SHORT RESEARCH AND DISCUSSION ARTICLE

Prediction of the confirmed cases and deaths of global COVID-19 using artificial intelligence

Qingchun Guo[1,2] · Zhenfang He[1,3]

Abstract

The outbreak of coronavirus disease 2019 (COVID-19) has seriously affected the environment, ecology, economy, society, and human health. With the global epidemic dynamics becoming more and more serious, the prediction and analysis of the confirmed cases and deaths of COVID-19 has become an important task. We develop an artificial neural network (ANN) for modeling of the confirmed cases and deaths of COVID-19. The confirmed cases and deaths data are collected from January 20 to November 11, 2020 by the World Health Organization (WHO). By introducing root mean square error (RMSE), correlation coefficient (R), and mean absolute error (MAE), statistical indicators of the prediction model are verified and evaluated. The size of training and test confirmed cases and death base employed in the model is optimized. The best simulating performance with RMSE, R, and MAE is realized using the 7 past days' cases as input variables in the training and test dataset. And the estimated R are 0.9948 and 0.9683, respectively. Compared with different algorithms, experimental simulation shows that trainbr algorithm has better performance than other algorithms in reproducing the amount of the confirmed cases and deaths. This study shows that the ANN model is suitable for predicting the confirmed cases and deaths of COVID-19 in the future. Using the ANN model, we also predict the confirmed cases and deaths of COVID-19 from June 5, 2020 to November 11, 2020. During the predicting period, the R, RMSE, and MAE for new infected confirmed cases of COVID-19 are 0.9848, 17,554, and 12,229, respectively; the R, RMSE, and MAE for new confirmed deaths of COVID-19 are 0.8593, 631.8, and 463.7, respectively. The predicted confirmed cases and deaths of COVID-19 are very close to the actual confirmed cases and deaths. The results show that continuous and strict control measures should be taken to prevent the further spread of the epidemic.

Keywords SARS-CoV-2 · COVID-19 · Epidemic · Infected cases · Deaths · Artificial intelligence

Responsible Editor: Lotfi Aleya

✉ Qingchun Guo
guoqingchun@lcu.edu.cn

✉ Zhenfang He
hezhenfang@lcu.edu.cn

1   School of Environment and Planning, Liaocheng University, Liaocheng 252000, China

2   State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710061, China

3   State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

Introduction

In December 2019, an outbreak of coronavirus disease 2019 (COVID-19) appeared in Wuhan, which was caused by severe acute respiratory syndrome coronavirus 2 (SARSCoV-2) (Li et al. 2020a; Zhou et al. 2020; Zhu et al. 2020). The epidemic spread quickly, with over 51,251,907 cases and 1,270,930 deaths reported globally, and there are more than 216 countries or regions with cases (as of November 11, 2020, WHO). Common symptoms are fever, cough, and myalgia or fatigue (Huang et al. 2020). The proportion of black patients with COVID-19 in the USA is much higher than that of white patients (Price-Haywood et al. 2020). Human to human transmission plays a role in the disease (COVID-19) (Chan et al. 2020). Environmental monitoring of two hospitals in Wuhan reveals the airborne hot spots of SARS-CoV-2 RNA (Liu et al.

2020). And summer does not stop the spread of SARSCoV-2 (Baker et al. 2020).

Although researchers believe that novel coronavirus (SARSCoV-2) is the natural host of bat, the final origin of the virus is not yet fully established (Zhou et al. 2020). The intermediate host of SARSCoV-2 may be pangolins (Lam et al. 2020), cat (Halfmann et al. 2020), ferret (Shi et al. 2020), dog (Sit et al. 2020), and hamster (Sia et al. 2020).

It is important to understand the potential epidemic of SARSCoV-2 by assessing the infectivity of the unrecorded SARSCoV-2 infection. It is estimated that 86% of infections before travel restrictions in China on January 23, 2020 have not been recorded. Seventy-nine percent of recorded infections originate from unrecorded infections, resulting in rapid spread of SARSCoV-2 and difficult to block (Li et al. 2020b). Using the real-time mobility data and cases in Wuhan, the initial population mobility data can well explain the spatial distribution of cases. After the prevention and control measures were taken, the growth rate in most regions became negative, and the control measures greatly alleviated the spread of the disease (Kraemer et al. 2020). The global metapopulation disease transmission model is used to predict the influence of travel restrictions on the domestic and international spread of COVID-19. Measures to close Wuhan slowed the outbreak of COVID-19 in China and gave China 3–5 days to prepare for it. At the international level, it had more radiation effects, and by the middle of February, it had reduced nearly 80% of the imported cases (Chinazzi et al. 2020).

Most empirical models use hypothetical parameters, so they are not suitable for COVID-19 data. Using these models to predict future cases of COVID-19 may not be very accurate. There is still great uncertainty about the amount of new confirmed cases in the future. When the epidemic will be over is an urgent question for the public to know. Artificial intelligence (AI) has been proven of their feasibility in capturing nonlinear relationships. Artificial intelligence has been widely used in many fields (Cassimon et al. 2020; Guo et al. 2020; He et al. 2014), and AI can quickly diagnose COVID-19 (Zhang et al. 2020). To overcome limitations of the epidemiological model, we develop artificial intelligence (AI) for real-time predicting of the new confirmed cases of COVID-19 all over the world. Our goal is to provide better estimation methods to assist medical and government agencies in effective response and timely adjustment in the event of the epidemic. We hope our research results will be able to inform the world of the future trend of the epidemic.

Materials and method

Data collection

Data on the daily new infected confirmed cases of COVID-19 and new confirmed deaths of COVID-19 around the world

from January 21, 2020 to November 11, 2020 are obtained from the World Health Organization (WHO). Data of COVID-19 worldwide are divided into training, testing, and predicting sequences. Training data are from January 21 to May 22, 2020, testing data are from May 23 to June 4, 2020, and forecasting data are from June 5 to November 11, 2020.

Artificial intelligence

Artificial intelligence (AI) is used to predict the amount of new infected confirmed cases of COVID-19 and new confirmed deaths of COVID-19. With one hidden layer, the artificial neural network (ANN) that hires a backpropagation algorithm is constructed (Rumelhart et al. 1986). The architecture of ANN contains three layers (input layer, hidden layer, output layer) and a layer contains some nodes. $C(1) \dots C(n)$ is the data of daily new confirmed cases as the input variable, and $C(n+1)$ is the new confirmed cases predicted for $+1$ day (Fig. 1). C_i is the node input, z expresses the node output, and W_{ji} expresses the weight, where D expresses the node excitation threshold, and v and f express the basic and activation functions, respectively. A node assesses the weighted summation of the inputs as

$$v = \left(W_{ji} C_i \right) + D \tag{1}$$

The activation function appraises output by

$$z = f(v) \tag{2}$$

To determine the relationships between previous days' cases and the next day's (tomorrow) variations, the ANN model is used to estimate global patients with COVID-19. The input parameters are the past days' global patients, and the output variable produces predictions of global patients with COVID-19 for $+1$ day. The activation functions employed in the input layer, hidden layer, and output layer are hyperbolic tangent sigmoid function (tansig) and rectified linear unit (ReLU, or poslin). To avoid overfitting and validate
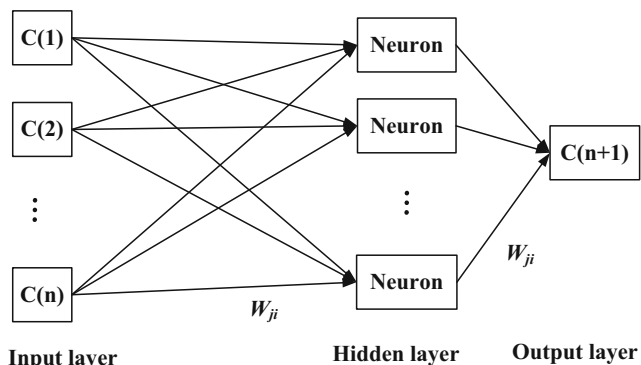

Fig. 1 The architecture of ANN predicting the COVID-19 epidemic

the reliability of the developed model, we use 90% of the
cases for training and 10% of the cases for testing.

Performance criteria

The performance of ANN is assessed employing three metrics
including: root mean square error (RMSE), correlation coef-
ficient (R), and mean absolute error (MAE). The R values are
employed to determine the model accuracy, and the root mean
square error (RMSE) values are employed to determine the
residuals between predictions and actual cases:

$$RMSE = \sqrt{\frac{\sum(A_j - G_j)^2}{L}} \tag{3}$$

$$R = \frac{\sum(A_j - \overline{A})(G_j - \overline{G})}{\sqrt{\sum(A_j - \overline{A})^2(G_j - \overline{G})^2}} \tag{4}$$

$$MAE = \frac{1}{L}\sum|A_j - G_j| \tag{5}$$

A_j denotes the actual amount of new confirmed cases of
COVID-19, G_j denotes the predicted amount of new con-
firmed cases of COVID-19, A is the mean of the actual cases,
and G is the mean of the predicted cases.

Results

The cycles of the global COVID-19 epidemic

The cycles of new infected confirmed cases of COVID-19 are
calculated using wavelet analysis (Fig. 2). The wavelet vari-
ance diagram can reflect the distribution of wave energy of time
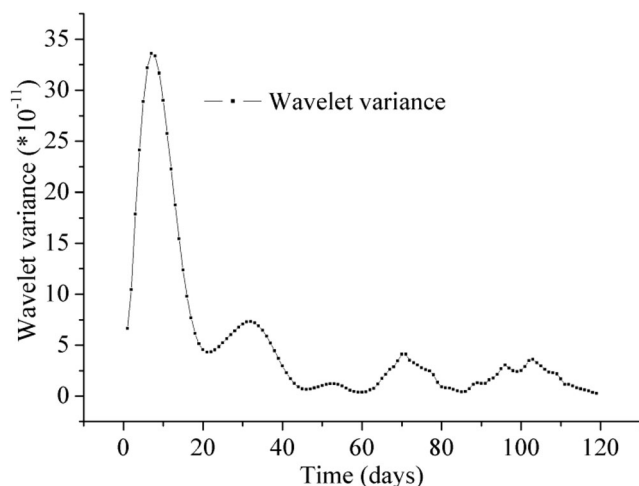series with the different scales. It can be used to determine the

Fig. 2 The cycles of new infected confirmed cases of COVID-19

main periods of new infected confirmed cases of COVID-19.
There are five obvious peaks in the wavelet variance diagram,
which correspond to the time scales of 7 days, 32 days, 71 days,
96 days and 103 days (Fig. 2). These are the cycles of new
infected confirmed cases of COVID-19. Among them, the
maximum peak value corresponds to the 7 days (time scale),
which means that the period oscillation of about 7 days (time
scale) is the strongest, which is the first main cycle; the second
peak corresponds to the 32 days (time scale), which is the
second main cycle; the third, fourth, and fifth peak value cor-
respond to 71 days, 96 days, and 103 days (time scale), respec-
tively, and they last for the third, fourth, and fifth main cycle in
turn. This shows that the fluctuation of the above five periods
controls its variation characteristics in the whole time domain.

Model development using the ANN

Many studies have shown theoretically that three-layered ANN
can describe any nonlinear mapping relation with precision (He
et al. 2014). A typical neural network consisting of three layers
was applied to forecast the global COVID-19 epidemic. The
number of neuron in input and hidden layer is decided by trial
and error. Figure 3 shows optimization of network topologies.
The performance of various numbers of days in the input and
hidden layer is compared (Tables 1 and 2). Tables 1 and 2 show
the variables and the R values of the proposed ANN model.
Tables 1 and 2 also show simulation of the amount of new
infected confirmed cases of COVID-19 during the training and
test period. Interestingly, using only the most recent 7 days
reproduces the best global patients with COVID-19 simulations.
Using more than 7 days entirely confuses the model and pro-
duces unexpected cases. The final model included 7 past days'
data. Seven variables were selected for the model input, and the
amount of nodes of the hidden layer is similarly 6. Finally,
network topologies (7-6-1 for ANN) are better than others.

Training algorithms of the ANN model are also chosen by
trial and error. Figure 4 shows optimization of training algo-
rithms for predicting the COVID-19 epidemic. The simulation
values are close to the actual values using trainbr in both
training period and testing period. trainbr is performed best
in predicting the COVID-19 epidemic. Table 3 shows the
performances of the training algorithms, revealing that the
trainbr algorithm has the best performance in simulating glob-
al patients with COVID-19. Table 3 also shows the simulation
performance using trainbr for the developed ANN model. The
simulation case is very close to the actual case. To avoid
overfitting, we conduct a test. The model has similar R values,
so there were no overfitting issues with the model. The RMSE
value for the ANN model using trainbr for the training dataset
is 3859.4, and that for the test dataset is 3102.9. The R for the
ANN using trainbr during the training and test period are
0.9948 and 0.9683, respectively. The MAE for the ANN using
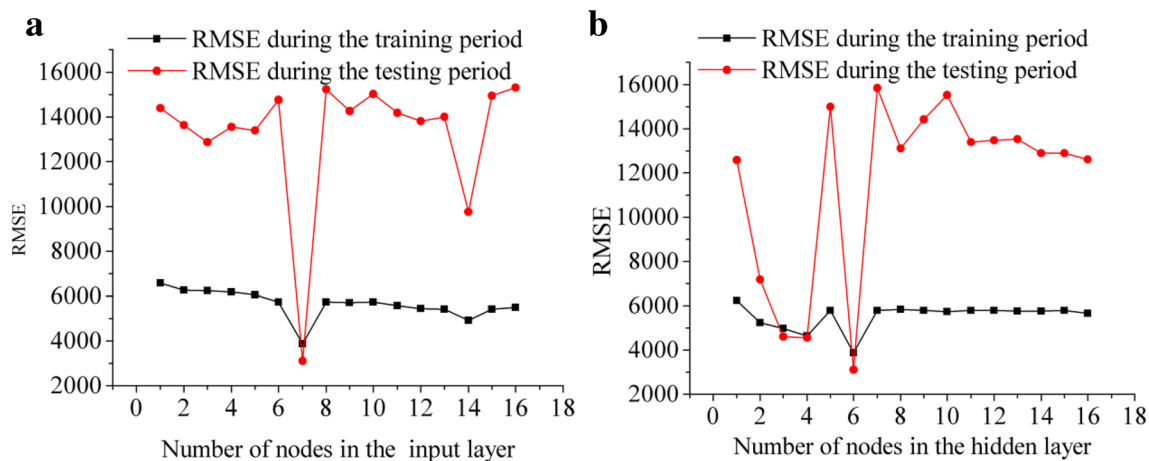trainbr are 2303.7 and 2090.6, respectively.

Figure 3 Optimization of network topologies for predicting the COVID-19 epidemic. **a** RMSE for different nodes in the input layer. **b** RMSE for different nodes in the hidden layer

Transfer functions of the ANN model are also chosen by trial and error. Table 4 shows transfer function (tansig-poslin) is better than others during training, testing, and predicting periods. Purelin and poslin is a linear transfer function and positive linear transfer function, respectively. Tansig and logsig is a hyperbolic tangent sigmoid transfer function and logarithmic sigmoid transfer function, respectively.

Prediction of the number of infected cases and deaths of COVID-19

In the forecast period, global infected cases and deaths of COVID-19 in the next day are predicted cumulatively using previous days' predicted infected cases and deaths. Table 5 shows the predicting performance using trainbr for the developed ANN model. For new infected confirmed cases of COVID-19 during the predicting period, the R, RMSE, and MAE are 0.9848, 17,554, and 12,229, respectively. For total infected confirmed cases of COVID-19 during the predicting period, the R, RMSE, and MAE are 0.9999, 453,720, and 324,000, respectively. For new infected deaths of COVID-19 during the predicting period, the R, RMSE, and MAE are 0.8593, 631.8, and 463.7, respectively. For total infected deaths of COVID-19 during the predicting period, the R, RMSE, and MAE are 0.9999, 4861.6, and 4213.4, respectively.

Table 1 Variables selected using the proposed ANN model

Variables	R		RMSE		MAE	
	Training	Testing	Training	Testing	Training	Testing
1	0.9853	0.1546	6585.6	14,394	3897.3	11,018
2	0.9868	0.2229	6253.5	13,643	3837.5	11,132
3	0.9867	0.2273	6228.8	12,866	3862	10,449
4	0.9872	0.1873	6178.8	13,552	4158.7	11,458
5	0.9875	0.2081	6061.4	13,390	3730.7	11,236
6	0.9891	0.2443	5730.7	14,773	3532.5	12,958
7	0.9948	0.9683	3859.4	3102.9	2303.7	2090.6
8	0.9887	0.6762	5728.6	15,227	3668.8	13,052
9	0.9892	0.6719	5701.8	14,251	3560.8	12,021
10	0.989	0.677	5748.6	15,019	3668.3	12,665
11	0.9895	0.6714	5591.4	14,184	3536.5	11,863
12	0.9899	0.6673	5452.8	13,825	3363.7	11,643
13	0.9901	0.6736	5410.4	14,007	3277.4	11,738
14	0.9916	0.8446	4918.8	9752	3143.8	7519
15	0.9899	0.7232	5416.1	14949	3441	12,610
16	0.9898	0.6652	5507.2	15,309	3570.3	12,884

Table 2 Comparison between different nodes in hidden layer

Nodes	R		RMSE		MAE	
	Training	Testing	Training	Testing	Training	Testing
1	0.9867	0.353	6242.7	12,579	4061.2	10,446
2	0.9903	0.8034	5246.3	7171.9	3373.7	6282.5
3	0.9913	0.9343	4986.2	4610.7	3038.5	3523.5
4	0.9925	0.934	4626.7	4557.7	2696.8	3612.5
5	0.9891	0.566	5788.6	15,006	3616	12,771
6	0.9948	0.9683	3859.4	3102.9	2303.7	2090.6
7	0.989	0.5573	5794.8	15,851	3697.7	13,501
8	0.9887	0.6104	5835.5	13,106	3703.4	10,976
9	0.9891	0.587	5796.1	14,413	3863.9	12,237
10	0.9894	0.6362	5743	15,529	3790.8	13,108
11	0.9891	0.6005	5780.5	13,390	3797	11,274
12	0.9891	0.5958	5790.7	13,479	3860.9	11,362
13	0.9891	0.6062	5750.2	13,537	3703.2	11,401
14	0.989	0.5786	5768.4	12,898	3669.9	10,832
15	0.9889	0.5794	5796.9	12,891	3662.6	10,820
16	0.9895	0.7329	5663.9	12,602	3554.3	10,433

Figure 4 Optimization of training algorithms for predicting the COVID-19 epidemic
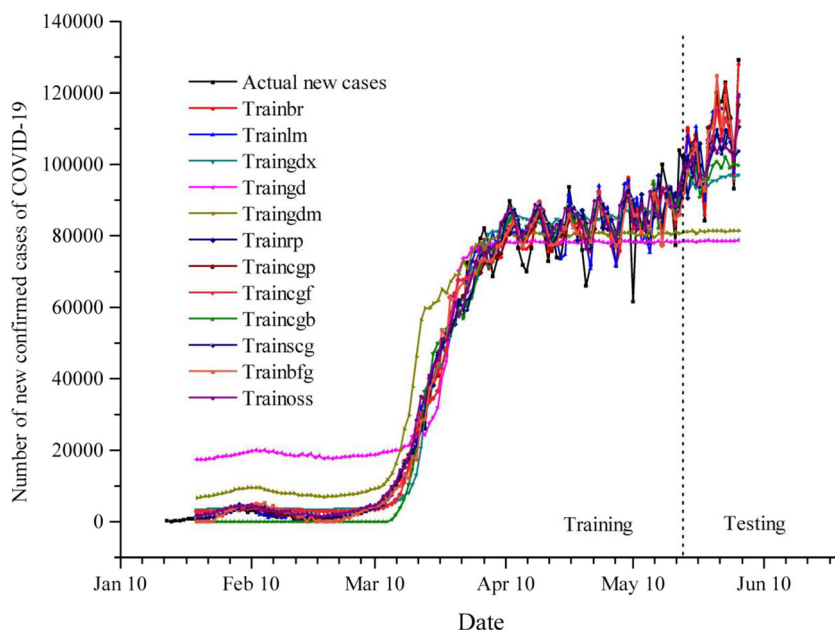

Figure 5 expresses predicted new infected confirmed cases of COVID-19 all over the world. The training and testing cases are from January 20 to June 4, 2020. We began to predict from June 5 to November 11, 2020. In the 5 months, the average amount of predicted new infected confirmed cases of COVID-19 is 271,761 every day in the world, and the average amount of actual infected cases is 271,528. Figure 6 expresses predicted total infected confirmed cases of COVID-19 worldwide. The total infected confirmed cases continue to grow during the predicting period. The actual total amount of cumulative infected confirmed cases is more than 51 million by November 11, 2020, and the forecasting total infected cases are similar to the actual infected cases.

Figure 7 expresses predicted new infected deaths of COVID-19 all over the world. During the predicting period, the average amount of predicted new infected deaths of COVID-19 is 5460 every day in the world, and the average amount of actual infected deaths is 5466. Figure 8 expresses predicted globally total infected deaths of COVID-19. The actual amount of total infected deaths is 1,270,930 by November 11, 2020, and the forecasting total infected deaths are 1,258,700. In summary, the predicting infected deaths are very close to the actual infected deaths.

Table 6 shows the predicting performance of new infected confirmed cases of COVID-19 in 10 countries. During the predicting period, the R, RMSE, and MAE for USA are 0.9696, 5139.1, and 4074.6, respectively; the R, RMSE, and

Table 3 Simulation performance and optimization of training algorithms for the ANN model.

Training functions	R		RMSE		MAE	
	Training	Testing	Training	Testing	Training	Testing
trainbr	0.9948	0.9683	3859.4	3102.9	2303.7	2090.6
trainlm	0.9926	0.9152	3859.4	5359.8	2588.4	4018.7
traingdx	0.9859	0.5529	6474.4	16,696	4512.7	14,399.2
traingd	0.9774	0.3153	12497	31,499	10,950	29,203.5
traingdm	0.9748	0.2415	9698.8	28,951	7763	26,445.3
trainrp	0.9877	0.5393	5929.9	10,751	3747.6	8208.7
traincgp	0.9903	0.842	5255.8	6721.8	3387.3	6227.2
traincgf	0.9898	0.7956	5408.6	8154.2	3632.3	6477.2
traincgb	0.9883	0.4287	6150.5	14,420	4644.4	12,517.3
trainscg	0.9897	0.7045	5512	10,617	3546.2	8860.9
trainbfg	0.9902	0.8085	5287.6	6974.6	3530.5	6120.7
trainoss	0.9929	0.7654	4817.3	9533.7	3143.9	8154.2

Table 4 Comparison between various transfer functions

Transfer function	R		RMSE		MAE	
	Training	Testing	Training	Testing	Training	Testing
tansig-purelin	0.9886	0.6121	5923.6	16,494	4011.5	2570.3
tansig-logsig	0.8519	0.2859	49,990	98,425	36,582	10,950
tansig-tansig	0.8987	0.3475	49,992	98,415	36,640	10,949
logsig-purelin	0.9893	0.5656	5666.5	14,566	3636.8	2388.9
logsig-tansig	0.794	0.3671	50,015	98,425	37,098	10,950
logsig-logsig	0.8577	0.4409	50,096	98,401	38,285	10,948
purelin-tansig	0.9185	0.2577	49,983	98,397	36,478	10,947
purelin-logsig	0.936	0.2941	50,026	98,397	37,482	10,947
purelin-purelin	0.9848	0.3944	6572.7	11,027	4491.6	2963.9
tansig-poslin	0.9948	0.9683	3859.4	3102.9	2303.7	2090.6
logsig-poslin	0.9885	0.5262	5873.7	14,344	4013.8	2365.3

MAE for Russian Federation are 0.9978, 311.9, and 205.3, respectively; the R value of Russian Federation is higher than that of other countries, and the RMSE and MAE value of Russian Federation is smaller than that of other countries. Figure 9 (a-j) expresses predicted new infected confirmed cases of COVID-19 in 10 countries. The second wave of COVID-19 outbreak occurred in 10 countries. The largest number of new cases in a single day is in the USA.

Compared with existing COVID prediction models

Different models for predicting COVID-19 infections are evaluated. The R, RMSE, and MAE for new confirmed cases of COVID-19 using our model are 0.9848, 17,554, and 12,229, respectively. The R^2, MSE, and MAPE, for COVID-19 infections using W-LSTM, are 0.93, 3.13E+05, and 39.29, respectively (Tuli et al. 2020b). In the LSTM model, they trained and tested LSTM network on Canadian COVID-19 infections dataset; the RMSE error is 34.83 with an accuracy of 93.4% for short-term predictions in Canada. Meanwhile, based on the testing/validation COVID-19 infections dataset, the RMSE error is about 45.70 with an accuracy of 92.67% for long-term predictions (Chimmula and Zhang 2020). The LSTM model predicted that new infections will peak on

February 4, resulting in 95,000 cases by the end of April in mainland China. Both the SEIR and LSTM-model predicted a peak of 4000 daily infection between February 4 and 7. The SEIR model also predicted several smaller peaks of new infections in mid to late February. The actual number of cumulative confirmed cases in mainland China on April 30, 2020 is 82,874 (Yang et al. 2020). A set of models using 9 different machine learning algorithms for predicting the rise in new cases, having an average accuracy of 87.9 ± 3.9%, was developed for 10 high population and high-density countries. The highest accuracy of 99.93% was achieved for Ethiopia using ARMA averaged over the next 5 days (Khakharia et al. 2020). Not every machine learning algorithm could give a very high accuracy for predicting the cases for each country. Machine learning models based on decision tree, random forest, logistic regression, and support vector machines are developed and show accuracies between 76.2 and 92.9% to predict early signs of infection containment (Kasilingam et al. 2020). Although our model has better performance when compared with other predicting models, it is unfortunate that cumulative confirmed cases are following increasing trend.

Many machine learning methods or artificial intelligence techniques have been employed to forecast the number of confirmed cases of COVID-19. However, there are lots of

Table 5 COVID-19 prediction of the amount of new infected confirmed cases, total infected confirmed cases, new infected deaths, total infected deaths in 2020

COVID-19	R			RMSE			MAE		
	Training	Testing	Predicting	Training	Testing	Predicting	Training	Testing	Predicting
New cases	0.9948	0.9683	0.9848	3859.4	3102.9	17,554.0	2303.7	2090.6	12,229.0
Total cases	1	0.9998	0.9999	16,586.0	14,694.0	453,720.0	13,903.0	1223.3	324,000.0
New deaths	0.979	0.6028	0.8593	567.2	828.8	631.8	320.0	58.7	463.7
Total deaths	0.9998	0.9989	0.9999	4125.2	1028.4	4861.6	2806.0	94.6	4213.4

Fig. 5 Prediction of the amount of new infected confirmed cases of COVID-19 in 2020

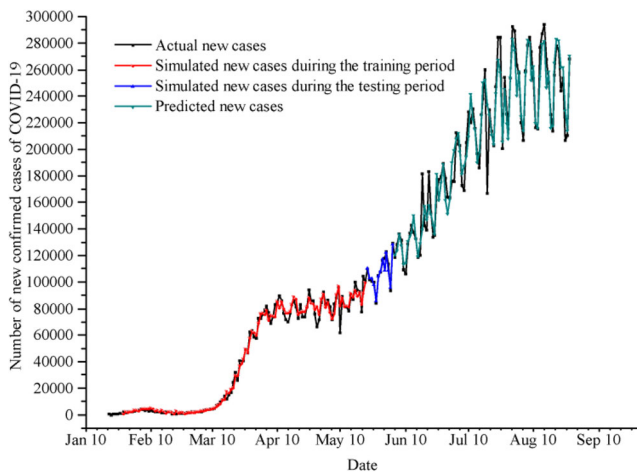**Fig. 7** Prediction of the amount of new infected deaths of COVID-19 in 2020

challenges for the accurate prediction by machine learning methods. It is difficult to cultivate accurate machine learning models with small datasets. Deep learning methods are successful because of big training data which is not available for prediction task of COVID-19 confirmed cases. It is difficult to select suitable architectures and parameters for deep learning neural networks with small datasets. It is disputed that lots of countries are not doing enough testing. Therefore, it is impossible to have correct confirmed cases in these countries. Using poor-quality datasets to train machine learning algorithms will lead to wrong conclusions (Ahmad et al. 2020).

Discussion

We predict the amount of the infected confirmed cases and deaths of COVID-19 for the next day using artificial intelligence. The accuracies of simulation of the infected confirmed cases and deaths of COVID-19 are high. So, the ANN can be
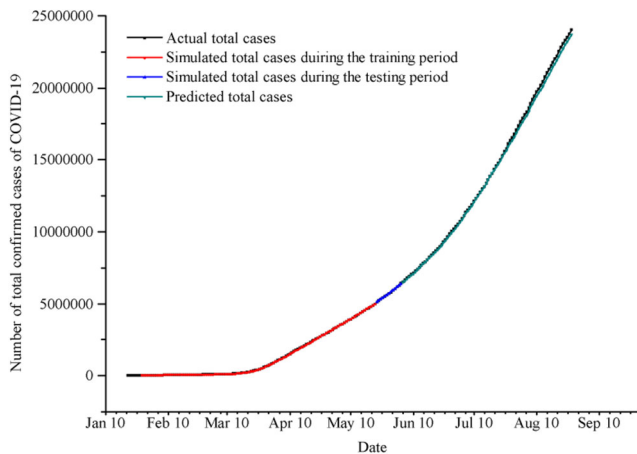
used for simulations. The results show that using trainbr algorithm has the best performance. The 7 input variables are selected to create an ANN model. We also can use the ANN with known parameters to predict the number of the virus epidemic outbreaks in the future.

The lowest RMSE is accomplished by using previous 7 days data in the training and trest stage. If more input parameters are used, the error gets higher until seven input parameters for + 1 day prediction. The RMSE goes up slightly as more input parameters are used in the prediction of + 1 day cases of COVID-19.

We provide a simple AI model for policy makers and researchers to understand the infected confirmed cases and deaths of the COVID-19 in the next 3 months based on specific estimates of global past historical data. The actual data of the confirmed case of COVID-19 in progress is well matched with that of AI, which strongly shows that it is suitable for simulating the epidemic caused by SARS-CoV-2. These results help authorities to control the COVID-19 epidemic.

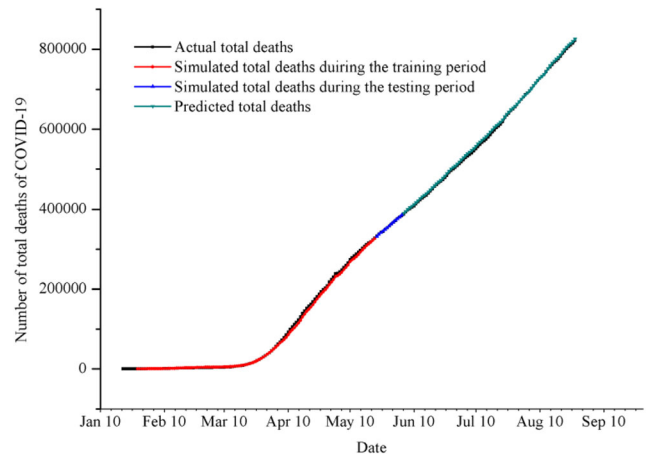**Fig. 6** Prediction of the amount of total infected confirmed cases of COVID-19 in 2020

Fig. 8 Prediction of the amount of total infected deaths of COVID-19 in 2020

Table 6 COVID-19 prediction of the amount of new infected confirmed cases in 10 countries in 2020

Country	R			RMSE			MAE		
	Training	Testing	Predicting	Training	Testing	Predicting	Training	Testing	Predicting
USA	0.956	0.9165	0.9696	3987.4	1938.5	5139.1	2331.3	1567.4	4074.6
India	0.9848	0.9285	0.9933	269.6	359.7	2969.9	137.5	291.5	2106.0
Brazil	0.9752	0.9964	0.8505	1037.2	985.6	7187.2	511.5	830.2	5264.9
Russian Federation	0.976	0.9991	0.9978	312.8	24.0	311.9	86.2	18.9	205.3
France	0.8047	0.9709	0.9971	968.0	405.0	1153.0	574.3	333.0	767.8
Spain	0.9438	0.9766	0.9914	839.0	75.0	864.5	514.6	61.2	578.6
UK	0.9702	0.9894	0.9884	537.7	137.6	1243.2	316.1	123.7	640.7
Argentina	0.8717	0.9853	0.9812	51.3	95.6	877.0	27.6	71.6	625.6
Colombia	0.9436	0.9819	0.9366	74.1	118.6	1015.6	41.3	109.2	744.8
Italy	0.9664	0.9485	0.9983	503.4	108.5	533.7	337.5	98.7	304.3

Without any measures, and the relative risk of infection is 1.5, 2.0, and 3.0, in the next year, the COVID-19 death toll will be 146,996, 293,991, and 587,982, respectively (Banerjee et al. 2020). In the baseline scenario, the basic regeneration number of COVID-19 was 2.68, and the model predicted that the number of people infected in Wuhan was 75,815 as of January 25, 2020. If the transmission characteristics of COVID-19 do not change significantly, the outbreak period of other major cities in China will be 1–2 weeks later than that of Wuhan (Wu et al. 2020). Using the reported data from January 11 to February 10, 2020 to calibrate the susceptible infected recovered dead model, and predict the evolution of the epidemic in Hubei. As of February 29, it is predicted that at least 45,000 people will be infected, and 2700 people will die in Hubei. In fact, about 67,000 people have been infected, and 2800 people have died in this period (Anastassopoulou et al. 2020). The suspension of urban public transport, the closure of places of entertainment, and the prohibition of public gatherings are related to the reduction of COVID-19 cases. Without the Wuhan travel ban and China's emergency response, more than 70,000 people will be infected with the virus outside Wuhan by February 19, 2020. China's prevention and control measures seem to have succeeded in breaking the transmission chain and preventing contact between infectious and susceptible people (Tian et al. 2020). The three major non-drug interventions used in China not only contain the development of the epidemic in China but also win a time window for the world. If we do not implement a strong non-drug intervention "combination boxing," the amount of COVID-19 cases in China may exceed 7 million (Lai et al. 2020). Spatial temporal "risk source" model predicts confirmed cases and identifies high-risk areas in the early stage by using population mobility data (Jia et al. 2020). If the UK government does nothing, it could face more than 500,000 deaths. Without intervention, the USA could face 2.2 million

deaths (Adam 2020). The AI app allows individuals to report symptoms themselves to effectively predict whether they have COVID-19, with an accuracy of nearly 80% (Menni et al. 2020). If the average growth rates 30.6% in the USA during the past 14 days, we will be looking at 3.9 million cases by April 12, 2020 (Perc et al. 2020). After the epidemic, COVID-19 will break out again in winter 2020. The USA may still need long-term or intermittent social alienation interventions by 2022. SARS-CoV-2 should continue to be monitored, as the new outbreak is likely to occur again later in 2025 (Kissler et al. 2020). The 3 biomarkers of COVID-19 are used to predict the mortality rate of COVID-19 patients at least 10 days in advance, and the accuracy rate is over 90% (Yan et al. 2020).

The ANN model with EEMD-based decomposition technique for predicting COVID-19 epidemic is developed. The training and testing period property of the ANN model obtained by R^2 values 0.9997 and 0.99982, respectively. And the R^2 of validation is 0.99981 (Hasan 2020). Cloud computing and machine learning (ML) is deployed to forecast COVID-19 epidemic. The results show that the severity of the global spread of COVID-19 (Tuli et al. 2020a).

The outbreak of COVID-19 has seriously affected the environment, ecology, economy, and society (Kluge et al. 2020). COVID-19 is a major menace to the world economy. In March 2020, the outbreak of the epidemic caused a huge earthquake in the US stock market, triggering the circuit breaker mechanism four times a month. At present, there is an urgent need to know what the future transmission trend of COVID-19 might be.

Conclusions and Implication

We examined the feasibility of using AI with past days' cases as input variables to predict global infected confirmed cases

Fig. 9 Prediction of the amount of new infected confirmed cases of COVID-19 in 10 countries in 2020
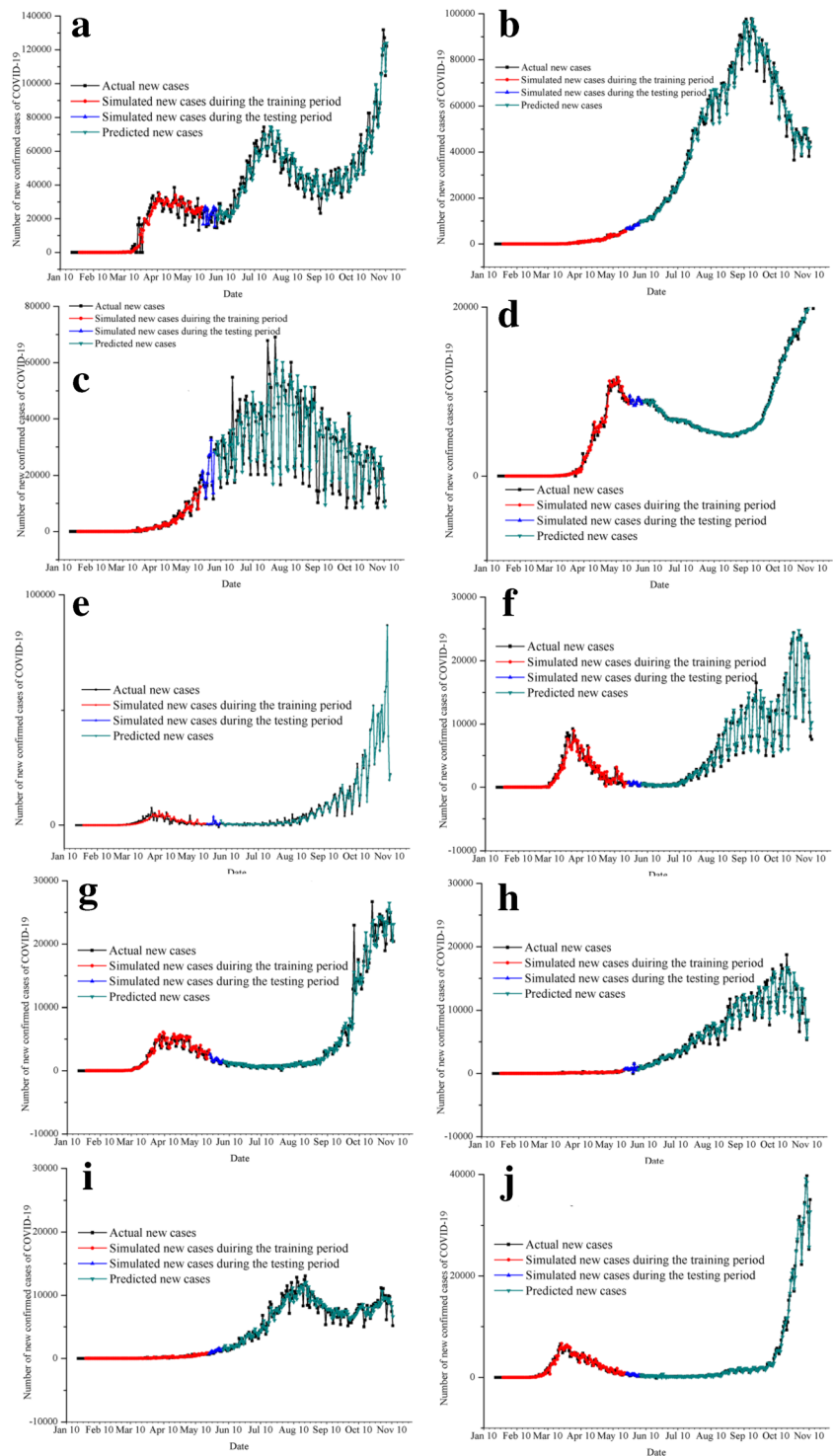
and deaths of COVID-19. The performance of the ANN model is assessed employing three statistical criteria. A simple ANN model with 7 past days' global patients as input variables is identified. The acceptable predicting capability of the ANN is also verified. In view of the global epidemic of COVID-19, we need to further strengthen the research on the prediction model.

At present, the risk of the world economic recession mainly comes from the spread of the epidemic, and the recovery process depends on when the epidemic is contained in the world. Special drugs of COVID-19 are still not successfully developed in the current situation. It is urgent to launch joint action to fight the epidemic. China's experience will undoubtedly provide effective help for the global fight against the epidemic.

We will predict the epidemics trend of COVID-19 for different countries using deep learning approaches, such as the recurrent neural network (RNN), the gated recurrent unit (GRU), and the long short-term memory (LSTM) and compare how their model performs in diverse demographics. We are planning to get a single standard model that can be used for any country, which may be a combination of different algorithms.

People should gather less. People should avoid the places where people gather, especially the places with poor air mobility, and reduce unnecessary going out. If going out, personal protection and frequent hand washing should be done. In densely populated public places, people try to keep a certain social contact distance with others, and it is recommended to wear medical surgical masks. Only by adhering to the concept of human community, following the trend, responding to the times, avoiding the pitfalls of protectionism rationally, strengthening international joint defense and joint control, coordinating macroeconomic policies, and strengthening global supply chain cooperation, can the international community work together to promote the early recovery of the world economy.

Data availability Extra data is available by emailing to guoqingchun@lcu.edu.cn in on reasonable request.

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

Ethics approval and consent to participate All data were obtained with the agreement of world health organization (WHO). The data in this article are obtained from an open database of the World Health Organization (links are provided in the resources section). These can be accessible locally for educational purposes.

Consent to publish Not applicable.

References

Adam D (2020) Special report: the simulations driving the world's response to COVID-19. Nature 580:316–318

Ahmad A, Garhwal S, Ray SK, Kumar G, Malebary SJ, Barukab OM (2020): The number of confirmed cases of Covid-19 by using machine learning: methods and challenges. Arch Comp Methods Eng

Anastassopoulou C, Russo L, Tsakris A, Siettos C (2020) Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PLoS One 15:e0230405

Baker RE, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT (2020) Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. Science 369:315–319 eabc2535

Banerjee A, Pasea L, Harris S, Gonzalez-Izquierdo A, Torralbo A, Shallcross L, Noursadeghi M, Pillay D, Sebire N, Holmes C, Pagel C, Wong WK, Langenberg C, Williams B, Denaxas S, Hemingway H (2020) Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study. Lancet 395:1715–1725

Cassimon T, Vanneste S, Bosmans S, Mercelis S, Hellinckx P (2020): Designing resource-constrained neural networks using neural architecture search targeting embedded devices. Internet of Things, 100234

Chan JF-W, Yuan S, Kok KH, To KKW, Chu H, Yang J, Xing F, Liu J, Yip CCY, Poon RWS, Tsoi HW, Lo SKF, Chan KH, Poon VKM, Chan WM, Ip JD, Cai JP, Cheng VCC, Chen H, Hui CKM, Yuen KY (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 395:514–523

Chimmula VKR, Zhang L (2020) Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons Fractals 135:109864

Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, Pastore Y, Piontti A, Mu K, Rossi L, Sun K, Viboud C, Xiong X, Yu H, Halloran ME, Longini IM, Vespignani A (2020) The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 368:395–400

Guo Q, He Z, Li S, Li X, Meng J, Hou Z, Liu J, Chen Y (2020) Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. Aerosol Air Qual Res 20:1429–1439

Halfmann PJ, Hatta M, Chiba S, Maemura T, Fan S, Takeda M, Kinoshita N, S-i H, Sakai-Tagawa Y, Iwatsuki-Horimoto K, Imai M, Kawaoka Y (2020) Transmission of SARS-CoV-2 in domestic cats. N Engl J Med 383:592–594

Hasan N (2020) A methodological approach for predicting COVID-19 epidemic using EEMD-ANN hybrid model. Internet Things 11:100228

He Z, Zhang Y, Guo Q, Zhao X (2014) Comparative study of artificial neural networks and wavelet artificial neural networks for groundwater depth data forecasting with various curve fractal dimensions. Water Resour Manag 28:5297–5317

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395:497–506

Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA (2020) Population flow drives spatio-temporal distribution of COVID-19 in China. Nature 582:389–394

Kasilingam D, Sathiya Prabhakaran SP, Rajendran DK, Rajagopal V, Santhosh Kumar T, Soundararaj A (2020) Exploring the growth of COVID-19 cases using exponential modelling across 42 countries and predicting signs of early containment using machine learning. Transbound Emerg Dis 67(5):1754–1755

Khakharia A, Shah V, Jain S, Shah J, Tiwari A, Daphal P, Warang M, Mehendale N (2020) Outbreak prediction of COVID-19 for dense and populated countries using machine learning. Ann Data Sci 7(4):697–712

Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M (2020) Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368:860–868

Kluge HHP, Wickramasinghe K, Rippin HL, Mendes R, Peters DH, Kontsevaya A, Breda J (2020) Prevention and control of non-communicable diseases in the COVID-19 response. Lancet 395: 1678–1680

Kraemer MUG, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, du Plessis L, Faria NR, Li R, Hanage WP, Brownstein JS, Layan M, Vespignani A, Tian H, Dye C, Pybus OG, Scarpino SV (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368:493–497

Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, Wesolowski A, Santillana M, Zhang C, Du X, Yu H, Tatem AJ (2020) Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 585:410–413

Lam TT-Y, Shum MH-H, Zhu H-C, Tong Y-G, Ni X-B, Liao Y-S, Wei W, Cheung WY-M, Li W-J, Li L-F, Leung GM, Holmes EC, Hu Y-L, Guan Y (2020) Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. Nature 583:282–285

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z (2020a) Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 382:1199–1207

Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J (2020b) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368:489–493

Liu Y, Ning Z, Chen Y, Guo M, Liu Y, Gali NK, Sun L, Duan Y, Cai J, Westerdahl D, Liu X, Xu K, Ho K-F, Kan H, Fu Q, Lan K (2020) Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. Nature 582:557–560

Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, Ganesh S, Varsavsky T, Cardoso MJ, el-Sayed Moustafa JS, Visconti A, Hysi P, Bowyer RCE, Mangino M, Falchi M, Wolf J, Ourselin S, Chan AT, Steves CJ, Spector TD (2020) Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 26:1037–1040

Perc M, Gorišek Miksić N, Slavinec M, Stožer A (2020) Forecasting COVID-19. Front Phys 8

Price-Haywood EG, Burton J, Fort D, Seoane L (2020) Hospitalization and mortality among black patients and white patients with Covid-19. N Engl J Med 382:2534–2543

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536

Shi J et al (2020) Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. Science 368:1016–1020 eabb7015

Sia SF, Yan L-M, Chin AWH, Fung K, Choy K-T, Wong AYL, Kaewpreedee P, Perera RAPM, Poon LLM, Nicholls JM, Peiris M, Yen H-L (2020) Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. Nature 583:834–838

Sit THC, Brackman CJ, Ip SM, Tam KWS, Law PYT, Tol EMW, Yu VYT, Sims LD, Tsang DNC, Chu DKW, Perera RAPM, Poon LLM, Peiris M (2020) Infection of dogs with SARS-CoV-2. Nature 586:776–778

Tian H, Liu Y, Li Y, Wu CH, Chen B, Kraemer MUG, Li B, Cai J, Xu B, Yang Q, Wang B, Yang P, Cui Y, Song Y, Zheng P, Wang Q, Bjornstad ON, Yang R, Grenfell BT, Pybus OG, Dye C (2020) An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 368:638–642

Tuli S, Tuli S, Tuli R, Gill SS (2020a) Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. Internet Things 11:100222

Tuli S, Tuli S, Verma R, Tuli R (2020b): Modelling for prediction of the spread and severity of COVID-19 and its association with socioeconomic factors and virus types. medRxiv, 2020.06.18.20134874

Wu JT, Leung K, Leung GM (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 395: 689–697

Yan L, Zhang HT, Goncalves J, Xiao Y, Wang M, Guo Y, Sun C, Tang X, Jing L, Zhang M, Huang X, Xiao Y, Cao H, Chen Y, Ren T, Wang F, Xiao Y, Huang S, Tan X, Huang N, Jiao B, Cheng C, Zhang Y, Luo A, Mombaerts L, Jin J, Cao Z, Li S, Xu H, Yuan Y (2020) An interpretable mortality prediction model for COVID-19 patients. Nat Mach Intell 2:283–288

Yang Z, Zeng Z, Wang K, Wong SS, Liang W, Zanin M, Liu P, Cao X, Gao Z, Mai Z, Liang J, Liu X, Li S, Li Y, Ye F, Guan W, Yang Y, Li F, Luo S, Xie Y, Liu B, Wang Z, Zhang S, Wang Y, Zhong N, He J (2020) Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thor Disease 12:165–174

Zhang K et al (2020) Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell 181:1423–1433.e11

Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579:270–273

Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W (2020) A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 382:727–733