO was reported as satisfactory. Consequently, for further assessment of the ML models, the logistic model fitted with GWO was used for comparative analysis.

In the next step, for introducing the machine learning methods for time-series prediction, two scenarios were proposed. Scenario 1 considered four data samples from the progress of the infection from previous days, as reported in table 3. The sampling for data processing was done weekly for scenario 1. However, scenario 2 was devoted to daily sampling for all previous consecutive days. Providing these two scenarios expanded the scope of this study. Training and test results for the two machine learning models (MLP and ANFIS) were considered for the two scenarios. A detailed investigation was also carried out to explore the most suitable number of neurons. For the MLP, the performances of using 8, 12 and 16 neurons were analyzed throughout the study. For the ANFIS, the membership function (MF) types of Tri, Trap, and Gauss were analyzed throughout the study. The five counties of Italy, China, Iran, Germany, and USA were considered. The performance of both ML models for these countries varied amongst the two different scenarios. Given the observed results, it is not possible to select the most suitable scenario. Therefore, both daily and weekly sampling can be used in machine learning modeling. Comparison between analytical and machine learning models using the deviation from the target value (figures 17 to 21) indicated that the MLP in both scenarios delivered the most accurate results. Extrapolation for long-term prediction of up to 150 days using the ML models was tested. The actual prediction of MLP and ANFIS for the five countries was reported which showed the progression of the outbreak.

5. Conclusions

The global pandemic of the severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) has become the primary national security issue of many nations. Advancement of accurate prediction models for the outbreak is essential to provide insights into the spread and consequences of this infectious disease. Due to the high level of uncertainty and lack of crucial data, standard epidemiological models have shown low accuracy for long-term prediction. This paper presents a comparative analysis of ML and soft computing models to predict the COVID-19 outbreak. The results of two ML models (MLP and ANFIS) reported a high generalization ability for long-term prediction. With respect to the results reported in this paper and due to the highly complex nature of the COVID-19 outbreak and differences from nation-to-nation, this study suggests ML as an effective tool to model the outbreak.

For the advancement of higher performance models for long-term prediction, future research should be devoted to comparative studies on various ML models for individual countries. Due to the fundamental differences between the outbreak in various countries, advancement of global models with generalization ability would not be feasible. As observed and reported in many studies, it is unlikely that an individual outbreak will be replicated elsewhere [1].

Although the most difficult prediction is to estimate the maximum number of infected patients, estimation of the \( n(\text{deaths}) / n(\text{infecteds}) \) is also essential. The mortality rate is particularly important to accurately estimate the number of patients and the required beds in intensive care units. For future research, modeling the mortality rate would be of the utmost importance for nations to plan for new facilities.
Nomenclatures

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Acronym</th>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-layered perceptron</td>
<td>MLP</td>
<td>Grey wolf optimization</td>
<td>GWO</td>
</tr>
<tr>
<td>Adaptive network-based fuzzy inference system</td>
<td>ANFIS</td>
<td>Mean square error</td>
<td>MSE</td>
</tr>
<tr>
<td>Susceptible-infected-recovered</td>
<td>SIR</td>
<td>Root mean square error</td>
<td>RMSE</td>
</tr>
<tr>
<td>Call data record</td>
<td>CDR</td>
<td>Artificial intelligence</td>
<td>AI</td>
</tr>
<tr>
<td>Classification and regression tree</td>
<td>CART</td>
<td>Artificial neural network</td>
<td>ANN</td>
</tr>
<tr>
<td>Evolutionary algorithms</td>
<td>EA</td>
<td>Triangular</td>
<td>Tri.</td>
</tr>
<tr>
<td>Genetic algorithm</td>
<td>GA</td>
<td>Gaussian</td>
<td>Gauss.</td>
</tr>
<tr>
<td>Particle swarm optimization</td>
<td>PSO</td>
<td>Trapezoidal</td>
<td>Trap.</td>
</tr>
<tr>
<td>Membership function</td>
<td>MF</td>
<td>Machine learning</td>
<td>ML</td>
</tr>
</tbody>
</table>

Conflicts of Interest: The authors declare no conflict of interest.

References


47. Deb, K. *Multi-objective optimization using evolutionary algorithms*; John Wiley & Sons: 2001; Vol. 16.


80. Taghavifar, H.; Mardani, A. Wavelet neural network applied for prognostication of contact pressure between soil and driving wheel. \textit{Information Processing in Agriculture} \textbf{2014}, \textit{1}, 51-56.


