Results in Physics 27 (2021) 104484

Contents lists available at ScienceDirect

Results in Physics

journal homepage: www.elsevier.com/locate/rinp

Forecast of the outbreak of COVID-19 using artificial neural network: Case study Qatar, Spain, and Italy

Moayyad Shawaqfah [a], Fares Almomani [b,*]

[a] Department of Civil Engineering, Faculty of Engineering, Al al-Bayt University, Mafraq 25113, Jordan
[b] Department of Chemical Engineering, College of Engineering, Qatar University, P.O. Box 2713, Doha, Qatar

ARTICLE INFO

Keywords:
COVID-19
Outbreak
Prediction
Political
The decision-maker

ABSTRACT

The present study illustrates the outbreak prediction and analysis on the growth and expansion of the COVID-19 pandemic using artificial neural network (ANN). The first wave of the pandemic outbreak of the novel Coronavirus (SARS-CoV-2) began in September 2019 and continued to March 2020. As declared by the World Health Organization (WHO), this virus affected populations all over the globe, and its accelerated spread is a universal concern. An ANN architecture was developed to predict the serious pandemic outbreak impact in Qatar, Spain, and Italy. Official statistical data gathered from each country until July 6th was used to validate and test the prediction model. The model sensitivity was analyzed using the root mean square error (RMSE), the mean absolute percentage error and the regression coefficient index R^2, which yielded highly accurate values of the predicted correlation for the infected and dead cases of 0.99 for the dates considered. The verified and validated growth model of COVID-19 for these countries showed the effects of the measures taken by the government and medical sectors to alleviate the pandemic effect and the effort to decrease the spread of the virus in order to reduce the death rate. The differences in the spread rate were related to different exogenous factors (such as social, political, and health factors, among others) that are difficult to measure. The simple and well-structured ANN model can be adapted to different propagation dynamics and could be useful for health managers and decision-makers to better control and prevent the occurrence of a pandemic.

Introduction

The outbreak of the novel coronavirus (COVID-19) which began December 2019 in the city of Wuhan (China) and later spread rapidly in March 2020 to different countries, has had several impacts on public health and the worldwide economy [1–3]. The COVID-19 virus was detected and isolated from a single patient in late December, and afterward identified and verified in additional patients [4,5]. The genome analysis of the novel virus known as COVID-19 or SARS-CoV-2 revealed that the DNA sequence has up to 96% similarity to bat coronaviruses, sharing properties with other pathogenic viruses such as SARS-CoV and MERS-CoV [6–10]. Different studies have shown that the main concerns related to SARS-CoV-2 are its high transmission potential that has been responsible for the global COVID-19 outbreak [11–13]. The fast-spreading rate of COVID-19 is a result of different transmission methods, including one that is proved to be the most effective: direct contact between humans[14]. This close contact facilitates transmission more rapidly as droplets are expelled through individuals coughing or sneezing [15,16]. As this virus is proven to be airborne, it has increased chance of transmission, rate of contagiousness and a high survival period, lasting for up to 9 days within different materials [17–19]. Based on preliminary observation, it was proposed that COVID-19 has an incubation period ranging from 3 to 10 days with an average incubation period of 5.2 days [20,21]. Other studies suggest a period from 2 to 14 days in length [22,23]. Furthermore, personnel during the incubation period can transmit the infection, creating a very arduous identification process resulting in the number of infected people being higher than the official count [24].

The spread of COVID-19 reached more than 177,419,000 cases on July 16, 2021 in more than 110 countries, with a total death toll of up to 3,839,600 people. As a result of the high risk the fast outbreak of this virus, it was declared a global pandemic by the Director-General of WHO, which had a direct effect on public health requiring urgent international attention [25,26]. This declaration encouraged all countries to take serious remedial action to prevent the spread of the virus among their citizens, protect public health, and if required ban travels and close

* Corresponding author.
E-mail address: falmomani@qu.edu.qa (F. Almomani).

https://doi.org/10.1016/j.rinp.2021.104484
Received 19 March 2021; Received in revised form 17 June 2021; Accepted 18 June 2021
Available online 21 June 2021

borders. The virus outbreak in Europe started in the northern zone on February 22, 2020, reaching closure on March 9, 2020, and suspending all national activities on March 21, 2020. Following that, Spain implemented the closure of commercial activities on March 15th 2020 [27]. In the Middle East, initial cases were reported in Qatar at the beginning of March and the number increased to more than 400 in the 25 days leading to the closure of the industrial city, suspension of schools, universities, followed by a full closure within 20 days. In other countries, the precautions were focused on performing initial patient screenings and performing massive tests to all suspected people and for anyone who had been in contact with them, isolating only the affected people and areas [28,29]. With all these measures taken, the virus continued to spread throughout the world marking the largest quarantine in history. The statistics showed that the confirmed infected cases increased during the period of May 10, 2020 to June 15, 2021 from 3,900,000 to more than 177,00,000 cases, with more than 270,000 to 3,800,000 death cases reported in more than 180 affected countries, respectively [30].

Compared to SARS or MERS, COVID-19 is more infectious with high R_0 values that are superadded across the globe at an alarming rate causing more infections and high mortality rates [24]. The fast spread rate and the high chance of transmission indicate that the employed prevention and control strategies including isolation, detection tests, and prophylactic measures, although helping flatten the epidemic curve are not effective to limit, prevent or stop the spread of the virus throughout the world. Therefore, there is an urgent need to develop models that can be used for the prediction of the outbreak of the pandemic disease in different areas. The developed model will help decision-makers and physicians to be prepared, understand the real magnitude of the risk, and take the appropriate prevention measures. Prediction tools can also help to estimate the risk volume and prepare the required control measures with sufficient advance time.

Different models were used in epidemiology history to calculate the outbreaks of different diseases [31,32]. One of the models that is often used is the classic model of susceptible-infected-recovered (SIR) developed by Kermack and McKendrick [33,34]. Various prediction models have been studied, based on the SEIR model [35,36] and the logistic model has been used to successfully predict for 20-day infections [37]. Other subsequent studies using more complicated models with multiple variants of SIR patterns were used to predict the outbreaks of SARS [38] and Cholera [39]. Vanderpast al. [37] incorporated the exposed population and their corresponding immunity within the model. Furthermore, another logistic model has been used to predict the growth and development of diseases similar to bacterial growth [40] or infectious diseases [41]. Other studies have used the Gompertz model, which is usually used for bacterial sprouting in predicting virus outbreaks [42,43]. Recently, Wu et al. [44] studied the growth of COVID-19 and predicted the national spread of the pandemic in China using the SEIR model. Yang and Wang [45] used the confinement variable to predict the national spread of COVID-19 in China. Other studies have also studied and estimated infected and non-infected population in different areas of China [46] and flights from Japan [47]. Additionally, the epidemic evolution model has also been studied following the system of differential equations for the susceptible-infected-recovered-dead (SIRD) variables. However, all of these models have required accurate initial data, temporal dynamics of the disease, and growth rate to achieve precise prediction, which is not possible in the case of COVID-19.

Recently, artificial neural networks (ANN) have been successfully used for the prediction of the evolution of different systems with a high degree of accuracy[48]. The ANN demonstrated an excellent prediction accuracy short time, for different engineering-based processes such as wastewater treatment [49,50], electro-dialysis separation, metal removal [51], biosorption of heavy metals [52], anaerobic digestion [53], biofuel production [54–57,52], cell growth rate [58] and population growth [59]. The ANN architecture is a simple and fast methodology to predict the process output compared with the complicated physically related models. In general, ANN predictions depend on reasoning and formulation of a mathematical relationship between the inputs collected data (ICD) and system output function (SOF) without the need for previous physical correlation. The neutrons (*Ns*) connect the store and manipulate the ICD to produce the SOF using different combinations of the transfer functions (TFs). The intensities of the signals from different *Ns* determine the contribution of the ICD to the SOF through the different layers of the ANN. Therefore, this study presents the first time development of ANN architecture to predict the spread of COVID-19 to understand the evolution of epidemics over time. The developed ANN architecture is capable of forecasting the outbreak behaviors (number of infections and mortality) to help health systems and politicians in predicting future situations for better decision-making in the control and prevention of this pandemic. The ANN architecture was applied on two different European countries (Spain and Italy) and one Middle Eastern country (Qatar) to test its application.

Model manipulation

Data collection

The prediction of the pandemic growth of COVID-19 in different countries was followed using the ANN algorithm. The breakout was tested in two different continents: the Middle East represented by Qatar with low population density, hot and humid climate, and a high number of reported cases per million of the population (PM_{IN}), compared to the European countries, Spain, and Italy, with high population density, Mediterranean mild climate conditions and medium to high number of reported cases PM_{IN}. The reported data of the infected/death cases were obtained from WHO [30], with daily reports presented worldwide from the European Center for Disease Prevention and Control [60] and daily reports issued by the Ministry of Health in Qatar, which are publicly available in the ministry website.

Artificial neural network model

The ANN architecture simulates both linear and non-linear systems by combining and reasoning the influence of the inputs collected data on the SOF. The ICD was introduced into the ANN architecture to generate a mathematical SOF without the requirement for a previous physical relationship. The ICDs are connected within the ANN via nodes known as neutrons (*Ns*). The ICDs are received, stored, and manipulated by *Ns* via TFs at changeable intensities depending on the contribution of the ICD to the SOF. The forecast of the ANN is refined by error tuning using a feed forward-back propagation neural network (FF-BPNN) algorithm. In the FF-BPNN algorithm, the prediction of the ANN is tuned by deploying the flow of ICD within the ANN layers (input layer [IL], hidden layer [HL], and the output layer [OL]. The *Ns* between layers are related by the connection weights (CW_{ij}) attuned by mapping competency of the trained network and activated by a bias value (β_j). The effect of the ICD entered at the IL is calculated as ICD (X_i) and transferred to HL*s* to categorize the correlation between the dependent and independent parameters. The summation of the weighted output (Σ CID (X_i) is added to a threshold bias (B_i) (i.e., $\sum_{j=1}^{n} \mathrm{CID}(\mathrm{Xi}) + B_i$) and the outcome is transferred to OL via specific TF. The SOF in the OL is compared with measured data and the associated error (AE) determined. After that, the AE signal returned backward to the IL via HL renewing both the weighted ICD and B_i at each *Ns* to minimize the error signal. This iterative tactic continues to update the Σ CID (X_i) in different *Ns* until the minimum required error is reached. After that, training of the ANN architecture is complete and another set of data known as testing data is used to verify the prediction ability.

The algorithm used in developing the ANN architecture is presented in Fig. 1a. Initially, the number of cases of the spread and death were collected for each country until June 16, 2021. Then the initial pandemic and death dates were adjusted for each country and the rate of

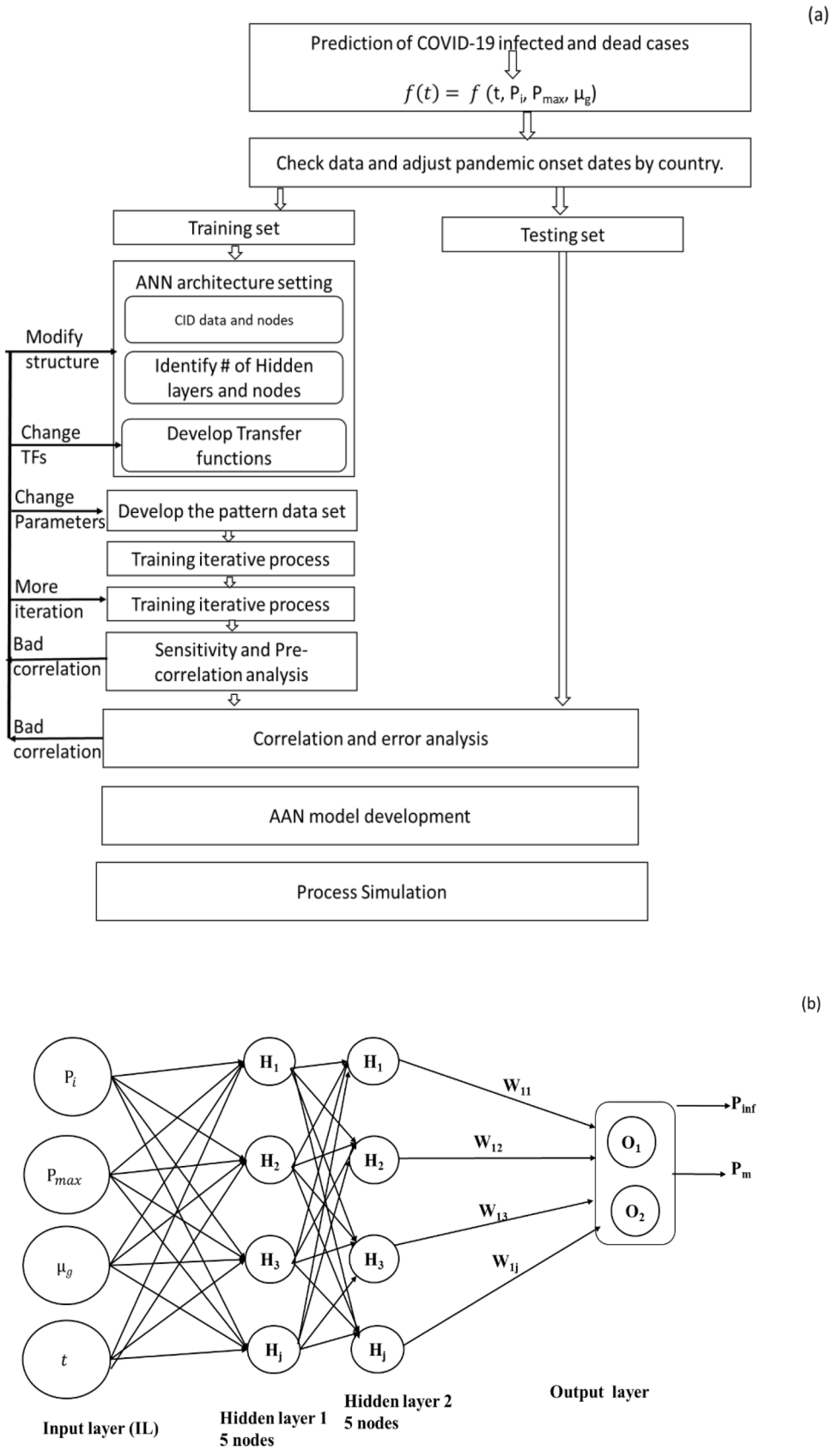

Fig. 1. (a) Processing neural network model (b) Configuration of the ANN for the COVID-19 prediction.

spread and the maximum number of cases were determined. The data was entered into the IL and moved within the ANN along with the calculated CW_{ij} and the sum of the weighted output ($\sum_{j=1}^{n} W_{ij}^{l} . X_{j}^{l-1}$). The CW_{ij} between layers was performed using the Levenberg-Marquardt back-propagation tuning algorithm (L-M-BB-TA) and the prediction was refined using feed-forward back-propagation neural network (FF-BP-NN) until the minimum acceptable error achieved. Sigmoid (S), hyperbolic secant (HS), hyperbolic tangent (HT), and Gaussian (G) transfer functions were utilized to determine the number of infected/death cases to an acceptable error. The ANN architecture was established and validated with a separate set of data and the model prediction was determined.

The arrangement of the ANN employed in the present study is shown in Fig. 1b. The growth of the pandemic disease and mortality can be represented as a function of time (t), the number of initial patients (infected or dead) (P_i), the maximum predicted number of patients infected or dead (P_{max}), and the growth rate characteristic of the pandemic (μ_g) as per equation (1).

$$f(t) = f(\mathrm{t}, \mathrm{Pi}, \mathrm{Pmax}, \mu\mathrm{g}) \tag{1}$$

The training, testing, and validation of the ANN was conducted using MATLAB® software (MathWork, Inc.,Version: R2010a) and following L-M-BB-TA algorithms. A total of 5400 ICD from the three countries were used in the calculations. The ICD manipulation was separated into the 56% training set, 24% testing, and the balance for validation subsets. The ICD data were normalized with respect to minimum/maximum values previously used in the ANN to decrease the chance of local minima.

Statistical analysis

The ANN prediction during the training and testing calculation was judged using the root mean square (RMS), determination coefficient (R^2), and the mean absolute percentage error ($|\%ER|$) outlined in equations (2)–(4).

$$RMS = \sqrt{\frac{\sum P_{read} - P_{prid}}{N \sum CMP_{read}^{2}}} \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m} \left(P_{prid} - P_{prid}\right)^2}{\sum_{i=11}^{m} \left(P_{prid} - \overline{P}_{prid}\right)^2} \tag{3}$$

$$|\%RE| = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_{prid} - P_{prid}}{P_{prid}} \right| \times 100 \tag{4}$$

where P_{read} is the number of patients infected or dead at any time, P_{prid} is the number of predicted patients infected or dead from the ANN algorithm, and N, the number of data. The acceptable RMS limit for the testing data was set in the range 10^{-4} and 10^{-2}.

Results

Statistical data

Table 1 presents the number of infected/death cases in Qatar, Spain, and Italy as of July 16, 2021. It was noticed that the number of infected cases in Qatar is higher than in Spain and Italy due to the high spread rate of COVID-19. However, the mortality cases are much lower than those obtained in Spain and Italy. The reported infected cases per million of populations (PM_{IN}) were very high for Spain (80073) and Qatar (78365) compared to Italy (70342). On the other hand, Qatar showed a lower number of reported death cases per million of populations (~206) compared with Spain (~1723) and Italy (~2105). The reported values suggest that the population density follow the right procedure of social distance and precautions and the responsibilities of the individual have a major contribution to the pandemic evolution. The highest mortality rate per reported infected people in the European countries, 2.992% for Spain and 2.15% for Italy is one order of magnitude of Qatar (0.263), suggesting that the population intensity has a direct effect on the COVID-19 breakout between the two studied continents.

The ANN model

The training of the collected data of infection and death cases was carried out using a wide variety of TFs combinations (sigmoid (S), sigmoid (S), hyperbolic tangent (HT), and hyperbolic secant (HS) and different iterations as presented in Table 2. It should be indicated that the training was carried out with the data until July 2020. Results revealed that the most appropriate ANN architecture that predicts the number of infected/death cases was achieved by the combination of 5-4-4-2 and 5-5-5-2) using HT, S, and S/HT, S, and S as TFs in the HL-1, HL-2, and OL, respectively. The statistical analysis showed that the RMS and maximum percentage deviation error ($\%Max_{div,\ erro}$) for the infection data were 0.36% and 45.75%, with no more than 0.23% of the manipulated data fell within $\%Max_{div,\ erro}$ of ±10% error. Death cases showed RMS data at 0.40% and 22% of the manipulated data fell within $\%Max_{div,\ erro}$ of ±10% error. Previous studies showed that the mathematical modeling of the spread of the pandemic is based on the Gompertz model, which belongs to the family of the Sigmoid curve [61]. The combinations used in this study, although belonging to Sigmoid curve, support and refine the prediction of the ANN architecture [62]. The result showed that a single exponential model was not very adequate for the description of virus outbreak. On the other hand, the double exponential Gompertz model can accurately describe the biological growth. Indeed, the Sigmoid TF with double exponential expression has been used to model human mortality, bacterial growth curves, population growth [63], growth of animal fetuses [64], growth of chickens [65] and weight growth of fish [66], including anaerobic digestion kinetic [53]. The developed ANN architecture combines two families of TF to predict the growth and development of the COVID-19. This combination of S and HT generated SOF function that follows the instantaneous disease growth as a function of time, presenting a point of inflection where the growth curve transfers from concave to convex. Similarly, the training

Table 1
Summary of country COVID-19 data as June 16th, 2021.

Country	Population (10^6)	# of tests	Population % aged greater than 65	aged ≥ 70_	CVD death rate	Life expectancy	Infected (2021–6-16)	Dead (2020–06-16)	%Fatality rates (CFR)	Infected /M_{IN}	Dead /M_{IN}	DP
Qatar	2.88	2,084,963	1.31	0.62	176.69	80.23	219,887	579	0.263	76349.7	201.0	227.32
Italy	59.64	68,684,085	23.02	16.24	113.15	83.51	4,245,779	127,038	2.992	71190.1	2130.1	205.86
Spain	47.33	44,966,367	19.44	13.80	99.40	83.56	3,741,767	80,517	2.152	79057.0	1701.2	93.11

M_{IN}: Millions of Population.
CVD: Cardular vascular disease.
DP: Population density.

Table 2
Summary of the iterative tests during the training and testing of the ANN.

Test number	Infection							Death					
	# of Iterations	ANN structure	TF	$RMS_{training}$	% $Max_{div,erro}$	# ID with % $Max_{div,erro} > 15$	# ID with % $Max_{div,erro} > 10$	# of Iterations	ANN structure	TF	$RMS_{testing}$	% $Max_{div,erro}$	# ID with % dev > 15	# ID with % $Max_{div,erro} > 10$
1	6034	5-2-2-1	S-S-S	1.22	45.75	15.79	26.93	4827	5-2-2-3	S-S-S	1.34	46.03	14.79	28.47
2	6442	5-2-2-2	S-G-S	1.24	48.76	15.56	27.03	5153	5-2-2-3	S-G-S	1.36	44.67	14.43	27.97
3	6705	5-2-4-3	S-HS-S	1.17	40.38	10.66	24.41	5364	5-2-4-3	S-HS-S	1.31	40.82	9.55	25.70
4	7011	5-2-4-1	S-HT-S	0.91	45.83	11.69	23.30	5609	5-2-4-3	S-HT-S	0.99	45.64	14.43	23.87
5	7158	5–3-2–2	G-HS-HT	0.83	39.90	3.69	13.64	5726	5–3-2–3	G-HS-HT	0.92	31.62	2.97	14.24
6	8059	5–3-2–3	G-S-HT	0.86	40.90	5.49	14.48	6447	5–3-2–3	G-S-HT	1.01	32.67	4.88	15.77
7	8403	5-3-4-3	HT-HT-S	0.61	20.69	3.22	7.69	6722	5-3-4-3	HT-HT-S	0.67	29.17	2.76	8.03
8	8767	5-3-5-3	S-G-HT	0.55	28.97	2.32	5.75	7014	5-3-5-3	S-G-HT	0.62	31.02	2.12	5.77
9	9702	5-3-4-3	S-HT-HS	0.59	25.28	2.07	6.31	7762	5-3-4-3	S-HT-HS	0.67	22.67	1.63	6.50
10	10,098	5-4-3-3	G-HT-HS	0.50	28.70	2.3	3.87	8078	5-4-3-3	G-HT-HS	0.58	29.31	1.84	3.87
11	12,649	5-4-6-3	HT-S-G	0.36	23.09	0.23	1.62	10,119	5-4-6-3	HT-S-G	0.40	20.49	0.15	1.53
12	13,077	5-4-6-3	HT-HT-G	0.48	13.94	0.15	1.17	10,461	5-4-6-3	HT-HT-G	0.53	17.33	0.21	1.24
13	14,157	5-4-4-1	HT-S-G	0.40	16.12	0.15	1.14	11,326	5-5-3-3	HT-S-G	0.45	18.80	0.28	0.95
14	16,682	5-5-4-2	HT-S-HS	0.39	23.96	0.11	0.84	13,345	5-5-4-3	HT-S-HS	0.42	22.07	0.15	0.87
15	16,812	5-4-4-2	HT-S-S	0.36	45.75	0.08	0.23	13,450	5-5-5-2	HT-S-S	0.40	12.00	0.11	0.22

S: Sigmoid, **HS:** Hyperbolic secant, **HT:** Hyperbolic tangent, **G:** Gaussian.

tests with different ICD achieved a maximum RMS and $\%Max_{div,erro}$ of 0.55% and 5.1% for infection cases, and 0.65% and 3.1% for death cases, respectively. Following the results in Table 2, it was observed that ANN architecture with one IL, two HLs, and one OL has an excellent tendency for predicting the number of infected/dead cases in the studied countries with high accuracy.

The data in Table 2 also shows that the RMS decreases by increasing the number of iterations to 14,157 and 10,461 when infected and death cases were reached. Thereafter, the error values remain almost constant, suggesting that the number of trials is appropriate to predict the expected infected/death cases. Statistical analysis of the predicted versus collected data showed high agreements in the numbers with a maximum absolute error of 0.34% and 0.25% for infected and death cases. The plot of the residual between predicted versus reported cases generated points scattered around horizontal zero-line, confirming high prediction accuracy between the calculated and reported data. Up to 98.1% and 98.2% of the data points were within ±10% deviation of the mean values, respectively. This confirms that the selected TFs and iteration limits applied to the ANN are accurate for the prediction of the number of cases. The statistical manipulation of data showed that the R^2 and $|\%ER|$ during the training stage were 0.993% and 0.076%, respectively. Equally, these values were 0.995% and 0.189% for the testing stage, respectively. The frequency count versus relative error followed a Gaussian distribution with relative error in the ranges −27.6% to 32.5% and −28.8% to 31.9% for training and testing stages, respectively. The plot of residuals between the reported and predicted number of cases showed scattered points around the horizontal zero-lines, with an $|\%ER|$ in the range of −0.5 to 1.0 for training, and −1.0 to 1.0 for testing data, suggesting a very small deviation from the reported data.

Prediction of the infected cases

The developed ANN architecture was used to calculate the number of infected/death cases in Qatar, Spain, and Italy. Table 3 shows the predicted versus reported cases of COVID-19 during the period of January 2020 to June 2021. As the initial date of the breakout of COVID-19 in the three examined countries was different, the prediction calculations were based on the real date where at least three confirmed cases were reported. The analysis of variance (ANOVA) indicated that the developed ANN architecture accurately predicts the number of infected/death cases in the three countries. The student T-test ($P = 0.05$) disclosed an insignificant difference between the predicted and reported number of

Table 3
Summary of predicted versus reported cases of COVID-19 infections/death till June 2021.

Country		Infected						Death				
	Pandemic spread rate	Recorded cases PM_{IN}	Forecast Cases PM_{IN}	RMS	$\|\%ER\|$	R^2	Fatality rates PM_{IN}	Recorded cases PM_{IN}	Forecast Death PM_{IN}	RMS	$\|\%ER\|$	R^2
Spain	0.081	79056.98289	77,476	4.24	1.63	0.96	2.2	1701	1667	0.52	0.19	0.97
Italy	0.059	71190.12408	70,478	3.50	1.23	0.94	3.0	2130	2109	0.50	0.18	0.94
Qatar	0.085	76349.65278	76,044	4.31	1.65	0.98	0.3	201	200	0.53	0.19	0.99

cases until June 16, 2021. Calculations showed that up to 98.5% of the infected cases and 95.7% of death cases were within the ±2.5% and ±3.5% of the maximum deviation of the reported cases ($\%Max_{div,\ erro}$). Table 3 also shows that the RMS, $|\%ER|$ and R^2 of the forecasted infected cases were ≤4.31, ≤1.65 and ≥ 0.94 for the three countries. The calculated residuals between recorded and forecasted cases were in the range ±0.75 and ±0.750, respectively. As indicated before, having the residual scattered around the horizontal zero-line suggests high prediction accuracy of the data. Although the Pandemic spread rate and the number of reported cases were the highest in Qatar compared with Spain

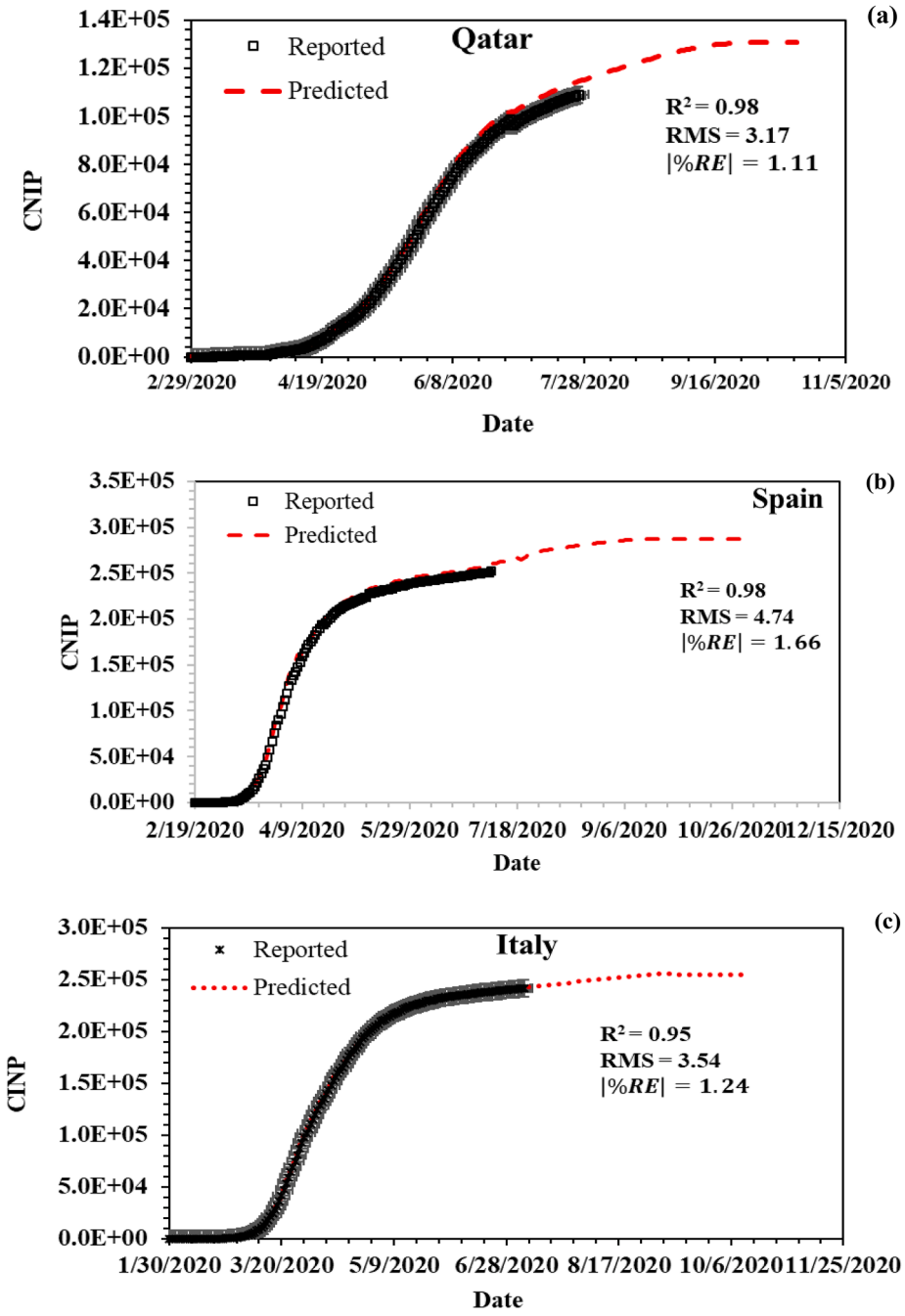

Fig. 2. Reported and predicted cumulative number of infected people (CNIP) (a) Qatar,(b) Spain and (c) Italy.

and Italy, the low mortality number suggest that the population density, infected people age, social distances precautions, weather conditions, and the responsibilities of individual have a major impact towards the critical pandemic evolution.

Fig. 2 presents the reported and predicted cumulative number of infected people (CNIP) for Qatar, Spain, and Italy during the training period. The CNIP shows that Spain and Italy are entering the stabilization region of the pandemic curve with very little increase in the reported cases, while Qatar is still within the exponential spread range. The slopes of the exponential spread curve during the exponential outbreak phase were 1311, 7065 and 3794 for Qatar, Spain, and Italy respectively. However, the slope of the pandemic curve during the last 5 days of this study show slopes of 0.05, 0.065 and 0.36 confirming that Qatar is still within the exponential spread range. The reported and predicted new daily infected cases (NDICs) by country are presented in Fig. 3. It can be seen that the ANN model follows the same trend of the real reported cases. Although, the model presents a forecast for the expected decrease in the NDICs for the coming period. Figs. 2 and 3 show that the ANN prediction follows the reported case evaluation and can forecast the future number of cases with factors of $R^2 \geq 0.96$, RMS $\geq$ 0.055 and $Ab_{RE} \leq 1.66$ for all countries. The results in Fig. 3 show that the NDICs in Qatar during the exponential outbreak phase are lower than in Spain and Italy, mainly due to the time elapsed for the latter countries to take measures to stop the pandemic, the differences between the climate conditions, and population density differences between the European countries and Qatar. Results from other countries such as South Korea and China present very low daily infected cases due to the strict measure confinements and monitoring strategies applied by these countries at the time of discovering the breakout, which has managed to paralyze the pandemic rapidly.

The developed ANN model was used to predict the evolution of the cases after July 6, 2021 and verify if the end of a pandemic can be determined based on the recorded data and the developed model. The NDICs have been calculated in prediction for the date that the pandemic disappears and the number of infections reaches zero. According to the ANN model with the corresponding coefficients of each country, the end of the pandemic situation where the number of cases will be stable and $\leq$200 is predicted to be October for Qatar, mid-September for Spain, and early September for Italy. However, if the precautions are released and/or curfew and social distancing regulations are relaxed, a second wave can start and the number of cases will be increase.

Mortality forecast

The ANN architecture was also used to model and calculate the number of death cases in the three studied countries from the initial date of the pandemic until the number of cases reached zero. In this case, similar trends were observed between Spain and Italy but different than Qatar, with Qatar possessing the lowest expected deaths. Table 3 includes the reported mortality numbers per million of the population (MN-M_{IN}) for Qatar, Spain, and Italy until June 16, 2021. The highest MN-M_{IN} was observed in Italy at 2130 followed by Spain at 1701, with no more than 200 deaths in Qatar. The percentage mortality of the infected population was 2.2% for Spain, 3.0% for Italy, and no more than 0.30% for Qatar. The MN-M_{IN} obtained by the ANN showed a very high correlation adjustment coefficient R^2, with an almost perfect adaptation to the real values. Fig. 4 shows the actual and predicted values of MN-M_{IN} for Qatar, Spain, and Italy. Based on the MN-M_{IN} it was observed that all of the studied countries are not yet at the stationary or decay phase. In the case of Qatar, the MN-M_{IN} is still within the exponential phase but approaching the flat zone with a very low growth rate coefficient ($\leq$0.45 day^{-1}). Spain and Italy show that the MN-M_{IN} is not decreasing yet and the approximate date of no deaths from the pandemic requires at least 3 to 4 months. The shape of the curve of the daily mortality numbers per million of the population (DMN-M_{IN}) for Qatar shows stabilization, while it is not decreasing or stabilized yet. Again, the correlation factors R^2 for all countries are very high, from the predicted DMN-M_{IN} concerning the real values.

Discussion

The developed ANN model showed promising outcomes in terms of predicting the number of infected/death cases without requiring the assumptions needed for the epidemiological models. The ANN as an alternative to epidemiological models showed an accurate prediction of the COVID-19 breakout in three different countries. It is also worth mentioning that ANN structure can further be useful in dealing with the challenges faced by other models used to predict the breakout of COVID-19. For example, different world meter incorrectly considered the number of cases reported as the number of people tested. Also, determining the number of infected people is a challenging matter, as many people who might be infectious may not turn up for testing especially if they do not show evident symptoms. Considering this data issue, it is extremely difficult to use physical models satisfactorily. Although we stress the necessity for credible data to be used in the ANN for better prediction, it is understood that self-reporting of symptoms is always prejudiced. Therefore, with the growth of the COVID-19 pandemic, it is essential to continuously collect and share robust data with the scientific community and public organizations. At the same time, it is highly recommended to consider the contribution of other factors of the pandemic breakout to be added into future models. Since the estimated and predicted data are determined by the description of the event and the data gathered, data collection and definition must be maintained in real time.

The ANN architecture can be effectively used to verify growth in the number of infected cases and to analyze the growth and development of unexpected pandemics and mortality. Specifically, the protocol shows how the ANN methodology can incorporate mathematical modeling and different combinations of TFs to anticipate the number of cases of COVID-19. The simulation model presents good results and excellent prognoses, with the correlation index $\geq$ 0.99 in all cases. The model predicted both infected and mortality cases for the three studied countries and can be extended to other countries. The study was conducted on two different continents (Europe and the Middle East) and it was observed that there are different trends in the breakout of COVID-19 that can be related to the population density, climate conditions and age of the infected people.

The ANN model used short-term data to predict the long-term spread of the pandemic. The ANN was also valid in determining the rate of mortality and infections of the pandemic with respect to infected individuals, who were detected with symptoms. The FFBPNN tuning algorithm used in this study eliminates the limitations of input data uncertainty that generates problems for the models based on physical relationships. The FFBPNN tuning algorithm also reduces the dependency of the number of evolution data to achieve accurate estimates, which is a large problem in different growth models including the Gompertz function. Besides, the structure of the ANN can be manipulated to incorporate the external pandemic containment factors applied by each country and determine their effect on the virus breakout. A structure such as this will assist in anticipating how each containment factor will affect the disease growth patterns (decreasing or increasing) and determine the correct measures to be used to contain cases and stop the spread of the virus.

The obtained results demonstrated the use of ANN algorithm to describe the pandemic growth of COVID-19 in the number of infections and deaths, outdo dynamic progression rate, predicting the peak of the trend change and anticipating the end of the outbreak. The algorithm is adaptable to different countries with different population density, climate conditions, and socio-political circumstances. The developed ANN algorithm generated COVID-19 outbreak curves that are very close to the real reported dates, attaining future forecasts for both infections and deaths with high correlation coefficients ($\geq$0.95) with respect to

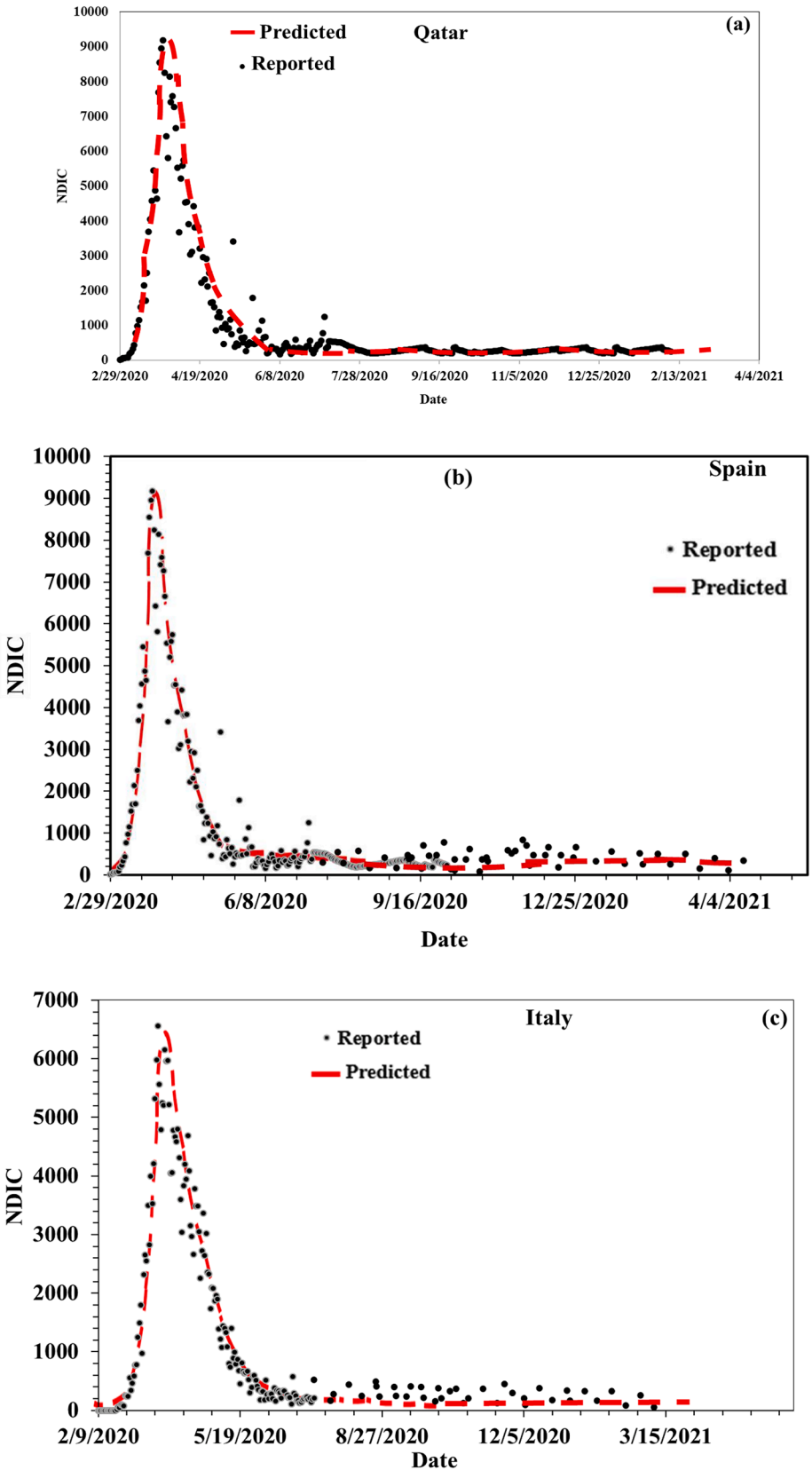

Fig. 3. Reported versus predicted daily infected cases (NDICs) (a) Qatar, (b) Spain and (c) Italy.

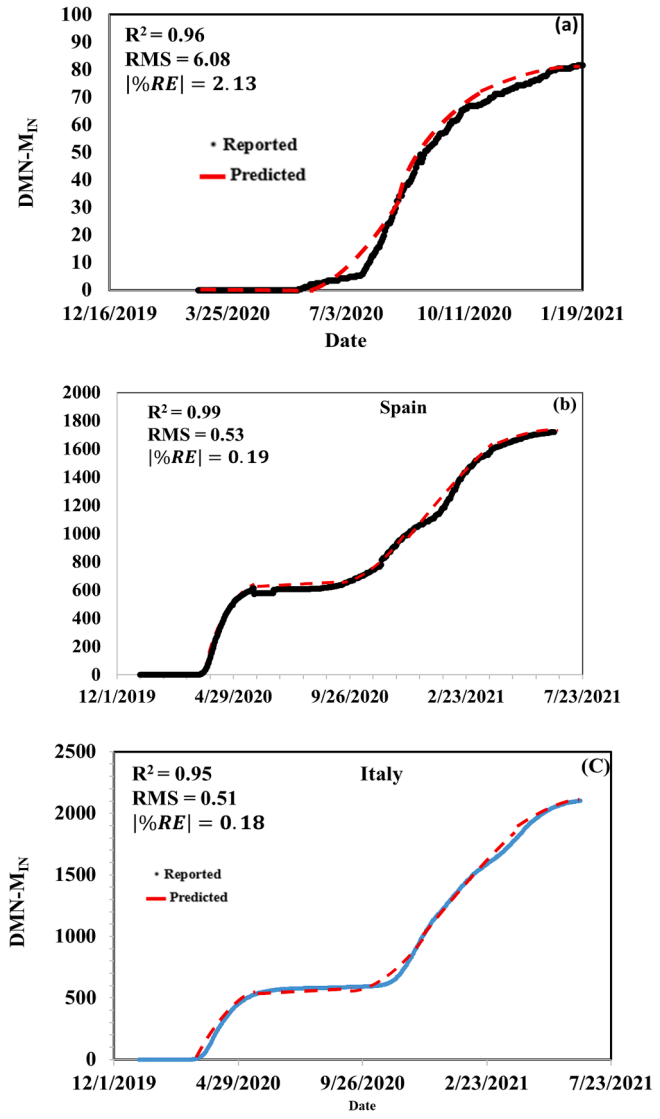

Fig. 4. The reported and predicted mortality numbers per million of population (MN-M_{IN}) cases (a) Qatar,(b) Spain and (c) Italy.

reported cases in all of the countries studied. Therefore, the ANN algorithm can be used as an appropriate procedure to investigate the growth and progress of unexpected pandemics with sudden and general outbreaks, and to also further describe epidemiological stability indices based on data collected from different countries that have already reached the peak of growth.

Considering that the studied countries have treated the outbreak of the pandemic in different ways and with different actions, the obtained results show that Qatar's strategy has been more efficient in controlling the number of death cases while recording a high number of infected people compared to European countries, where the death rates are high compared to the number of DMN-M_{IN}. It was observed that the treatment strategy is much more efficient in these European countries. Our results show that the procedure applied in Qatar, which was based on the systematic control of all those infected, their monitoring, control and isolation is very effective in decreasing the number of critical cases and reducing the rate of mortality. The procedure combats this exponential invasion of the COVID-19 by isolating the groups that have been in contact with the virus, the use of preventive protections for the main contagion through masks, and control through massive population tests. In contrast, the European system with partial confinement and spot testing has been very ineffective in avoiding or mitigating pandemic damage causing a high mortality rate. The ANN algorithm proposed in this study can be extended to other affected countries. With a continuous tuning algorithm, the ANN can be periodically updated during the outbreak to adopt any change in the pandemic conditions. The obtained results could be extended for future studies related to the evaluation of the impact of the pandemic on the economy, the ecosystem, and renewable energies, among others.

In comparison with other models, the ANN model provides an accurate prediction of the infection and death cases without the need for previous physical correlation, nor the assumptions required by the epidemiological models. In this regard, the model is considered an easy tool for the prediction of different diseases. Different studies were used in what is called a compartment model to quantitatively estimate the impact of interventions on the pandemic [67], most of these models considered population perspective and considered either deterministic or stochastic models (e.g. SEIR/SLIR, SIRD)[68,69]. Models based on Bayesian method [70,71], agent-based model and generalized growth model[72] were used for COVID-19 prediction. Based on the available results the ANN is considered within the top five models for pandemic prediction. The focuses of all the epidemic models were to determine the time for the infective pandemic and the estimation of key time periods of COVID-19 infection, the short and long-term prediction of the rate of spread as well as assessment of the influence of public health interventions

Conclusions

Due to the high level of uncertainty and lack of crucial data, analytical epidemiological models have shown low accuracy for long-term pandemic prediction. This work presents for the first time the development of ANN algorithms to describe the pandemic growth of COVID-19, both in the number of infections and deaths, outdo dynamic progression rates, predicting the peak of the trend change and anticipating the end of the breakout. The algorithm is adaptable to different countries with different population density, climate conditions and socio-political circumstances. The developed ANN algorithm generated COVID-19 outbreak curves that are very close to the real reported dates, attaining future forecasts for both infections and deaths with high correlation coefficients ($\geq$0.99) with respect to reported cases in all studied countries. The ANN architecture established using feed-forward back-propagation algorithms provided good estimates of the infected/death COVID-19 cases in Qatar, Spain, and Italy. The ANN predictions were acceptable, despite the limited amount of data. The results showed the high generalization ability of the ANN for long-term predictions of the pandemic outbreak of COVID-19. The trends of the breakout are different from country to country due to the highly complex nature of the COVID-19 outbreak and differences in nation-to-nation controls and precautions. The present work provides an initial benchmark to demonstrate the potential of ANN for future research.

Declaration of Competing Interest

The author of this work confirm the following:

1. The data used in the present study was obtained from daily reports issued by the Ministry of Health in Qatar, which is publicly available through the ministry website.
2. All methods used in this study were carried out in accordance with relevant guidelines and regulations.
3. The experimental protocols used in this study were developed solely by the author, who is working at the department of chemical engineering Qatar University.
4. The author confirms that there is no information taken from a third party persons.

References

[1] Chan J-F-W, Yuan S, Kok K-H, To K-K-W, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 2020;395(10223):514–23.
[2] Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). Int J Surgery 2020.
[3] Lu J, Gu J, Li K, Xu C, Su W, Lai Z, et al. COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020. Emerg Infect Dis 2020;26(7):1628.
[4] Organization WH. WHO statement regarding cluster of pneumonia cases in Wuhan. China. Beijing: WHO 2020;9.
[5] She J, Jiang J, Ye L, Hu L, Bai C, Song Y. 2019 novel coronavirus of pneumonia in Wuhan, China: emerging attack and management strategies. Clinical Transl Med 2020;9(1):1–7.
[6] Zhang Y-Z, Holmes EC. A genomic perspective on the origin and emergence of SARS-CoV-2. Cell 2020.
[7] Gralinski LE, Menachery VD. Return of the Coronavirus: 2019-nCoV. Viruses 2020;12(2):135.
[8] Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Hu Y, Song Z-G, Tao Z-W, Tian J-H, Pei Y-Y, Complete genome characterisation of a novel coronavirus associated with severe human respiratory disease in Wuhan, China. bioRxiv (2020).
[9] Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. BioRxiv 2020.
[10] Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 2020;395(10224):565–74.
[11] Mizumoto, K., Kagaya, K., Chowell, G.: Early epidemiological assessment of the transmission potential and virulence of 2019 Novel Coronavirus in Wuhan City: China, 2019-2020. medRxiv (2020).
[12] Mizumoto K, Kagaya K, Chowell G, Early epidemiological assessment of the transmission potential and virulence of coronavirus disease 2019 (COVID-19) in Wuhan City: China, January-February, 2020. medRxiv (2020).
[13] Sanche S, Lin YT, Xu C, Romero-Severson, E., Hengartner, N.W., Ke, R.: The novel coronavirus, 2019-nCoV, is highly contagious and more infectious than initially estimated. arXiv preprint arXiv:2002.03268 (2020).
[14] Khan M, Adil SF, Alkhathlan HZ, Tahir MN, Saif S, Khan M, et al. COVID-19: A Global Challenge with Old History, Epidemiology and Progress So Far. Molecules 2021;26(1):39.
[15] Hindson J. COVID-19: faecal–oral transmission? Nature Rev Gastroenterology Hepatology 2020;17(5):259.
[16] Ghinai I, McPherson TD, Hunter JC, Kirking HL, Christiansen D, Joshi K, et al. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA. Lancet 2020.
[17] Cheng Z, Shan J, Novel coronavirus: where we are and what we know. Infection [Internet]. 2020 Apr [cited 2020 Apr 30]; 48 (2): 155-63. In. (2019).
[18] Sahin AR, Erdogan A, Agaoglu PM, Dineri Y, Cakirci AY, Senel ME, et al. 2019 novel coronavirus (COVID-19) outbreak: a review of the current literature. EJMO 2020;4(1):1–7.
[19] Kampf G, Todt D, Pfaender S, Steinmann E. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. J Hosp Infect 2020;104(3):246–51.
[20] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020.
[21] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med 2020;172(9):577–82.
[22] Shao Y, Liu G-H, Wang Y, Zhang Y, Wang H, Qi L, et al. Sludge characteristics, system performance and microbial kinetics of ultra-short-SRT activated sludge processes. Environ Int 2020;143:105973.
[23] Patel A, Jernigan DB. Initial public health response and interim clinical guidance for the 2019 novel coronavirus outbreak—United States, December 31, 2019–February 4, 2020. Morb Mortal Wkly Rep 2020;69(5):140.
[24] Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N Engl J Med 2020;382(10):970–1.
[25] Organization, W.H.: Novel Coronavirus (2019-nCoV): situation report, 3. (2020).
[26] Organization, W.H.: WHO statement regarding the outbreak of novel coronavirus (2019-nCoV), 2020. Available from: https://www.who.int/newsroom/ detail/30-01-2020-statement-on-the-\second meeting-of-the-international-health-regulations-(2005)-emergency-\committeeregarding-the-outbreak-of-novel-coronavirus-\(2019-ncov). (Accessed on 30 March 2020). (2020).
[27] Spain, G.o.: Royal Decree 463/2020, of March 14, declaring the state of alarm for the management of the health crisis situation caused by COVID-19. Official Bulletin of the State (67): 25390-25400. March 14, 2020. ISSN 0212-033X. (2020).
[28] Lee M, You M. Psychological and behavioral responses in South Korea during the early stages of coronavirus disease 2019 (COVID-19). Int J Environ Res Public Health 2020;17(9):2977.
[29] Deshwal VK. COVID 19: A comparative study of Asian, European, American continent. Int J Sci Res Enginee Dev 2020;3(2):436–40.
[30] Organization WH. Coronavirus Disease (COVID-19) Situation Report-111. In. 2020.
[31] Watkins RE, Eagleson S, Veenendaal B, Wright G, Plant AJ. Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. BMC Med Inf Decis Making 2008;8(1):1–11.
[32] Geidelberg L, Boyd O, Jorgensen D, Siveroni I, Nascimento FF, Johnson R, et al. X.: Genomic epidemiology of a densely sampled COVID-19 outbreak in China. Virus Evol 2021;7(1):veaa102.
[33] Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics—I. Bull Math Biol 1991;53(1–2):33–55.
[34] Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. Containing papers of a mathematical and physical character Proc R Soc London. Series A 1932;138(834):55–83.
[35] Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J. Clinical Med. 2020;9(2):462.
[36] Tang, B., Wang, X., Li, Q., Bragazzi, N.L., Tang, S., Xiao, Y., Wu, J.: Estimation of the Transmission Risk of 2019-nCov and Its Implication for Public Health Interventions (2019-nCov 的传播风险估计及其对公共卫生干预的意义). (2020).
[37] Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman J, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infectious Disease Modelling 2020;5:256–63.
[38] Kwok KO, Leung GM, Lam WY, Riley S. Using models to identify routes of nosocomial infection: a large hospital outbreak of SARS in Hong Kong. Proc R Soc B: Biological Sci 2007;274(1610):611–8.
[39] Andrews JR, Basu S. Transmission dynamics and control of cholera in Haiti: an epidemic model. The Lancet 2011;377(9773):1248–55.
[40] McMeekin T, Olley J, Ratkowsky D, Corkrey R, Ross T. Predictive microbiology theory and application: Is it all about rates? Food Control 2013;29(2):290–9.
[41] Viboud C, Simonsen L, Chowell G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. Epidemics 2016;15:27–37.
[42] Kamel DG, Gomma NH, Osman DM, Hassan A. Growth of Lactococcus lactis subsp lactis in Milk Under Control of Culture Acidity. Assiut J Agric Sci 2017;48(6):32–9.
[43] Conde-Gutiérrez R, Colorado D, Hernández-Bautista S. Comparison of an artificial neural network and Gompertz model for predicting the dynamics of deaths from COVID-19 in México. Nonlinear Dyn 2021;1–15.
[44] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 2020;395(10225):689–97.
[45] Yang C, Wang J. A mathematical model for the novel coronavirus epidemic in Wuhan, China. Math Biosci Eng 2020;17(3):2708–24.
[46] Zhao S, Musa SS, Lin Q, Ran J, Yang G, Wang W, et al. Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak. Journal of clinical medicine 2020;9(2):388.
[47] Nishiura H, Kobayashi T, Yang Y, Hayashi K, Miyama T, Kinoshita R, Linton NM, Jung S-M, Yuan B, Suzuki A. The rate of underascertainment of novel coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights. Multidisciplinary Digital Publishing Institute. 2020.
[48] Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. NPJ Digital Med 2021;4(1):1–5.
[49] Cristea V-M, Pop C, Agachi PS. Artificial Neural Networks Modelling of PID and Model Predictive Controlled Waste Water Treatment Plant Based on the Benchmark Simulation Model No. 1. In: Computer Aided Chemical Engineering. Elsevier; 2009. p. 1183–8.
[50] Almomani F. Prediction the performance of multistage moving bed biological process using artificial neural network (ANN). Sci Total Environ 2020. 140854.
[51] Sahinkaya E. Biotreatment of zinc-containing wastewater in a sulfidogenic CSTR: performance and artificial neural network (ANN) modelling studies. J Hazard Mater 2009;164(1):105–13.
[52] Prakash N, Manikandan S, Govindarajan L, Vijayagopal V. Prediction of biosorption efficiency for the removal of copper (II) using artificial neural networks. J Hazard Mater 2008;152(3):1268–75.
[53] Almomani F. Prediction of biogas production from chemically treated co-digested agricultural waste using artificial neural network. Fuel 2020;280:118573.
[54] Holubar P, Zani L, Hager M, Fröschl W, Radak Z, Braun R. Advanced controlling of anaerobic digestion by means of hierarchical neural networks. Water Res 2002;36(10):2582–8.
[55] Akbaş H, Bilgen B, Turhan AM. An integrated prediction and optimization model of biogas production system at a wastewater treatment facility. Bioresour Technol 2015;196:566–76.
[56] Dibaba OR, Lahiri SK, T'Jonck S, Dutta A. Experimental and artificial neural network modeling of a Upflow Anaerobic Contactor (UAC) for biogas production from Vinasse. Int J Chem Reactor Eng 2016;14(6):1241–54.
[57] Betiku E, Okunsolawo SS, Ajala SO, Odedele OS. Performance evaluation of artificial neural network coupled with generic algorithm and response surface methodology in modeling and optimization of biodiesel production process parameters from shea tree (Vitellaria paradoxa) nut butter. Renewable Energy 2015;76:408–17.
[58] Yan Z, He A, Hara S, Shikazono N. Modeling of solid oxide fuel cell (SOFC) electrodes from fabrication to operation: Microstructure optimization via artificial neural networks and multi-objective genetic algorithms. Energy Convers Manage 2019;198:111916.
[59] Hassen, E.B., Asmare, A.M.: Predictive performance modeling of Habesha brewery wastewater treatment plant using artificial neural network. (2019).
[60] ECDC: European Centre for Disease Prevention and Control (ECDC). COVID-19 - Situation update – worldwide. Stockholm: ECDC; 1 April 2020. Available from: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases (Accessed on 12 April 2020). (2020).
[61] Gompertz, B.: XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a

letter to Francis Baily, Esq. FRS &c. Philosophical transactions of the Royal Society of London(115), 513-583 (1825).
[62] Laird AK. Dynamics of tumour growth: comparison of growth rates and extrapolation of growth curve to one cell. Br J Cancer 1965;19(2):278.
[63] Román-Román P, Romero D, Torres-Ruiz F. A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data. J Theor Biol 2010;263(1):59–69.
[64] McCredie J. The rate of tumor growth in animals. Growth 1965;29:331–47.
[65] Rogers S, Pesti G, Marks H. Comparison of three nonlinear regression models for describing broiler growth curves. Growth 1987;51(2):229–39.
[66] Silliman R. Comparison between Gompertz and von Bertalanffy curves for expressing growth in weight of fishes. J Fisheries Board Canada 1969;26(1):161–5.
[67] Xiang Y, Jia Y, Chen L, Guo L, Shu B, Long E. COVID-19 epidemic prediction and the impact of public health interventions: A review of COVID-19 epidemic models. Infectious Disease Modelling 2021;6:324–42. https://doi.org/10.1016/j.idm.2021.01.001.
[68] Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020;368(6489):395–400.
[69] Fanelli D, Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos, Solitons Fractals 2020;134:109761.
[70] Kai D, Goldstein, G-P, Morgunov A, Nangalia V, Rotkirch A, Universal masking is urgent in the covid-19 pandemic: Seir and agent based models, empirical validation, policy recommendations. arXiv preprint arXiv:2004.13553 (2020).
[71] Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. Lancet Infect Dis 2020;20(7):793–802.
[72] Kraemer MU, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020;368(6490):493–7.