The predicted trend of COVID-19 in the United States of America under the policy of “Opening Up America Again”

Kejia Yan a, *, Huqin Yan b, Rakesh Gupta a

a Department of Accounting, Finance and Economics, Griffith University, Nathan, Australia, 4111
b Xiamen National Accounting Institute, Xiamen, 361005, China

ARTICLE INFO

Article history:
Received 27 May 2020
Received in revised form 22 May 2021
Accepted 27 May 2021
Available online 6 June 2021
Handling Editor: Jianhong Wu

Keywords:
COVID-19
The United States of America
Policy and forecasting

ABSTRACT

COVID-19 virus has been spreading worldwide for more than a year. At present, the situation of the new crown pneumonia epidemic remains full of tension and uncertainty. It is of concern is that the worst outbreak in the world is in the United States. The total number of confirmed new cases of COVID-19 and the total number of new deaths in the United States have entered their second and third cyclical peaks since the White House announced the “Open America Again” guidelines on April 16, 2020, and the start of the US presidential election season in August 2020. This paper combines the generalized exponential model (EXPM) with Chebyshev polynomials to develop a special generalized growth model (GGM) to predict the total number of daily new confirmed cases and the total number of new deaths in the United States for three periods under a 14-day sensitivity regression model. In this paper, the US epidemic is divided into three periods from early January 2020 to early January 2021, and three forecasts are made for the three periods. The first two prediction periods have already occurred and the predictions match well with known results, and the third period predicts that the total number of new confirmed cases of COVID-19 and the total number of new deaths in the United States will fall to a minimum level by next July, when the supply of COVID-19 vaccine has already begun. The results suggest that the “Open America Again” policy and the events of the 2020 US presidential election season have contributed to the worsening of the COVID-19 in the United States.

© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Since human-to-human transmission of pneumonia was detected in Wuhan City in China at the end of December 2019, COVID-19, or coronavirus, has spread with astonishing speed around the world (He et al., 2020; Zhao et al.,). On January 13, 2020, Thailand reported the first imported case of COVID-19 outside China. At the end of January 2020, the WHO and the US declared that the COVID-19 outbreak was a public health emergency. On 2 February 2020, the Philippines confirmed the first death from COVID-19 outside China. Six days later, on 26 February 2020, South Korea was the first country to confirm more than one hundred cases of COVID-19 outside China. Six days later, on 26 February 2020, South Korea further confirmed more than one
Coronaviruses, as positive-sense enveloped RNA viruses, belong to the Coronaviridae family, which can infect birds, mammals and humans with acute resolved or fatal pneumonia (Beck et al., 2020; Cavanagh, 2007; Lim et al., 2016; Weiss & Leibowitz, 2011). The Coronaviridae family also includes two other famous coronaviruses, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) ((Kuiken et al., 2003); (Zaki et al., 2003); Zhao, Zhuang, et al., 2020). Even though the current mortality rate of COVID-19 is around 2.20% (based on the data from WHO) which is lower than SARS-CoV (9.6%) and MERS-CoV (34.4%) (WHO, 2004; WHO, 2019), compared with the infection rate of SARS-CoV and MERS-CoV in more than 10,000 people, as of 3 January 2021, COVID-19 has officially infected 83,326,479 people and caused 1,831,703 deaths (WHO, 2021).

The main reasons for this overwhelming increase are that the spread of the coronavirus occurs much faster than people expect, and also because the US is suffering from a severe lack of preparedness for an epidemic (Marzia et al., 2020; WHO, 2004; WHO, 2019; (Zhao, Li et al., 2020)). By 18 May 2020, New York, the most affected state, had reported 361,266 confirmed cases with 28,480 deaths (CDC, 2020). New York City's mortality rate at that time was over 7.88%, higher than the world average of 6.75%. This implies that medical resources are not best allocated in New York.

However, at a time when the number of new confirmed cases per day had reached more than 20,000 and the number of deaths per day had reached more than 1,000, Americans were being urged to return to work because of a looming economic recession. Based on the data of the US Department of Labor (U.S. Bureau of Labor Statistics report; 2020), the unemployment rate jumped to 14.7% in April, the highest in the history of recording. While it is certain that the first turning point in the total number of confirmed new cases and new deaths in the United States occurred in April and early May, the premature "Opening Up America Again" allowed the outbreak to lead to a complete outbreak, the second high point.

Policy confusion and social unrest are important factors in the spreading backlash against COVID-19 in many states. As elections approach, incumbents who can run for re-election impose less stringent epidemic prevention restrictions, particularly in the US and India (Pulejo & Querubín, 2021). Protesters and crowds gathered for the presidential election campaign not only exacerbated the social rifts, but also directly increased the difficulty of controlling the outbreak. When people are concentrated in large numbers in crowded and inadequately ventilated environments, large-scale outbreaks are inevitable (Gonsalves & Yamey, 2020; Ran et al., 2021). Indeed, there is urgent need for a COVID-19 vaccine, and for accurate prediction of the trend of the outbreak.

The fear people have of the virus is not only that the quantity and effects of vaccine for COVID-19 is still unclear, but also that the final scale of the disease for the forthcoming period has not been appropriately forecast. If there is no relative accurate measurement of the range of the final scale of the coronavirus disease, the policymakers and the medical system will be unable to allocate medical resources to maximum effect, causing the rhythm of economy recovery to be even further delayed. In order to provide guidance for policymakers and the medical system about the prevention and control of this pandemic, this study analyze the data of COVID-19 for the United States of America, as released by WHO, using Chebyshev polynomial functions to construct generalized exponential models of the trend of coronavirus disease transmission under different growth patterns. Moreover, this study forecast the scale of the coronavirus disease in the US for the next 180 days up to 1 July 2021, and find that the curves of the proposed model well simulate the US’s official data curves.

2. Methods

2.1. Data and observation

The empirical data of the coronavirus disease (COVID-2019) was taken from the World Health Organization’s situation reports (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/), and Microsoft Excel and Python programming language were used to build and analyze time-series database. The World Health Organization released the first outbreak report on novel coronavirus (2019-nCoV) on January 21, 2020. Since then, WHO has issued daily or weekly COVID-19 outbreak reports.

In the US, the first five new cases of coronavirus pandemic was confirmed on 20 January 2020. In this paper, the total number of daily new confirmed cases was selected from 20 January 2020 to 3rd January 2021, and the total number of observed cases was 350. Until 3rd January 2021, the total confirmed cases of COVID-19 in the US had reached 19,974,413 cases. The curve of the new confirmed cases of COVID-19 in the United States to date can be divided into three periods: period 1 was between January and May 2020, with a peak incidence of 35,386 cases on 11 April 2020; period 2 was between June and September 2020, with a peak incidence of 74,354 cases on 19 July 2020; period 3 was beginning from October 2020, with a peak incidence of 402,270 cases on 20 Dec 2020.

In the US, the first two new deaths of coronavirus pandemic was confirmed on 3rd March 2020. The data in this paper were selected for the total number of new deaths per day between 3rd March 2020 and 3rd January 2021, with a total 307 observations. Until 3 January 2021, the total deaths of COVID-19 in the US had reached 345,253 cases. The curve of the total new deaths in the US COVID-19 can be clearly divided into three periods: the first period is from March–June 2020, where the peak
incidence is on April 17, with 6409 cases; the second period is from July—October 2020, with a peak incidence of 1486 cases on August 14; and the third period starts from November 2020, and the highest peak is 5703 cases on January 1, 2021.

The second peaks of the total new confirmed cases occurred in July 2020 and the total new deaths occurred in August 2020 might be caused by the policy of Donald Trump’s “opening up America again”. On 16 April 2020, the US President Donald Trump announced plan to reopen virus-ravaged U.S. economy, published guidelines for reopening the coronavirus-ravaged U.S. economy, asked states take a phased approach to let Americans return to work as conditions allow. Trump’s guidelines (Whitehouse, 2020) include a three-stage prescription for restarting normal life in America: In Phase 1, restaurants, movie theaters, sports venues, and places of worship would be allowed to reopen and resume work if possible; in Phase 2, schools could reopen and non-essential travel could resume, but most employees would be encouraged to continue to work remotely; and in Phase 3, workplaces would be restored for unrestricted staff. The rush to return to work without effective disease control methods caused this spike in infections. The third peak in the total number of new confirmed cases occurs in December 2020, while the total number of new deaths in January 2021 may be due to the US presidential election season. Since August 2020, the United States entered into a presidential election season.

To observe three cyclical trend changes in new confirmed cases and new deaths, regression models will be built periodically in this paper. The basic empirical data include the total number of new confirmed cases and the total number of new deaths. Since total confirmed cases and total deaths can be accumulated from the basic empirical data, this paper will focus on building regression models for new confirmed cases and new deaths to obtain the predicted values of total confirmed cases and total deaths. The predict period will begin from 4 January 2021 until 2 July 2021 and last 180 days.

For doing a comparison, this paper not only predicts the total number of new confirmed cases, the total number of confirmed cases, and the total number of deaths, but also compares the empirical data with the predicted values.

2.2. GGM model

The generalized exponential model provides an upper bound for future scenarios. This model has been used to characterize and forecast the early epidemic growth patterns across a diversity of disease outbreaks in many diseases such as Ebola, foot-and-mouth disease, HIV/AIDS, influenza, measles, plague and smallpox ((Viboud et al., 2016); Wu et al., 2020).

If \(Y(t)\) represents a growth function, \(r(t)\) is a relative rate function at time \(t\), and if the pair \(Y(t), r(t)\) satisfy ordinary differential equation (ODE) function, then the model will be suitable for describing a growth procedure (Koya & Gosho, 2013), thus a generalized growth model (GGM) can be defined as:

\[
Y'(t) = \frac{dY(t)}{dt} = rY^p(t)
\]

(1)

where \(Y'(t)\) describes the incidence growth phase over time \(t\); \(Y(t)\) describes the cumulative number of disease cases at time \(t\); \(r\) \((r > 0)\) is the growth rate which controls the characteristic time scale of the dynamic, and \(p \in [0,1]\) is the exponent that allows the model to capture different growth conditions from constant incidence to exponential growth for three cases:

**Case 1.** for a constant growth rate incidence, \(p = 0\)

**Case 2.** for a sub-exponential growth, \(0 < p < 1\), its solution equation is:

\[
Y(t) = Y(0)\left(1 + \frac{rt}{A}\right)^b \quad \text{and} \quad b = \frac{1}{1-p} \quad A = \frac{Y(0)^{1-p}}{1-p}
\]

(2)

**Case 3.** for an exponential growth, \(p = 1\). Then the GGM model will be written as the exponential growth model (EXPM) with the solution:

\[
Y(t) = Y(0)e^{rt}
\]

(3)

The outbreak will decrease transmission rates when proper epidemiological control and treatment is in place. At such time, logistic type models, such as the generalized logistic model (GLM) and the generalized Richards model (GRM) will be used to characterize the performance of the outbreak. Compared with the GGM model which has only one free parameter \(r\), logistic type models have more free parameters (Chowell, 2017). For example, a GRM model can be examined as:

\[
Y'(t) = \frac{dY(t)}{dt} = rY^p(t)\left[1 - \left(\frac{Y(t)}{K}\right)^a\right]
\]

(4)

where parameter \(r\) represents the intrinsic growth rate, \(K\) represents the size of the epidemic, and \(a\) represents the extent of deviation. As Burger et al. (2019) summarized, differential equations can be categorized by the selections of the parameter \(r,p,a,k\). To simplify the equation, when \(a = 1\), then a GRM model becomes a generalized growth model (GGM).
It is first essential to determine the value of the pre-required parameter $K$ in the analysis. However, parameter $K$, as the final asymptotic total number of infections during the whole epidemic, is very hard to predict for COVID-19 at this stage. Similarly, an appropriate time scale of the epidemic’s growth process $r$ is very hard to set under the uncertainty of an epidemic prevention policy when the White House is eagerly engaged in “opening up American again”. The volatility for total numbers of the new confirmed cases and total numbers of new deaths became very high since then.

Thus, in order to avoid the estimation of epidemic growth process $r$, we will combine the exponential growth model (EXPM) and a polynomial function together to forecast the early epidemic growth patterns.

Assume variable $x(t)$ is defined as a continuous function of the time variable $t$:

$$x(t) = \frac{1}{1 + t}, \quad \text{or} \quad t = \frac{1}{x} - 1 \quad (5)$$

Function $Y(x)$ can be defined as:

$$Y(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_q x^q = a_0 + a_1 \left( \frac{1}{1 + t} \right) + a_2 \left( \frac{1}{1 + t} \right)^2 + \ldots + a_q \left( \frac{1}{1 + t} \right)^q \quad (6)$$

Assume that the variable $y(t)$ is a COVID-19 related variable, such as the total numbers of confirmed new cases and new deaths. To avoid a zero value of the variable, $y(t)$ may be calculated in a logarithm function. We will define the variable $Y(t)$ as:

$$Y(t) = \ln(y(t)), \quad y(t) = e^{Y(t)} \quad (7)$$

Then combine equations (6) and (7):

$$y(t) = e^{a_0 + a_1 x + a_2 x^2 + \ldots + a_q x^q} \quad (8)$$

For a discrete time system, if the empirical sample time is $t_i \in \{0, 1, 2, \ldots, n\}$, $n \in \mathbb{N}$, $x_i = x(t_i)$, then:

$$Y_i = a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_q x_i^q + \epsilon_i \quad (9)$$

Or it can be directly represented by the time variable $t_i$ as:

$$Y_i = a_0 + a_1 \left( \frac{1}{1 + t_i} \right) + a_2 \left( \frac{1}{1 + t_i} \right)^2 + \ldots + a_q \left( \frac{1}{1 + t_i} \right)^q + \epsilon_i \quad (10)$$

Here the variable $\epsilon_i$ is the residual item or error of the regression. It is the difference between the observed value and the predicted value. For improving the accuracy of calculation, the function $Y(x)$ will use a Chebyshev polynomial. When the time point $t_i$ is used as the Chebyshev polynomial roots within the full time period, $x_i^q$ can be replaced by the $q$– order basic Chebyshev polynomial.

Time series models are significant in predicting the impact of the COVID-19 outbreak and in taking the necessary measures to respond to this crisis (Ceylan, 2020). In dealing with the time series data, many researchers have applied the autoregressive integrated moving average (ARIMA) model to predict the trend of epidemic outbreaks. Ribeiro et al. (2020) used an ARIMA model to forecast the short-term trend of COVID-19 cumulative confirmed cases in Brazil, and they suggest that a 5-day lag may have some sensitive impacts on regression results. Chakraborty and Ghosh (2020) used ARIMA models to forecast the COVID-19 trend for Canada, France, India, South Korea, and the UK, and they suggest that the lags may be 1 or 2 days. Ahmar et al. (2020) predicted the short-term effects of confirmed cases of COVID-19 in Spain by using the Sutte ARIMA method, and suggest that the lags may be 0, 1, or 2 days. Ceylan (2020) used an ARIMA model to predict the epidemiological trend of COVID-19 prevalence in Italy, Spain, and France, and suggest that the lags may be 0 or 1 day. Because the ARIMA model is dependent on the order of time series stationarity, it is not suitable to analyze longer lags of time variables in a regression. Different to the ARIMA models, we will analyze regression models independently when the size of the empirical window is less than the size of the full sample window by 14 day observations. The reason for choosing 14 days is that the effects of isolation and social distancing will take 14 days to show up, just as the “opening up American again” requires a lag of time to characterize the impact of policy on the epidemic.

To analyze how well the regression models are fitted to a country’s epidemic outbreaks, the root mean squared error (RMSE) is usually used in an assessment (Burger et al., 2019; Ceylan, 2020). Many other similar assessments, including the mean absolute error (MAE) and mean absolute percentage error (MAPE), are used to evaluate the effectiveness of adopted models (Chakraborty & Ghosh, 2020; Ribeiro et al., 2020). We will do a similar thing, and moreover, we will consider the values of the coefficient of determination ($R^2$).
2.3 Regression time windows

Fig. 1 lists the structure of the time windows for forecasting. The size of the full sample window is $N$, which will be combined by three sub windows, including an empirical sample window, a test window and a forecasting window.

The size of the empirical sample window is $n$. The regression model will be based on the empirical sample window. At the time point $t_i \in \{0, 1, 2, ..., n\}$, the COVID-19 related variable $y_i \in \{y_0, y_1, ..., y_n\}$ will be introduced to the regression model. When the time point $t_i \in \{n + 1, ..., N\}$, the COVID-19 related variable $y_i \in \{y_{n+1}, ..., y_N\}$ will be the prediction value from the regression model.

The size of the test window is $m$. Since the total number of new confirmed cases and new deaths in the last 14 days are quite sensitive to building regression models and predicting future trends, this paper will perform a sensitivity test by examining the effect of the test window on the prediction time $m = 0, 1, 2, ..., 13$.

The size of the expected forecasting window is $NN$. Since the purpose of this paper is to test the prediction results for three time periods, different time windows will be chosen. Table 1 lists the structure of the three empirical sample windows and three forecasting windows for the total new confirmed cases in the US. Table 2 lists the structure of the three empirical sample windows and three forecasting windows for the total new deaths in the US. The structure of the empirical analysis is the same for all three periods.

2.4 Assessment of the degree of freedom

The coronavirus pandemic has swept through the world each country and every day over a year. It is difficult to predict the exact number of new confirmed cases or new deaths which will occur today or tomorrow. For the purposes of this research, we will determine an appropriate value of $q$ by using the following assessments.

First, the value of $q$ is selected if it is fitted to an appropriate model that has a greater value of the coefficient of determination $R^2$ than the other models when $q$ is changed and satisfied the conditions of $q \leq \frac{1}{2}(n - m - 1)$.

Second, the value of $q$ is selected if it is fitted to an appropriate model that has a smaller value of the root mean square error (RMSE), the mean absolute error (MAE) and mean absolute percentage error (MAPE), than the other models when the value of $q$ is changed.

Third, because the degree of freedom is also the power of the independent variable, we will choose a lower value of $q$ if both of the values of $R^2$ or $R$ are the same.

Fourth, for Chebyshev polynomial, when the number of power $q$ increases, the degree of accuracy will also increase, we will choose a reasonable value of $q$ by observing the trend of curve, it is required that the prediction value should not be less than zero or close to an infinite value.

3. Results and discussion

3.1 Forecasting regression of the total numbers of new confirmed cases and confirmed cases

3.1.1 First period forecast results (20/01/2020–19/05/2020)

In order to forecast the total numbers of new confirmed cases which are likely to occur every day during the period between 20 January 2020 and 19 May 2020, A suitable polynomial power $q$ needs to be found. Thus, after testing the models for each parameter of $q$ from different degrees of freedom, parameter $q = 29$ was found to be the best choice when the values

![Fig. 1. The structure of the time windows for empirical analysis and future forecasting.](image-url)
of RMSE, MAE, and MAPE are the smallest and the values of $R^2$, $R$ and $Adj. R^2$ are the largest. Table 3 lists the results of testing the degree of freedom in regression.

When the power of the polynomial $q$ is defined, the analysis of the sensitivity of the window size will be tested as $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. The sample data from within the last 14 days has strong sensitivity for building the regression models for forecasting future trends. Table 4 lists the test results by using the selected 29-order Chebyshev polynomial based on the total confirmed new cases for the period one. When the power of the polynomial is defined as $q = 29$, the sensitivity analysis window size is tested as $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. All of the 14 models’ values of $RMSE$, $MAE$, and $MAPE$ are quite small and the values of $R^2$, $R$ and $Adj. R^2$ are quite large.

Fig. 2 depicts the prediction curves of the total confirmed new cases from the selected 29-order Chebyshev polynomial for the period one. When the power of the polynomial is defined as $q = 8$, the other 12 prediction curves from the models $m = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$ are located in the middle between the two curves of $m = 0, 13$. The distribution of curve $m = 0$ to 13 represents the fact that the pandemic backlash has begun as more U.S. states respond to “Open America Again”.

Table 3 shows the prediction curves of the selected 29th order Chebyshev polynomial model for the total number of confirmed cases of COVID-19 in period one. It is clearly that the curve of the sample total confirmed cases represented in reference sample is between the lowest prediction curve when $m = 4$ and the highest prediction curve when $m = 8$. When comparing the predicted and actual values, the total confirmed cases of COVID-19 in the US of reference sample were
1,757,522 on 1 June, 2,573,393 on 1 July, 4,456,389 on 1 August and 5,936,572 on 1 September 2020; comparatively, the average predicted total confirmed cases of COVID-19 in the US were 1,813,887 on 1 June, 2,841,322 on 1 July, 4,416,297 on 1 August and 6,336,880 on 1 September 2020. It can be seen that the mean predicted total number of confirmed cases in period one is very close to the total number of confirmed cases in the reference sample.

3.1.2. Second period forecast results (20/01/2020–03/09/2020)

For the second period, the empirical sample window is between 20 January 2020 and 3 September 2020, and the forecasting window is between 4 September 2020 and 3 January 2021. Following the approach of the first prediction, we make the optimal choice of the original data in Tables 5 and 6. Table 5 lists the test result for choosing the best model of the Chebyshev polynomial based on the total confirmed new cases during the period two. The parameter \( q = 69 \) is the best choice when the

| \( q \) | 69* | 68 | 67 | 66 | 65 | 64 | 63 | 62 | 61 | 60 | 59 | 58 | 57 | 56 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| \( R^2 \) | 0.956 | 0.954 | 0.954 | 0.953 | 0.951 | 0.951 | 0.951 | 0.949 | 0.949 | 0.947 | 0.947 | 0.947 | 0.946 | 0.946 |
| \( R \) | 0.978 | 0.977 | 0.977 | 0.976 | 0.975 | 0.975 | 0.975 | 0.974 | 0.974 | 0.973 | 0.973 | 0.973 | 0.972 | 0.972 |
| \( Adj. R^2 \) | 0.937 | 0.934 | 0.934 | 0.934 | 0.931 | 0.931 | 0.932 | 0.930 | 0.930 | 0.928 | 0.928 | 0.929 | 0.929 | 0.927 | 0.928 |
| \( RMSE \) | 4467 | 4588 | 4600 | 4613 | 4727 | 4738 | 4748 | 4832 | 4840 | 4905 | 4912 | 4918 | 4979 | 4986 |
| \( MAE \) | 3281 | 3326 | 3325 | 3326 | 3356 | 3363 | 3372 | 3450 | 3458 | 3512 | 3516 | 3522 | 3560 | 3565 |
| \( MAPE \) | 246 | 227 | 224 | 225 | 257 | 254 | 252 | 235 | 235 | 239 | 237 | 236 | 239 | 238 |

Note: (1) * represents that the Chebyshev polynomial is the best model; (2) the listed models are representatives from the set of models.
values of $RMSE$, $MAE$, and $MAPE$ are the smallest and the values of $R^2$, $R$ and $Adj. R^2$ are the largest. \textbf{Table 6} lists the test results by using the selected 69-order Chebyshev polynomial for the total confirmed new cases for the period two. When the power of the polynomial is defined as $q = 69$, the sensitivity analysis window size is tested as $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. All of the 14 models’ values of $RMSE$, $MAE$, and $MAPE$ are quite small and the values of $R^2$, $R$ and $Adj. R^2$ are quite large.

\textbf{Fig. 4} depicts the prediction curves of the total confirmed new cases from the selected 69-order Chebyshev polynomial for the period two when $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. The second daily peak of the total confirmed new cases were 74,354 cases that occurred on 19 July 2020 in the US. After the second peak, the total number of new confirmed cases showed a downward trend in August and an upward trend after August. Until 20 December 2020, the daily total confirmed cases in the US increased reaching the peak of 402,270 cases. It is clearly that all of the 14 prediction curves have revealed the real rebound trend (reference sample) of the total confirmed new cases from September to December 2020.

\textbf{Fig. 5} shows the prediction curves of the total confirmed cases of COVID-19 from the selected 29-order Chebyshev polynomial model for the period two. Obviously, the curve of total confirmed cases in the sample represented by the reference sample is covered by the curve with the lowest predicted value when $m = 13$. When the total confirmed cases of COVID-19 in the US of reference sample were 7,115,491 on 1 October 2020, 8,952,086 on 1 November, 13,234,551 on 1 December 2020 and 19,346,790 on 1 January 2021. Comparatively, the prediction values of the total confirmed cases of COVID-19 from the model of $m = 13$ were 6,810,711 on 1 October 2020, 9,213,079 on 1 November, 14,105,290 on 1 December 2020 and 21,231,354 on 1 January 2021. The total number of predicted minimum confirmed cases in period two when $m = 13$ was very close to the total number of confirmed cases in the reference sample. The distribution of curve $m = 0$ to 13 represents the fact that the pandemic backlash has begun as U.S. presidential election season is coming up.

### 3.1.3. Third period forecast results (20/01/2020–03/01/2021)

Given the high accuracy of the first two forecasts, for the third period, the empirical sample window is between 20 January 2020 and 3 January 2021, and the forecasting window is between 4 January 2021 and 2 July 2021. Similarly, the optimal choice of model for the Chebyshev polynomial is given in \textbf{Tables 7 and 8. Table 7} lists the test result for choosing the best model of the Chebyshev polynomial based on the total confirmed new cases during the period three. The parameter $q = 66$ is the best choice when the values of $RMSE$, $MAE$, and $MAPE$ are the smallest and the values of $R^2$, $R$ and $Adj. R^2$ are the largest. \textbf{Table 8} lists the test results by using the selected 66-order Chebyshev polynomial based on the total confirmed new cases for the period three. When the power of the polynomial is defined as $q = 66$, the sensitivity analysis window size is tested as $m = \ldots$
0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13. All of the 14 models’ values of RMSE, MAE, and MAPE are quite small and the values of $R^2$, $R$ and Adj. $R^2$ are quite large.

Fig. 6 depicts the prediction curves of the total confirmed new cases from the selected 66-order Chebyshev polynomial for the period three when $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. From the prediction results, we can see that the prediction values of the total confirmed new cases of COVID-19 in the US will continuously increase when $m = 7, 8, 9, 10, 11, 12, 13$ or change to decrease when $m = 0, 1, 2, 3, 4, 5, 6$. Fig. 7 depicts the prediction curves of the total confirmed cases of COVID-19 from the selected 66-order Chebyshev polynomial model for the period three. On 1 January 2021, the sample total confirmed cases of COVID-19 in the US was 19,346,790 cases. When $m = 12$ or $m = 13$, the prediction values of the total confirmed cases of COVID-19 in the US will reach the highest values of 106,109,318 cases or 112,300,253 cases until to the day of 1 July 2021.

Since the COVID-19 vaccine has been mass-produced and the U.S. presidential election is over, there is reason to be optimistic in selecting the lowest predicted value of the model $m = 2$. 1 as the best models to predict the future values of the total confirmed new cases and total confirmed cases of COVID-19 in the US. According to the prediction results, the total confirmed new cases of COVID-19 in the US will decrease reaching the lowest level on 1 July 2021. When $m = 1$ and $m = 2$, the predicted range of the total confirmed new cases of COVID-19 in the US is between 3028 and 21388 daily. Moreover, the prediction values of the total confirmed cases of COVID-19 in the US will have reached between 38,102,565 and 40,271,338 cases by the day of 1 July 2021.

3.2. Forecasting regression of the total numbers of new deaths and total deaths

3.2.1. First period forecast results (03/03/2020–19/05/2020)

For the first period, the empirical sample window is between 3 March 2020 and 19 May 2020, and the forecasting window is between 20 May 2020 and 2 September 2020. Table 9 lists the test result for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period one. The parameter $q = 19$ is the best choice when the values of RMSE, MAE, and MAPE are the smallest and the values of $R^2$, $R$ and Adj. $R^2$ are the largest. Table 10 lists the test results by using the selected 19-order Chebyshev polynomial based on the total confirmed new cases for the period one. When the power of the polynomial is defined as $q = 19$, the sensitivity analysis window size is tested as $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. Fig. 8 depicts the prediction curves of the total new deaths from the selected 19-order Chebyshev polynomial model for the period one when $m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. Since the first two COVID-19 new cases were confirmed on 3 March 2020 in the US, the first daily peak of the total new deaths were 6409 cases that occurred on 17 April 2020. After decreasing in May and increasing in June and July, the second daily peak of the total new deaths were 1486 cases that occurred on 14 August 2020.

Fig. 9 shows the prediction curves of the total new deaths of COVID-19 from the selected 19-order Chebyshev polynomial model for the period one. It is clearly that the curve of the sample total deaths represented in reference sample is almost covered by the prediction curve of $m = 11$, which is between the prediction curves of $m = 12$. 13 and the prediction curves of $m = 0, 5, 6$. The reference sample was 182,162 cases on September 1, 2020, while the predicted number of deaths for $m = 11$ was 191,534 cases, and the predicted curve of deaths was close to the total sample deaths.
Table 7
Test results for choosing the best model of the Chebyshev polynomial based on the total new confirmed cases during the period 3.

| q  | 66* | 65 | 64 | 63 | 62 | 61 | 60 | 59 | 58 | 57 | 56 | 55 | 54 | 53 |
|----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| $R^2$ | 0.878 | 0.877 | 0.877 | 0.875 | 0.875 | 0.873 | 0.872 | 0.870 | 0.870 | 0.868 | 0.868 | 0.867 | 0.866 | 0.866 |
| R  | 0.937 | 0.936 | 0.936 | 0.935 | 0.935 | 0.934 | 0.934 | 0.933 | 0.933 | 0.932 | 0.932 | 0.931 | 0.931 | 0.931 |
| Adj. $R^2$ | 0.850 | 0.849 | 0.849 | 0.847 | 0.847 | 0.846 | 0.846 | 0.844 | 0.844 | 0.843 | 0.843 | 0.842 | 0.842 | 0.842 |
| RMSE | 22390 | 22514 | 22532 | 22695 | 22719 | 22905 | 22930 | 23111 | 23135 | 23293 | 23314 | 23442 | 23458 | 23473 |
| MAE | 9612 | 9793 | 9816 | 10021 | 10050 | 10208 | 10226 | 10376 | 10383 | 10561 | 10595 | 10726 | 10726 | 10718 |
| MAPE | 878 | 944 | 949 | 958 | 960 | 930 | 925 | 1014 | 1014 | 978 | 978 | 1024 | 1021 | 1015 |

Note: (1) * represents that the Chebyshev polynomial is the best model; (2) the listed models are representatives from the set of models.

Table 8
Test results by using the selected 66-order Chebyshev polynomial based on the total new confirmed cases during the period 3.

| m  | 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| $R^2$ | 0.878 | 0.893 | 0.891 | 0.925 | 0.925 | 0.926 | 0.931 | 0.932 | 0.939 | 0.940 | 0.940 | 0.940 | 0.943 | 0.944 |
| R  | 0.937 | 0.945 | 0.944 | 0.962 | 0.962 | 0.962 | 0.965 | 0.966 | 0.969 | 0.970 | 0.969 | 0.970 | 0.971 | 0.972 |
| Adj. $R^2$ | 0.850 | 0.867 | 0.865 | 0.908 | 0.907 | 0.907 | 0.908 | 0.914 | 0.916 | 0.924 | 0.926 | 0.924 | 0.926 | 0.929 | 0.931 |
| RMSE | 22390 | 20211 | 20192 | 16676 | 16653 | 16461 | 15857 | 15596 | 14789 | 14554 | 14525 | 14300 | 13894 | 13625 |
| MAE | 9612 | 9509 | 9526 | 8323 | 8440 | 8235 | 7922 | 7727 | 7403 | 7239 | 7181 | 7027 | 6845 | 6712 |
| MAPE | 878 | 882 | 885 | 247 | 241 | 231 | 210 | 196 | 173 | 160 | 156 | 142 | 126 | 115 |

Fig. 6. Prediction curves of the total new confirmed cases from the selected 66-order Chebyshev polynomial model for the period 3.

Fig. 7. Prediction curves of the total confirmed cases from the selected 66-order Chebyshev polynomial model for the period 3.
Table 9

Test results for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period 1.

| q  | 19* | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10  | 9  | 8  | 7  | 6  |
|----|-----|----|----|----|----|----|----|----|----|-----|----|----|----|----|
| R² | 0.589 | 0.588 | 0.587 | 0.585 | 0.585 | 0.584 | 0.584 | 0.583 | 0.578 | 0.567 | 0.549 | 0.524 | 0.497 | 0.477 |
| R  | 0.767 | 0.767 | 0.766 | 0.765 | 0.765 | 0.764 | 0.764 | 0.764 | 0.760 | 0.753 | 0.741 | 0.724 | 0.705 | 0.691 |
| Adj R² | 0.454 | 0.462 | 0.470 | 0.477 | 0.484 | 0.492 | 0.500 | 0.506 | 0.508 | 0.502 | 0.489 | 0.468 | 0.447 | 0.433 |
| RMSE | 751.4 | 751.9 | 752.9 | 754.1 | 755.0 | 755.2 | 755.2 | 756.4 | 760.9 | 770.7 | 787.0 | 808.4 | 830.3 | 846.9 |
| MAE | 357.9 | 357.7 | 360.3 | 362.6 | 361.6 | 357.7 | 361.0 | 371.5 | 387.5 | 406.3 | 434.6 | 464.6 | 488.7 | 487.1 |
| MAPE | 7.78 | 8.71 | 10.27 | 10.98 | 9.33 | 7.48 | 9.08 | 12.13 | 14.81 | 17.83 | 27.87 | 35.64 | 29.73 | 34.81 |

**Note:** (1) * represents the best model; (2) the listed models are representatives from the set of models.

3.2.2. Second period forecast results (03/03/2020–03/09/2020)

For the second period, the empirical sample window is between 3 March 2020 and 3 September 2020, and the forecasting window is between 4 September 2020 and 3 January 2021. Table 11 lists the test results for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period two. The parameter \( q = 35 \) is the best choice when the values of \( RMSE, MAE, \) and \( MAPE \) are the smallest and the values of \( R², R \) and \( Adj. R² \) are the largest. Table 12 lists the test results by using the selected 35-order Chebyshev polynomial based on the total new deaths for the period two. When the power of the polynomial is defined as \( q = 35 \), the sensitivity analysis window size is tested as \( m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 \).

Fig. 10 depicts the prediction curves of the total new deaths from the selected 35-order Chebyshev polynomial for the period two when \( m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 \). In the US, after the second peak occurred on 14 August 2020, the total new deaths tended to decrease in September and October and increase in November and December 2020. Until 1 January 2021, the daily total new deaths of COVID-19 in the US increased and reached 5703 cases. It is evident that all 14 predicted curves cover the area of the total new deaths in the sample from September to December 2020.

Fig. 11 depicts the prediction curves of the total deaths of COVID-19 from the selected 35-order Chebyshev polynomial model for the period two. Significantly, the total sample mortality curve represented by the reference sample lies between the predicted curves of the lowest and highest predicted values when \( m = 1 \) and \( m = 6 \). The mean predicted value of total deaths as of January 1, 2021 is 320,373, which is very close to the reference sample total of 335,789 deaths.

3.2.3. Third period forecast results (03/03/2020–03/01/2021)

As the accuracy of the first two forecasts is high, the empirical sample window is between 3 March 2020 and 3 January 2021, and the forecasting window is between 4 January 2021 and 2 July 2021. Identically, the optimal model choice for the Chebyshev polynomial is shown in Tables 13 and 14. Table 13 lists the test results for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period three. The parameter \( q = 63 \) is the best choice when the values of \( RMSE, MAE, \) and \( MAPE \) are the smallest and the values of \( R², R \) and \( Adj. R² \) are the largest. Table 14 lists the test results by using the selected 63-order Chebyshev polynomial based on the total confirmed new cases for the period three. When the power of the polynomial is defined as \( q = 63 \), the sensitivity analysis window size is tested as \( m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 \).

Fig. 12 depicts the prediction curves of the total confirmed new cases from the selected 63-order Chebyshev polynomial for the period three when \( m = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 \). The third daily peak of the total new deaths were 5703 cases that occurred on 1 January 2021 in the US. From the prediction models, we can see that the prediction values of the total new deaths of COVID-19 in the US will continuously increase when \( m = 0, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 \) or change to decrease when \( m = 2, 1 \).

Fig. 13 depicts the prediction curves of the total deaths of COVID-19 from the selected 63-order Chebyshev polynomial model for the period three. On 1 January 2021, the sample total deaths of COVID-19 in the US was 335,789 cases. Until the day of 1 July 2021, when \( m = 12 \) or \( m = 13 \), the prediction values of the total deaths of COVID-19 in the US will reach the highest values of 1,654,950 cases or 1,941,782 cases, which are 4 times more than the sample total deaths on 1 January 2021.

After the COVID-19 vaccine has been mass produced and the US presidential election is over, we have reason to be optimistic in choosing the predicted value of the model, i.e. \( m = 2,1 \) is the best model to predict the future value of total confirmed new cases of COVID-19 in the US. According to the prediction models of \( m = 2, 1 \), the total new deaths of COVID-19 in the US will decrease and reach the lowest level in July 2021. When \( m = 2 \) or \( m = 1 \), then the prediction values of the total new deaths of COVID-19 in the US will reach 1958 and 3348 cases on 1 July 2021 and the prediction values of the total deaths of COVID-19 in the US will have reached between 771,862 and 930,619 cases on 1 July 2021.

4. Summary and comments

At the present time the coronavirus disease situation remains tense and uncertain. There is great concern that in the US the situation of the COVID-19 pandemic is the worst in the world. But even worse, when the number of new confirmed cases reaches more than 20,000 daily and the number of deaths reaches more than 1000 daily, Americans were heading to back to work at a time of heavy economic recession. Since the White House published “opening up America again” guidelines on 16
Table 10
Test results by using the selected 19-order Chebyshev polynomial based on the total new deaths during the period 1.

| m  | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $R^2$ | 0.589 | 0.593 | 0.592 | 0.592 | 0.592 | 0.591 | 0.593 | 0.597 | 0.596 | 0.596 | 0.598 | 0.570 | 0.568 |
| $R_0$ | 0.767 | 0.770 | 0.770 | 0.770 | 0.770 | 0.769 | 0.769 | 0.770 | 0.773 | 0.772 | 0.772 | 0.773 | 0.755 | 0.754 |
| Adj. $R^2$ | 0.454 | 0.457 | 0.454 | 0.452 | 0.449 | 0.445 | 0.442 | 0.444 | 0.439 | 0.436 | 0.435 | 0.492 | 0.488 |
| RMSE | 751.4 | 751.8 | 756.6 | 761.2 | 763.9 | 768.4 | 773.7 | 777.5 | 784.4 | 789.6 | 782.0 | 811.0 | 811.5 |
| MAE | 357.9 | 359.6 | 363.8 | 366.9 | 361.7 | 361.5 | 366.3 | 367.4 | 365.9 | 371.0 | 373.6 | 360.5 | 409.6 | 411.1 |
| MAPE | 7.78 | 8.58 | 8.81 | 9.18 | 8.73 | 8.53 | 8.63 | 9.27 | 10.25 | 10.46 | 11.19 | 9.42 | 20.28 | 22.05 |

Note: (1) * represents that the Chebyshev polynomial is the best model; (2) the listed models are representatives from the set of models.

Fig. 8. Prediction curves of the total new deaths from the selected 19-order Chebyshev polynomial models for the period 1.

Fig. 9. Prediction curves of the total deaths from the selected 19-order Chebyshev polynomial models for the period 1.

Table 11
Test results for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period 2.

| q  | 35* | 34 | 33 | 32 | 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 |
|----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| $R^2$ | 0.558 | 0.558 | 0.558 | 0.558 | 0.557 | 0.557 | 0.556 | 0.552 | 0.549 | 0.548 | 0.548 | 0.547 | 0.546 | 0.545 |
| $R_0$ | 0.747 | 0.747 | 0.747 | 0.747 | 0.747 | 0.747 | 0.746 | 0.743 | 0.741 | 0.740 | 0.740 | 0.740 | 0.739 | 0.738 |
| Adj. $R^2$ | 0.455 | 0.458 | 0.461 | 0.464 | 0.464 | 0.471 | 0.473 | 0.472 | 0.472 | 0.473 | 0.476 | 0.479 | 0.481 | 0.484 |
| RMSE | 544.3 | 544.5 | 544.9 | 544.9 | 545.0 | 545.0 | 545.0 | 545.0 | 548.2 | 549.9 | 550.8 | 551.0 | 551.1 | 551.8 |
| MAE | 292.4 | 292.4 | 292.4 | 292.4 | 292.6 | 293.0 | 293.0 | 293.0 | 296.2 | 299.3 | 298.7 | 298.7 | 298.7 | 300.3 |
| MAPE | 2.90 | 2.80 | 2.11 | 2.01 | 2.54 | 2.52 | 3.72 | 4.19 | 2.95 | 3.08 | 3.08 | 2.52 | 3.20 | 3.94 |

Note: (1) * represents that the Chebyshev polynomial is the best model; (2) the listed models are representatives from the set of models.
April 2020, the number of total confirmed new cases and the number of total new deaths had consequently entered the second periodic peaks in July 2020 and August 2020. Furthermore, since the 2020 US presidential election season was beginning in August 2020, the number of total confirmed new cases and the number of total new deaths had then entered the third periodic peaks in December 2020 and January 2021.

It is a matter of great urgency that the final scale of the disease for the forthcoming period is forecast appropriately. A reliable forecast is very important for policymakers to allocate medical resources and step the economic recovery appropriately. The generalized exponential model (EXPM) is a special generalized growth model (GGM), it has been used to characterize and forecast the early epidemic growth patterns across a diversity of outbreaks of many diseases due to its simplicity and reliable performance (Wu et al., 2020). This paper has combined the EXPM model with the Chebyshev

| m  | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $R^2$ | 0.558 | 0.559 | 0.560 | 0.561 | 0.561 | 0.561 | 0.566 | 0.561 | 0.561 | 0.563 | 0.565 | 0.566 | 0.565 | 0.566 |
| $R$  | 0.747 | 0.747 | 0.748 | 0.749 | 0.749 | 0.749 | 0.752 | 0.749 | 0.749 | 0.751 | 0.752 | 0.752 | 0.752 | 0.752 |
| Adj. $R^2$ | 0.455 | 0.454 | 0.455 | 0.456 | 0.456 | 0.454 | 0.454 | 0.453 | 0.452 | 0.454 | 0.456 | 0.455 | 0.454 | 0.454 |
| RMSE | 544.3 | 545.7 | 545.8 | 545.9 | 547.4 | 548.8 | 557.1 | 551.7 | 553.2 | 552.5 | 552.5 | 553.9 | 555.5 | 557.1 |
| MAE  | 292.4 | 293.6 | 291.3 | 289.9 | 291.2 | 292.2 | 294.4 | 294.3 | 295.4 | 292.8 | 291.0 | 291.7 | 293.3 | 294.4 |
| MAPE | 2.90  | 2.93  | 2.86  | 2.79  | 2.80  | 2.80  | 2.73  | 2.88  | 2.91  | 2.81  | 2.73  | 2.72  | 2.73  | 2.73  |

**Fig. 10.** Prediction curves of the total new deaths from the selected 35-order Chebyshev polynomial models for the period 2.

**Fig. 11.** Prediction curves of the total deaths from the selected 35-order Chebyshev polynomial models for the period 2.
Table 13
Test results for choosing the best model of the Chebyshev polynomial based on the total new deaths during the period 3.

| q   | 63* | 62  | 61  | 60  | 59  | 57  | 58  | 56  | 55  | 53  | 52  | 51  | 50  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| R²  | 0.597 | 0.597 | 0.597 | 0.597 | 0.597 | 0.596 | 0.596 | 0.595 | 0.595 | 0.592 | 0.592 | 0.592 | 0.591 |
| R   | 0.773 | 0.773 | 0.773 | 0.773 | 0.773 | 0.772 | 0.772 | 0.772 | 0.772 | 0.770 | 0.770 | 0.769 | 0.769 |
| Adj. R² | 0.493 | 0.495 | 0.497 | 0.499 | 0.501 | 0.504 | 0.502 | 0.505 | 0.507 | 0.507 | 0.505 | 0.508 | 0.510 |
| RMSE | 559.4 | 559.5 | 559.5 | 559.5 | 560.3 | 560.3 | 560.7 | 560.8 | 562.8 | 562.9 | 563.1 | 563.2 | 563.7 |
| MAE  | 343.2 | 342.9 | 342.6 | 342.4 | 343.0 | 343.2 | 344.2 | 344.3 | 344.6 | 345.9 | 345.9 | 344.4 |      |
| MAPE | 9.62  | 9.63  | 9.64  | 9.64  | 9.55  | 9.56  | 9.66  | 9.68  | 9.86  | 9.88  | 9.92  | 9.91  | 9.70  |

*Note: (1) * represents that the Chebyshev polynomial is the best model; (2) the listed models are representatives from the set of models.

Table 14
Test results by using the selected 63-order Chebyshev polynomial based on the total new deaths during the period 3.

| m   | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| R²  | 0.597 | 0.598 | 0.592 | 0.621 | 0.614 | 0.617 | 0.626 | 0.638 | 0.647 | 0.642 | 0.634 | 0.627 | 0.632 | 0.640 |
| R   | 0.773 | 0.773 | 0.769 | 0.788 | 0.784 | 0.785 | 0.791 | 0.799 | 0.805 | 0.801 | 0.796 | 0.792 | 0.795 | 0.800 |
| Adj. R² | 0.493 | 0.493 | 0.485 | 0.522 | 0.512 | 0.515 | 0.527 | 0.541 | 0.553 | 0.546 | 0.535 | 0.525 | 0.531 | 0.541 |
| RMSE | 559.4 | 534.6 | 530.8 | 510.7 | 509.9 | 508.4 | 502.9 | 495.7 | 489.6 | 490.3 | 490.4 | 490.7 | 487.7 | 482.7 |
| MAE  | 343.2 | 338.1 | 336.7 | 326.5 | 325.9 | 323.3 | 317.9 | 311.9 | 309.4 | 309.7 | 309.6 | 309.6 | 306.6 | 302.7 |
| MAPE | 9.62  | 8.74  | 8.37  | 0.36  | 0.36  | 0.36  | 0.35  | 0.35  | 0.36  | 0.37  | 0.36  | 0.36  | 0.36  | 0.36 |

Fig. 12. Prediction curves of the total new deaths from the selected 63-order Chebyshev polynomial models for the period 3.

Fig. 13. Prediction curves of the total deaths from the selected 63-order Chebyshev polynomial models for the period 3.
polynomial together and built a special GGM model to predict the daily numbers of total new confirmed cases and total new deaths in the US for three periods under a 14-day sensitivity regression patterns.

The current and future situations of the COVID-19 pandemic in the US are presented as 14 scenarios (m = 0 to 13). This study provides a range of predictions about the prevalence of COVID-19 in the US under the uncertain effects of “opening up American again” and the 2020 US presidential election season. The prediction results for period one and period two have approved that the curves of the prediction values can cover the curves of the sample total new confirmed cases and the sample total new deaths.

Based on the prediction results for the period three, from the perspective of optimism and COVID-19 vaccine, the total confirmed new cases and total new deaths of COVID-19 in the US will decrease and reach the lowest level in July 2021; until to 1 July 2021, the daily total confirmed new cases will decrease to 3028 cases, the daily total new deaths will decrease to 1958 cases, the total confirmed cases will increase to 38,102,565 cases, and the total deaths will increase to 771,862 cases.

Declaration of competing interest

I Kejia Yan, declare that this paper, submitted in fulfillment of the requirement for the Infectious Disease Modelling. This work has not previously been submitted in any journal. To the best of my knowledge and belief, the dissertation contains no material previously published or written by another person except where due reference is made in the dissertation itself.

We know of no conflicts of interest associates with this paper, and there has been no significant financial support for this work that could have influenced its outcome. As Corresponding Author, I confirm that the manuscript has been read and approved for submission by all named authors.

References

Ahmar, A. S., & Val, E. B. D. (2020). SutteARIMA: Short-term forecasting method, a case: COVID-19 and stock market in Spain. The Science of the Total Environment, 729, 138883.

Beck, B., Shin, B., Choi, Y., Park, S., & Kang, K. (2020). Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. Computational and Structural Biotechnology Journal, 18, 784–790. https://doi.org/10.1016/j.csbj.2020.03.025

Burger, R., Chowell, G., & Lara-Diaz, L. Y. (2019). Comparative analysis of phenomenological growth models applied to epidemic outbreaks. Mathematical Biosciences and Engineering, 16(5), 4250–4273.

Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. Veterinary Research, 38, 281–297.

Cdc. (2020). Coronavirus disease (COVID-19) Cases in the U.S. Available on May 17, 2020 at the website: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html.

Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. Infectious Disease Modelling, 2(3), 379–398. https://doi.org/10.1016/j.idm.2017.08.001

Coutsoulas, G., & Yamey, G. (2020). Political interference in public health science during covid-19. BMJ, 371, 19–20.

He, D., Zhao, S., Li, Y., et al. (2020). Comparing COVID-19 and the 1918-19 influenza pandemics in the United Kingdom. International Journal of Infectious Diseases, 98, 67–70.

Hsu, P. R., & Goshu, A. T. (2013). Solutions of rate-state equation describing biological growths. American Journal of Mathematics and Statistics, 3(6), 305–311.

Kuiken, T., Fouchier, R. A. M., Schutten, M., et al. (2003). Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet, 362, 263–270. https://doi.org/10.1016/S0140-6736(02)13967-0.

Liu, Y. X., Ng, Y. L., Tam, J. P., & Liu, D. X. (2016). Human coronaviruses: A review of virus-host interactions. Diseases, 4(3).

Marzia, L., Egidio, B., Andrea, A., Federico, M., Fabio, C., & Gianluca, T. (2020). Delayed access or provision of care in Italy resulting from fear of COVID-19. Lancet, 4(5). https://doi.org/10.1016/S2535-4642(20)30108-5

Pulejo, M., & Querubín, P. (2021). Electoral concerns reduce restrictive measures during the COVID-19 pandemic. Journal of Public Economics, 198, 104387.

Ran, J., Zhao, S., Han, L., Chong, M. K. C., Qiu, Y., Yang, Y., Wang, J., Wu, Y. J. M., Wang, M., & H, D. (2021). The changing patterns of COVID-19 transmissibility during the social unrest in the United States: A nationwide ecological study with a before-and-after comparison. One Health, 12, 100201.

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

US Bureau of Labor Statistics report. (2020). U.S. Bureau of Labor Statistics report, 2020. USDL-20-0815. Available on May 17, 2020 at the website: https://www.bls.gov/news.release/pdf/empSit.pdf.

Viboud, C., Simonsen, L., & Chowell, G. (2016). A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. Epidemiics, 15, 27–37. https://doi.org/10.1093/epidemics/epw016

Weiss, S. R., & Ledovitz, J. L. (2011). Coronavirus pathogenesis. Advances in Virus Research, 81, 85–164.

Whitehouse. (2020). Guidelines: Opening up amercia again. Whitehouse, April 16, 2020. On 15 January 2021 available at: https://www.whitehouse.gov/openingamerica/.

WHO. (2003). Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. World Health Organization. Dec 31, 2003. Available on May 17, 2020 at the website: https://wwwwhooint/csr/sars/country/table2004_04_21/en/.

WHO. (2019). Middle East respiratory syndrome coronavirus (MERS-CoV). World Health Organization. November, 2019. Available on May 17, 2020 at the website: http://www.who.int/emergencies/mers-cov/en/.

WHO. (2020a). Coronavirus disease (COVID-19) situation reports. World Health Organization. Available on website: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

WHO. (2020b). Coronavirus disease (COVID-19) situation report-111. World health organization. Available on May 17, 2020 at the website: https://www.who.int/covid19.

WHO. (2021). Weekly epidemiological update - 5 January 2021. World Health Organization. January 5, 2021. On 15 January 2021 available at: https://www.who.int/publications/m/item/weekly-epidemiological-update---5-january-2021.

Wu, K., Darcet, D., Wang, Q., & Somette, D. (2020). Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. MedRxiv. Available on May 17, 2020 at the website: https://doi.org/10.1101/2020.03.11.20034363.
Zaki, A. M., Boheemen, S. V., Bestebroer, T. M., Osterhaus, A. D., & Fouchier, R. A. (2003). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine, 367*, 1814–1820.

Zhao, S., Cao, P., Gao, D., Zhuang, Z., Cai, Y., Ran, J., Chong, M. K. C., Wang, K., Lou, Y., Wang, W., Yang, L., He, D., & Wang, M. H. Serial interval in determining the estimation of reproduction number of the novel coronavirus disease (COVID-19) during the early outbreak. *Journal of Travel Medicine, 27*(3). doi: 10.1093/jtm/taaa033.

Zhao, S., Zhuang, Z., Cao, P., Ran, J., Gao, D., Lou, Y., Yang, L., Cai, Y., Wang, W., He, D., & W, M. H. (2020). Quantifying the association between domestic travel and the exportation of novel coronavirus (2019–nCoV) cases from Wuhan, China in 2020: A correlational analysis. *Journal of Travel Medicine, 27*(2). https://doi.org/10.1093/jtm/taaa022

Zhao, Z., Li, X., Liu, F., Zhu, G., Ma, C., & Wang, L. (2020). Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *The Science of the Total Environment, 25*, 138959. https://doi.org/10.1016/j.scitotenv.2020.138959