Mathematical model to predict COVID-19 mortality rate

Melika Yajada, Mohammad Karimi Moridani*, Saba Rasouli

Department of Biomedical Engineering, Faculty of Health, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 1 August 2022
Received in revised form 5 November 2022
Accepted 8 November 2022
Available online 13 November 2022
Handling Editor: Dr Y. Shao

Keywords:
Covid-19
Curve fitting
Mathematical modeling
Mortality
Prediction

ABSTRACT

Objective: Covid-19 is a highly contagious viral infection that has recently become a pandemic. Since the beginning of the pandemic, the disease has affected millions of people and taken many people’s lives. The purpose of this paper is to predict and compare the number of cases and mortality rate due to Covid-19 every quarter in 2020 and 2021 in three countries: Iran, the United States, and South Korea.

Materials and methods: The data of this study include the mortality rate of different countries of the world due to Covid-19, which has been approved by the World Health Organization (WHO). In this paper, to develop the mathematical model for mortality rate prediction, the data of the countries of Iran, the United States, and South Korea during the last two years from March 1, 2020, to March 1, 2022, have been used. In addition, the mortality trend was modeled using the MATLAB software toolbox version 2022b. During modeling, six methods including Fourier, Interpolant, Gaussian, Polynomial, Sum of Sine, and Smoothing Spline were implemented. Root Mean square error (RMSE) and final prediction error were used to evaluate the performance of these proposed methods.

Results: As a result of the analysis, it was shown that the Smoothing Spline model with the lowest error rate was capable of accurately evaluating and predicting Covid-19 incidence and mortality rate. Using RMSE, a prediction of the Covid-19 mortality rate for three countries is $3.76498 \times 10^{-5}$. The values of R-Square and Adj R-sq were 1 in all the experiments, which indicates the full compliance of the prediction model.

Conclusion: Using the proposed method, the incidence rate and mortality rate can be properly assessed and compared with each other in three countries. This provides a better view of the progression of the coronavirus outbreak in spring, summer, autumn, and winter. By using the proposed method, governments will be able to prevent disease and alert people to follow health guidelines more closely, thereby reducing infection numbers and mortality rates.

© 2022 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

The SARS Covid 2 virus causes severe respiratory syndrome in humans. From three cases detected on December 27, 2019, in Wuhan, China, the Covid-19 epidemic has spread rapidly around the world, reaching more than 267 million cases.
worldwide by the end of December 2021 and killing over 3.5 million people. SARS Covid 2 spreads from person to person more easily than SARS Covid, and reached almost every continent, resulting in the declaration of a public health emergency on January 30, 2020, and the WHO international concern. The disease was named Covid-19, which is a combination of coronavirus and the year 2019 (Coronavirus disease, 2019; st known case of coronavirus, 2020). Analysis of time series is widely used for various purposes, such as forecasting, event detection, and decision making. Research in time series forecasting particularly plays an important role in econometrics and operational research (Alghamdi et al., 2019).

Modeling tools can be used to estimate short-term and long-term care requirements to plan the number of materials and resources needed during the outbreak. To manage effective and timely medical care and overcome epidemics, public health officials must calculate the expected disease burden. Furthermore, such estimates can be used to guide the type and intensity of interventions required to reduce prevalence (Zhang et al., 2020).

In the past, a variety of high-accuracy statistical methods were used to predict various outcomes. In recent years, several statistical methods have been used to predict epidemics, like time series models, multivariate linear regression, grey prediction models, post-diffusion neural networks, and simulation models. Thus, the general prevalence is determined based on inclinations and randomness. Due to this, the mentioned statistical tools are insufficient to analyze random pandemics, and it is difficult to generalize the results. Autoregressive integrated moving average (ARIMA) models have been used successfully in the past in the health field and various other areas due to their simple structure, quick application, and ability to describe a data set. It has been used with success in the past to estimate the incidence, prevalence, and mortality of infectious diseases such as influenza, malaria, hepatitis, and others (Kurbalija et al., 2014; Liu et al., 2019; Nosseie et al., 2013; Orbann et al., 2017; Ren et al., 2013; Thomson et al., 2006; Wang et al., 2018; Zhang et al., 2013, 2017).

In order to predict the prevalence of Covid-19 in China, Li and Feng used an SEIR (Susceptible Exposed Infectious Recovered) model and data-driven analysis (Li et al., 2020a). Roosa et al. used valid phenomenological models and analyzed short-term predictions of the cumulative number of confirmed cases during the outbreak in Hubei, China. The team also found that Containment strategies implemented in China had successfully reduced transmission and that the epidemic has slowed down lately (Roosa et al., 2020). Fanelli and Piazza analyzed the time dynamics of the Covid-19 epidemic in China, Italy, and France. Furthermore, they calculated the susceptible infected recovered deceased (SIRD) model prediction, which had been modified by an expectation of fading infectivity after lockdown, and simulated the effects of severe infection limitations on the epidemic in Italy. It was found that lower infection rate suppressed the epidemic peak (Fanelli & Piazza, 2020). Roda et al. compared SIR (Susceptible, Infected and Recovered) and SEIR standard frameworks for modeling Covid-19 in Wuhan Province, China, and concluded that SIR performed much better in providing information in validated data than SEIR (Roda et al., 2020). In an assessment of the impact of quarantine in Wuhan and its surrounding areas, Wu et al. (2020) predicted the prevalence of Covid-19 nationally and globally, developed a SEIR- Metapopulational model, and simulated the Corona outbreak throughout China. Based on the data collected in publicly available outbreaks, Wang et al. developed an algorithm based on Patient Information Based Algorithm (PIBA) to estimate Covid-19 mortality in real-time (Wang et al., 2020).

Numerous studies have been conducted worldwide to estimate the potential impact of Coronavirus. A few of the major topics are stochastic simulation, Weibull distribution, exponential growth, normal logarithmic distribution, etc. But none of these studies were able to determine the exact reproduction rate. The Gene expression programming (GEP) model is proposed to predict the total number of cases in India based on five main parameters, which include confirmed cases and death rate (Ferreira, 2001). As a result of Ivanov’s simulation, he concluded that closing and opening facilities at various levels would be an important factor in determining the effects of the Covid-19 outbreak on global food supply chains (Ivanov, 2020). In a study by Lee et al. Gaussian distributions were used to analyze the transmission of Covid-19 in Hubei Province, China, and to predict the prevalence of the virus in South Korea, Italy, and Iran. It is evident from the results that the epidemic has evolved; therefore, enforcing controls would have a major impact (Li et al., 2020b). Petropoulos and Makridakis presented an objective method of predicting the continuity of the global Covid-19, based on the exponential smoothing family of models for proper planning and decision making (Petropoulos & Makridakis, 2020). Jia et al. used three mathematical models, including the Gompertz model and the logistics model, to estimate the progress of Covid-19 in Wuhan, China. As a result of these mathematical models, Covid-19 is predicted to end in Wuhan by the end of April 2020 (Jia et al., 2020). Castorina et al. utilized mathematical models, including the Gompertz model and logistics model to evaluate the effectiveness of curbing the Covid-19 epidemic in China, South Korea, Italy, and Singapore; using these models, we can determine the maximum number of infected individuals for each country, to develop a strong containment strategy (Castorina et al., 2020). According to Ahmadi et al. mathematical models were used to examine the predictions of Covid-19 in Iran; the results showed that by adapting and intervening in public behavior, the prevalence of Covid-19 in Iran from April 28, 2020, to July 2020 can be controlled and reduced through the Gompertz model (Ahmadi et al., 2020). Torrealba Rodriguez et al. were able to accurately predict the number of Covid-19 patients by the end of the epidemic by using Gompertz mathematical models and logistic models along with an inverted artificial neural network (ANN) model (Torrealba-Rodriguez et al., 2020).

Based on the reviews and studies mentioned above, different models have been proposed to analyze the prevalence rate, infection, and mortality rate associated with the Covid-19 disease around the world. Among the models studied were the Grey prediction model, ARIMA, SEIR model, SIRD model, Weibull distribution model, exponential growth model, normal logarithmic distribution, GEP model, Gaussian distribution, Gompertz model, and logistics model; we selected a smoothing spline model. This study was aimed at modeling and comparing the incidence and rate of mortality associated with Covid-19 disease in Iran, the United States, and South Korea during the last two years.
There were a variety of methods tried, comprising Fourier, Gaussian, interplant, polynomial, sum of sine, and smoothing spline, each of which is briefly described below. According to the results obtained by MATLAB software and after reviewing them, the best and least error method was selected for the study. The rest of the article is designed as follows:

In section 2, the database and the proposed method of this article, which is based on mathematical models for predicting mortality, are introduced. In section 3, the results obtained from the simulation of the proposed mathematical models are presented and an attempt is made to show a better view of the capabilities of the mathematical models by introducing the error of each model. Section 4 is devoted to discussion and conclusion. In this section, while comparing the proposed method with other recent studies, the applications of the method presented in this study will be discussed.

2. Materials and method

2.1. Database

Data is used in this study to determine the incidence and mortality rate of Covid–19 in Iran, the United States, and South Korea over the last two years, from March 1, 2020, to March 1, 2022. This data has been extracted from the Ministry of Health of Iran and our World in Data website as well (Ritchie et al., 2020).

2.2. Proposed method

In this section, mathematical models are introduced to predict the mortality trend of people due to Covid–19.

2.2.1. Spline

Splines have very special features that have long been the subject of attention of statisticians and mathematicians. Generally, k = 3 corresponds to cubic splines, which are the most common. These functions are continuous piecewise cubic functions and have continuous first and second derivatives. These splines exhibit fine smoothness because they have continuity in all of their lower-order derivatives. In theory, the nodes of a cubic spline can’t be distinguished with the naked eye since they are so smooth (R Tibshirani et al., 2014). Splines are divided into four types: Nominal, Ordinal, Polynomial, and Thin-Plate (Helwig, 2017).

Regression splines are classic tools and can work well as long as good node points t1, ..., tm are selected; But in general, choosing a node is a tricky business. One of the unique characteristics of splines is that they do not require nodes to be selected (R Tibshirani et al., 2014).

Smoothing splines, such as core regression and K-nearest neighbor regression, provides a flexible way to estimate the underlying regression function r(x) = E(Y|X = x). Although it can be defined for larger dimensions, for simplicity, X∈R is assumed throughout this paper, ie there is only one predictor variable. A spline of order k is a piecewise polynomial function of fractions of degree k whose continuous and discontinuous derivatives of the order 1, ..., k-1 at its node points. Formally the function f, R→R is a spline of order k with knot points at t1<m; If f is a polynomial of degree k in each of the intervals (-∞, t1] ...[tn0, +∞) and f (j) is a jth derivative of f, in [tn0, ..., t1 for every j = 0, 1,.. k-1 is continuous (R Tibshirani et al] ....... j [t1, t2] j 2014).

According to the non-parametric regression model in previous studies (Gu, 2013; Wahba, 1990), it is written as follows:

\[
y_i = \eta(x_i) + \epsilon_i \quad (1)
\]

For i = 1, ..., n where yi∈R answer is the real value for viewing the i-th sentence and xi∈X is the predictor for viewing the i-th sentence, where X is the predictor domain and \( \eta \in H \) is a smooth function in which H is an RKHS(Reproducing Kernel Hilbert Space) with inner-product (,). And N (0, σ2) ∈ i Gaussian error term.

A spline smoothing is a function \( \eta_s \), which penalized least squares functional.

\[
\frac{1}{n} \sum_{i=1}^{n} (y_i - \eta(x_i))^2 + \lambda j(\eta) \quad (2)
\]

Where j (.) Is a quadratic penalty function such that larger values correspond to less smoothness, and \( \lambda \geq 0 \) is a smoothing parameter that balances the trade-off between fitting and smoothing of the data. \( H = H_0 \oplus H_1 \) shows the decomposition of the total tensor H where \( H_0 = \{ \eta: J(\eta) = 0 \} \) null space J and \( H_1 = \{ \eta: 0 < j(\eta) < \infty \} \) contrast space.

Similarly, assuming \( \langle . . . \rangle_1 + \langle . . . \rangle_0 = \langle . . . \rangle \) Show the corresponding decomposition of the inner product of H. It is noteworthy that by definition, the quadratic error function j is the inner-product of the contrast space of H1, ie \( J(\eta) = \langle \eta, \eta \rangle_1 \). According to \( \lambda \), Kimeldorf-Wahba’s theorem states (Kimeldorf & Wahba, 1971) that \( \eta_1 \) in the minimization equation (2) is as follows:

\[
\eta_1(x) = \sum_{v=0}^{m-1} d v q_v + \sum_{i=1}^{n} c_i \rho_1 \quad (3)
\]
Where the functions \( \{ \varphi_i \}_{i=1}^{n-1} \) are in the null space of the error \( H_0 \), the function \( \rho_1 \) reproduces the kernel of the contrast space \( H_1 \) and \( d = \{ dv \} \) and \( c = \{ C \} \) are the unknown vectors coefficients (Gu, 2013; Wahba, 1990).

The reproduction property shows that the quadratic error function \( \mathbf{J}(\eta) = (\eta, \eta)_1 \) is as follows:

\[
\mathbf{J}(\eta) = \sum_{i=1}^{n} \sum_{j=1}^{n} C_i C_j \rho_1 (X_i, X_j)
\]

(4)

Where \( \rho_1, c, \) and \( x \) can be easily calculated (Helwig, 2017).

2.2.2. Fourier

The data generation process is considered in Equation (5).

\[
y_t = X_t \delta^2 + Z_t^2 y + \epsilon_t + \epsilon_t
\]

(5)

\[\mathbf{r}_t = \mathbf{r}_{t-1} + \mathbf{u}_t\]

Where \( \epsilon_t \) is Stationary errors and \( \mathbf{u}_t \) is independent and is evenly distributed with variance \( \sigma_u^2 \). Here \( X_t = \text{(Coronavirus disease, 2019)} \) is used for the fixed surface process for \( y_t \) and \( X_t = [1, t] \) for the fixed trend process. \( Z_t = \left( \sin \left( \frac{2k \pi t}{T} \right) + \cos \left( \frac{2k \pi t}{T} \right) \right) \) is selected to capture a break or other type of nonlinear unattended at a deterministic term in which \( k \) is the frequency and \( T \) is the sample size. Based on the null hypothesis, \( \sigma_u^2 = 0 \), so that the process described in Equation (5) is constant. The logic of choice \( Z_t = \left( \sin \left( \frac{2k \pi t}{T} \right) + \cos \left( \frac{2k \pi t}{T} \right) \right) \) is based on the fact that a Fourier expansion can approximate fully integrated functions with any degree of the desired accuracy. In particular, \( z(t) \) is allowed to represent the function with an unknown number of unknown breaks. Regardless of the nature of the breaks under very poor conditions, \( z(t) \) can be approximated to any degree of accuracy by the Fourier series long enough.

\[
a(t) = a_0 + \sum_{k=1}^{n} a_k \sin \left( \frac{2k \pi t}{T} \right) + \sum_{k=1}^{n} b_k \cos \left( \frac{2k \pi t}{T} \right) . n \left( \frac{T}{2} \right)
\]

(6)

Where \( n \) is the number of frequencies in the approximation and \( k \) represents a specific frequency. As noted at the beginning with \( n = 1 \), the approximation can always be improved by using additional frequencies. When it reaches \( n = T \), the fit in \( z(t) \) will be perfect. However, to keep the problem tractable, consider starting with the Fourier approximation using a single-frequency component such that:

\[
a(t) \equiv Z_t y = y_1 \sin \left( \frac{2k \pi t}{T} \right) + y_2 \cos \left( \frac{2k \pi t}{T} \right)
\]

(7)

Where \( k \) represents the frequency selected for the approximation and \( y = [y_1, y_2] \) measures the amplitude and displacement of the frequency component. The desirable feature of Equation (7) is that the standard linear profile appears as a special case with the setting \( y_1 = y_2 = 0 \). It is also concluded that in the event of a structural break, at least one frequency component must be present. Therefore, if the null hypothesis \( y_1 = y_2 = 0 \) can be rejected, the series must have a nonlinear component (Becker et al., 2006).

2.2.3. Gaussian

Gaussian functions are suitable for describing many processes in mathematics, science, and engineering, and make them very useful in the fields of image and signal processing. For example, random noise in a signal from complex physical factors can be simply modeled by Gaussian distribution according to the central limit theorem of probability theory. Another common example in image processing is an airy disk, which results from the diffraction of a finite circular aperture as the point function of an imaging system. Usually, an airy disk is represented by a nearly two-dimensional Gaussian function. Thus, fitting Gaussian functions to experimental data is crucial in many signal processing disciplines. The Gaussian function is as follows:

\[
y = A e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

(8)

This function can be plotted with a symmetrical bell-shaped curve at the center of position \( x = \mu \) where \( A \) is the height of the peak and \( \sigma \) controls its width, and on both sides of the curve (low-amplitude sections), the curve approaches the x-axis. The Gaussian function is used to accurately determine the parameters \( \mu, A, \) and \( \sigma \) (Guo, 2011).
### Table 1
The error in modeling the number of patients with Covid-19 in Iran (2020).

| Name of Fitting Method | RMSE        | R-Square  | Adj R-sq |
|------------------------|-------------|-----------|----------|
| **Spring**             |             |           |          |
| Fourier                | 177.5738    | 0.9431    | 0.9300   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 161.4425    | 0.9568    | 0.9421   |
| Polynomial             | 283.9764    | 0.8387    | 0.8210   |
| Sum of Sine            | 170.1688    | 0.9520    | 0.9357   |
| Smoothing Spline       | $1.188 \times 10^{-5}$ | 1 | 1 |
| **Summer**             |             |           |          |
| Fourier                | 156.2208    | 0.7535    | 0.6969   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 143.1479    | 0.8098    | 0.7455   |
| Polynomial             | 187.0390    | 0.6084    | 0.5655   |
| Sum of Sine            | 143.4187    | 0.8091    | 0.7445   |
| Smoothing Spline       | $7.653 \times 10^{-6}$ | 1 | 1 |
| **Autumn**             |             |           |          |
| Fourier                | 335.0015    | 0.9945    | 0.9932   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 333.4411    | 0.9950    | 0.9932   |
| Polynomial             | 389.6210    | 0.9917    | 0.9908   |
| Sum of Sine            | 262.2311    | 0.9969    | 0.9958   |
| Smoothing Spline       | $2.782 \times 10^{-5}$ | 1 | 1 |
| **Winter**             |             |           |          |
| Fourier                | 268.0796    | 0.9839    | 0.9800   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 222.6074    | 0.9898    | 0.9862   |
| Polynomial             | 260.4753    | 0.9831    | 0.9812   |
| Sum of Sine            | 180.2045    | 0.9933    | 0.9910   |
| Smoothing Spline       | $1.119 \times 10^{-5}$ | 1 | 1 |

NaN: Not a Number.

### Table 2
The error in modeling the number of patients with Covid-19 in Iran (2021).

| Name of Fitting Method | RMSE        | R-Square  | Adj R-sq |
|------------------------|-------------|-----------|----------|
| **Spring**             |             |           |          |
| Fourier                | 70.9454     | 0.6814    | 0.6082   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 58.9311     | 0.7980    | 0.7297   |
| Polynomial             | 70.8828     | 0.6476    | 0.6089   |
| Sum of Sine            | 41.6227     | 0.8992    | 0.8651   |
| Smoothing Spline       | $3.4036 \times 10^{-6}$ | 1 | 1 |
| **Summer**             |             |           |          |
| Fourier                | 190.2382    | 0.9023    | 0.8799   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 181.0192    | 0.9187    | 0.8912   |
| Polynomial             | 194.4279    | 0.8869    | 0.8745   |
| Sum of Sine            | 96.4386     | 0.9769    | 0.9691   |
| Smoothing Spline       | $7.9447 \times 10^{-6}$ | 1 | 1 |
| **Autumn**             |             |           |          |
| Fourier                | 331.7818    | 0.8460    | 0.8101   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | 412.4866    | 0.7783    | 0.7065   |
| Polynomial             | 388.0771    | 0.7662    | 0.7402   |
| Sum of Sine            | 252.1184    | 0.9184    | 0.8904   |
| Smoothing Spline       | $1.3242 \times 10^{-5}$ | 1 | 1 |
| **Winter**             |             |           |          |
| Fourier                | $7.3467 \times 10^{-3}$ | 0.9797    | 0.9749   |
| Interpolant            | NaN         | 1         | NaN      |
| Gaussian               | $5.2280 \times 10^{-3}$ | 0.9906    | 0.9873   |
| Polynomial             | $7.5698 \times 10^{-3}$ | 0.9761    | 0.9734   |
| Sum of Sine            | $7.0163 \times 10^{-3}$ | 0.9830    | 0.9771   |
| Smoothing Spline       | 0.0523      | 1         | 1        |
Fig. 1. The barplot of error in modeling the number of patients with Covid-19 in Iran (2020).

Fig. 2. The barplot of error in modeling the number of patients with Covid-19 in Iran (2021).
2.2.4. Sum of sine

The sine sum model is used for periodic functions and is a linear combination consisting of the sum of sine functions with constant coefficients. The unit profile can be defined by Equation (9).

\[ y_j = \sum_{i=1}^{s} a_i \sin(b_i x_j + c_i) + \epsilon_j, \quad i = 1, \ldots, s, \quad j = 1, \ldots, n \]  

(9)

Where \( j \) represents the number of observations, \( i \) represents the number of sine functions, and \( \epsilon \) represents the error expression; also \( a_i \) is the amplitude, \( b_i \) frequency, and \( c_i \) are the horizontal phase constant in each expression of the sine wave, and \( s \) is the number of series expressions generally \( 1 \leq s \leq 8 \). These unknown parameters, including \( a_i \)'s, \( b_i \)'s and \( c_i \)'s must be calculated using the nonlinear least squares estimation method.

The sum of sine functions serving as a fitting model has the advantage of adapting more accurate aesthetic and geometrically results. In addition, it’s fitting to nonlinear data, due to a variety of functional forms, is often performed with accurate results. The sum of sine functions also takes into account the variations between different points in the data. The

---

Table 3

The error in modeling the number of patients with Covid-19 in the USA (2020).

| Name of Fitting Method | RMSE | R-Square | Adj R-sq |
|------------------------|------|----------|----------|
| Spring                 |      |          |          |
| Fourier                | 2.4843 \times 10^{-3} | 0.9602 | 0.9510 |
| Interpolant            | NaN  | 1        | NaN      |
| Gaussian               | 1.9716 \times 10^{-3} | 0.9770 | 0.9692 |
| Polynomial             | 2.8078 \times 10^{-3} | 0.9436 | 0.9375 |
| Sum of Sine            | 1.6591 \times 10^{-3} | 0.9837 | 0.9782 |
| Smoothing Spline       | 1.3723 \times 10^{-4} | 1      | 1        |
| Summer                 |      |          |          |
| Fourier                | 5.5944 \times 10^{-3} | 0.9021 | 0.8795 |
| Interpolant            | NaN  | 1        | NaN      |
| Gaussian               | 5.5715 \times 10^{-3} | 0.9107 | 0.8805 |
| Polynomial             | 5.8283 \times 10^{-3} | 0.8822 | 0.8693 |
| Sum of Sine            | 3.5782 \times 10^{-3} | 0.9632 | 0.9507 |
| Smoothing Spline       | 3.0655 \times 10^{-4} | 1      | 1        |
| Autumn                 |      |          |          |
| Fourier                | 1.341 \times 10^4    | 0.9458 | 0.9332 |
| Interpolant            | NaN  | 1        | NaN      |
| Gaussian               | 1.355 \times 10^4    | 0.9492 | 0.9317 |
| Polynomial             | 1.2912 \times 10^{-4} | 0.9442 | 0.9380 |
| Sum of Sine            | 1.0350 \times 10^{-4} | 0.9703 | 0.9602 |
| Smoothing Spline       | 0.0010 | 1      | 1        |
| Winter                 |      |          |          |
| Fourier                | 2.7696 \times 10^{-4} | 0.8613 | 0.8285 |
| Interpolant            | NaN  | 1        | NaN      |
| Gaussian               | 2.6887 \times 10^{-4} | 0.8765 | 0.8384 |
| Polynomial             | 2.7768 \times 10^{-4} | 0.8451 | 0.8276 |
| Sum of Sine            | 2.0199 \times 10^{-4} | 0.9324 | 0.9088 |
| Smoothing Spline       | 0.0021 | 1      | 1        |

Fig. 3. Covid-19 incidence modeling during four seasons in Iran.
error minimization squares criterion is used to determine the best model. Normally, the coefficient of determination should be high enough to show a good result (Fan et al., 2012).

2.2.5. Polynomial

A polynomial method is a powerful tool for creating results in different regions by constructing an adequate polynomial. Due to the nature of the polynomial method, many arguments related to this method require that everything that happens in the types of polynomial productions be studied inductively. For this reason, it can lead to the study of how the polynomial method is compatible with a variety of general algebraic equations, not only if this natural arrangement is difficult, but also even if the question originally occurs in a constant variety such as \( \mathbb{R}^n \) (Walsh, 2020).

2.2.6. Interpolation

Spline interpolation is a useful and powerful tool in curve and surface design. In general, shared spline interpolation is a fixed interpolation, meaning that the curve or surface interpolation shape is fixed for the given interpolation data. If someone wants to change the shape, the interpolation data must change. Being able to change the shape under conditions where the given data does not change is in itself a big problem in computer-aided geometric design. Theoretically, the uniqueness of the interpolation function for the given interpolation data is contradictory. In recent years, rational splines, especially cube

Table 4
The error in modeling the number of patients with Covid-19 in the USA (2021).

| Name of Fitting Method | RMSE     | R-Square | Adj R-sq |
|------------------------|----------|----------|----------|
| Spring Fourier         | 1.0234 \times 10^{-4} | 0.7646   | 0.7105   |
| Interpolant Fourier    | NaN      | 1        | NaN      |
| Gaussian               | 8.9842 \times 10^{-3} | 0.8333   | 0.7769   |
| Polynomial             | 1.0107 \times 10^{-4} | 0.7456   | 0.7176   |
| Sum of Sine Fourier    | 5.1824 \times 10^{-3} | 0.9445   | 0.9258   |
| Smoothing Spline       | 5.6440 \times 10^{-4} | 1        | 1        |
| Summer                 | 4.4705 \times 10^{-4} | 0.6693   | 0.5933   |
| Interpolant Fourier    | NaN      | 1        | NaN      |
| Gaussian               | 4.2514 \times 10^{-4} | 0.7252   | 0.6322   |
| Polynomial             | 4.3008 \times 10^{-4} | 0.6608   | 0.6236   |
| Sum of Sine Fourier    | 1.6013 \times 10^{-4} | 0.9610   | 0.9478   |
| Smoothing Spline       | 0.0031   | 1        | 1        |
| Autumn                 | 4.8151 \times 10^{-4} | 0.3608   | 0.2120   |
| Interpolant Fourier    | NaN      | 1        | NaN      |
| Gaussian               | 4.7997 \times 10^{-4} | 0.4171   | 0.2170   |
| Polynomial             | 4.7303 \times 10^{-4} | 0.3155   | 0.2395   |
| Sum of Sine Fourier    | 2.3658 \times 10^{-4} | 0.8384   | 0.8098   |
| Smoothing Spline       | 0.0034   | 1        | 1        |
| Winter                 | 2.7606 \times 10^{-4} | 0.8613   | 0.8285   |
| Interpolant Fourier    | NaN      | 1        | NaN      |
| Gaussian               | 2.6887 \times 10^{-4} | 0.8765   | 0.8384   |
| Polynomial             | 2.7768 \times 10^{-4} | 0.8451   | 0.8276   |
| Sum of Sine Fourier    | 2.0199 \times 10^{-4} | 0.9324   | 0.9088   |
| Smoothing Spline       | 0.0021   | 1        | 1        |

Fig. 4. Covid-19 incidence modeling during four seasons in the USA.
splines with quadratic or cubic denominators, and their application in shape control, such as controlling curves to be in a given region, have been considered. Since the parameters of the interpolation function can be selected according to the control needs, finite shape control becomes possible. In this case, there are different interpolation curves for a given data set by selecting different parameters, this does not contradict the uniqueness of the given interpolation function. The uniqueness

| Name of Fitting Method | Spring | Summer | Autumn | Winter |
|------------------------|--------|--------|--------|--------|
| Fourier                | 44.4817| 25.7214| 33.7362| 91.6483|
| Interpolant            | NaN    | NaN    | NaN    | NaN    |
| Gaussian               | 30.2353| 23.3411| 24.2474| 25.7214|
| Polynomial             | 46.5698| 27.6213| 35.7181| 194.2479|
| Sum of Sine            | 46.1638| 23.9944| 27.8371| 75.2417|
| Smoothing Spline       | $3.73 \times 10^{-6}$ | $1.691 \times 10^{-6}$ | $1.815 \times 10^{-6}$ | $5.59 \times 10^{-6}$ |
| Fourier                | 70.9454| 190.2382| 331.7818| 37.3467|
| Interpolant            | NaN    | NaN    | NaN    | NaN    |
| Gaussian               | 58.9311| 181.0192| 412.4866| 25.7214|
| Polynomial             | 70.8828| 194.2479| 388.0771| 70.8828|
| Sum of Sine            | 41.6227| 27.8371| 96.4386| 252.1184|
| Smoothing Spline       | $3.4036 \times 10^{-6}$ | $1.3242 \times 10^{-5}$ | $7.9447 \times 10^{-6}$ | $7.3467 \times 10^{-3}$ |

Table 5
The error in modeling the number of patients with Covid-19 in South Korea (2020).

| Name of Fitting Method | RMSE   | R-Square | Adj R-sq |
|------------------------|--------|----------|----------|
| Spring                 | 44.4817| 0.9270   | 0.9102   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 30.2353| 0.9690   | 0.9585   |
| Polynomial             | 46.5698| 0.9113   | 0.9016   |
| Sum of Sine            | 46.1638| 0.9277   | 0.9033   |
| Smoothing Spline       | $3.73 \times 10^{-6}$ | 1        | 1        |
| Summer                 | 25.7214| 0.9509   | 0.9396   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 23.3411| 0.9628   | 0.9502   |
| Polynomial             | 27.6213| 0.9372   | 0.9303   |
| Sum of Sine            | 23.9944| 0.9607   | 0.9474   |
| Smoothing Spline       | $1.691 \times 10^{-6}$ | 1        | 1        |
| Autumn                 | 33.7362| 0.9347   | 0.9195   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 24.2474| 0.9690   | 0.9584   |
| Polynomial             | 35.7181| 0.9188   | 0.9098   |
| Sum of Sine            | 27.8371| 0.9592   | 0.9452   |
| Smoothing Spline       | $1.815 \times 10^{-6}$ | 1        | 1        |
| Winter                 | 91.6483| 0.8964   | 0.8720   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 93.7268| 0.9007   | 0.8661   |
| Polynomial             | 93.1003| 0.8813   | 0.8679   |
| Sum of Sine            | 75.2417| 0.9360   | 0.9137   |
| Smoothing Spline       | $5.59 \times 10^{-6}$ | 1        | 1        |

Table 6
The error in modeling the number of patients with Covid-19 in South Korea (2021).
of the interpolation function for the given data is replaced by the uniqueness of the interpolation curve for the given data and the selected parameters. For this type of interpolation, function values and derivatives are used in nodes. Unfortunately, in some manufacturing processes, derivatives are difficult to obtain, which is why most papers use an illustrated cube spline with a quadratic denominator based solely on function values (Qi et al., 2005).

3. Simulation results

Fourier, Interpolant, Gaussian, polynomial, the sum of sine, and smoothing spline functions have been tested to model and compare the incidence and mortality rate of Covid-19 over the past two years. Table 1 shows the error calculated during modeling by six different functions in Iran from March 1, 2020, to March 1, 2021, for the number of incidents. As you can see in Table 1 and Table 2, the least free function is the smoothing spline with the root mean square error of $1.188 \times 10^{-5}$ and $3.4036 \times 10^{-6}$, respectively for the spring season of 2020 and 2021, by which the modeling is performed.

Fig. 1 shows the barplot of the errors that include RMSE, R-Square, Adj R-sq; of incidence in the four seasons of spring, summer, autumn, and winter in Iran during the two years 2020.

### Table 7
The error of Covid-19 mortality rate modeling in Iran (2020).

| Name of Fitting Method | RMSE    | R-Square | Adj R-sq |
|------------------------|---------|----------|----------|
| Spring                 |         |          |          |
| Fourier                | 10.2503 | 0.9451   | 0.9324   |
| Interpolant            | NaN     | 1        | NaN      |
| Gaussian               | 8.9531  | 0.9615   | 0.9485   |
| Polynomial             | 11.3661 | 0.9242   | 0.9169   |
| Sum of Sine            | 8.7367  | 0.9633   | 0.9509   |
| Smoothing Spline       | 6.508 $\times 10^{-7}$ | 1 | 1 |
| Summer                 |         |          |          |
| Fourier                | 14.4304 | 0.9272   | 0.9104   |
| Interpolant            | NaN     | 1        | NaN      |
| Gaussian               | 13.9286 | 0.9376   | 0.9166   |
| Polynomial             | 14.9613 | 0.9122   | 0.9037   |
| Sum of Sine            | 12.9933 | 0.9457   | 0.9274   |
| Smoothing Spline       | 1.412 $\times 10^{-6}$ | 1 | 1 |
| Autumn                 |         |          |          |
| Fourier                | 19.3336 | 0.9815   | 0.9772   |
| Interpolant            | NaN     | 1        | NaN      |
| Gaussian               | 96.3613 | 0.7817   | 0.7067   |
| Polynomial             | 19.0066 | 0.9799   | 0.9780   |
| Sum of Sine            | 16.7472 | 0.9873   | 0.9829   |
| Smoothing Spline       | 1.395 $\times 10^{-6}$ | 1 | 1 |
| Winter                 |         |          |          |
| Fourier                | 12.5491 | 0.9817   | 0.9774   |
| Interpolant            | NaN     | 1        | NaN      |
| Gaussian               | 11.0360 | 0.9870   | 0.9825   |
| Polynomial             | 12.1425 | 0.9807   | 0.9788   |
| Sum of Sine            | 10.6285 | 0.9880   | 0.9838   |
| Smoothing Spline       | 7.9 $\times 10^{-7}$ | 1 | 1 |
Based on the four seasons of spring, summer, autumn, and winter in Iran during the two years, Fig. 2 displays the error bar plots including RMSE, R-Square, and adjusted R-Square.

In Fig. 3, the modeling of the incidence during spring, summer, autumn, and winter in Iran between 2020 and 2021 is shown.

Similar to Iran, there are Table 3 and Table 4 for the United States, show the error in modeling the incidence of Covid-19 from March 1, 2020, to March 1, 2021. As can be seen, the smoothing spline function with the root mean square error of 3.0655 $\times 10^{-6}$ and 0.0031 has the lowest error rates for the summer seasons of 2020 and 2021, respectively.

It is clear that the smoothing spline function is the least error function selected for modeling. Fig. 4 shows the modeling of Covid-19 infection in the four seasons of spring, summer, fall, and winter in the United States during the two years 2020 and 2021.

Table 5 and Table 6 show the modeling error of the Covid-19 infected number in South Korea by six different functions. According to the tables, the smoothing spline function with the root mean square error that respectively is $1.815 \times 10^{-6}$ and 0.0031 has the lowest error rates for the autumn season of 2020 and 2021, this model has the least error.

Fig. 5 shows the modeling of patient’s number infected by Covid-19 using the smoothing spline function in the four seasons of spring, summer, autumn, and winter during the two years 2020 and 2021 in South Korea.

Table 8
The error of Covid-19 mortality rate modeling in Iran (2021).

| Name of Fitting Method | RMSE          | R-Square | Adj R-sq |
|------------------------|---------------|----------|----------|
| Spring                 |               |          |          |
| Fourier                | 29.7950       | 0.9548   | 0.9444   |
| Interpolant            | NaN           | 1        | NaN      |
| Gaussian               | 23.1534       | 0.9749   | 0.9664   |
| Polynomial             | 29.1071       | 0.9522   | 0.9470   |
| Sum of Sine            | 22.6587       | 0.9760   | 0.9679   |
| Smoothing Spline       | $1.7694 \times 10^{-6}$ | 1        | 1        |
| Summer                 |               |          |          |
| Fourier                | 35.5047       | 0.9733   | 0.9672   |
| Interpolant            | NaN           | 1        | NaN      |
| Gaussian               | 28.2990       | 0.9844   | 0.9792   |
| Polynomial             | 35.5661       | 0.9704   | 0.9671   |
| Sum of Sine            | 27.9278       | 0.9848   | 0.9797   |
| Smoothing Spline       | $2.6673 \times 10^{-6}$ | 1        | 1        |
| Autumn                 |               |          |          |
| Fourier                | 24.4148       | 0.9782   | 0.9731   |
| Interpolant            | NaN           | 1        | NaN      |
| Gaussian               | 22.4583       | 0.9828   | 0.9722   |
| Polynomial             | 25.4143       | 0.9737   | 0.9708   |
| Sum of Sine            | 22.7455       | 0.9826   | 0.9766   |
| Smoothing Spline       | $1.7416 \times 10^{-6}$ | 1        | 1        |
| Winter                 |               |          |          |
| Fourier                | 8.7155        | 0.9858   | 0.9825   |
| Interpolant            | NaN           | 1        | NaN      |
| Gaussian               | 8.8531        | 0.9860   | 0.9819   |
| Polynomial             | 9.1339        | 0.9827   | 0.9808   |
| Sum of Sine            | 7.6448        | 0.9900   | 0.9865   |
| Smoothing Spline       | $6.4245 \times 10^{-7}$ | 1        | 1        |

Fig. 6. Covid-19 mortality rate modeling during four seasons in Iran.
Table 9
The error of Covid-19 mortality rate modeling in the USA (2020).

| Name of Fitting Method | RMSE   | R-Square | Adj R-sq |
|------------------------|--------|----------|----------|
| Spring                 |        |          |          |
| Fourier                | 268.8240 | 0.9195   | 0.9010   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 206.1591 | 0.9565   | 0.9418   |
| Polynomial             | 269.7810 | 0.9101   | 0.9003   |
| Sum of Sine            | 136.8963 | 0.9808   | 0.9743   |
| Smoothing Spline       | 2.8422 $\times 10^{-13}$ | 1 | 1 |
| Summer                 |        |          |          |
| Fourier                | 174.3172 | 0.7491   | 0.6915   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 215.7224 | 0.6470   | 0.5275   |
| Polynomial             | 265.6180 | 0.3546   | 0.2837   |
| Sum of Sine            | 114.8672 | 0.8999   | 0.8660   |
| Smoothing Spline       | 1.1315 $\times 10^{-5}$ | 1 | 1 |
| Autumn                 |        |          |          |
| Fourier                | 350.5188 | 0.5267   | 0.4164   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 319.0723 | 0.5695   | 0.4217   |
| Polynomial             | 309.6177 | 0.5099   | 0.4555   |
| Sum of Sine            | 145.7350 | 0.9102   | 0.8794   |
| Smoothing Spline       | 1.1595 $\times 10^{-5}$ | 1 | 1 |
| Winter                 |        |          |          |
| Fourier                | 522.2500 | 0.7488   | 0.6895   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 699.9555 | 0.5863   | 0.4422   |
| Polynomial             | 860.6407 | 0.2419   | 0.1566   |
| Sum of Sine            | 347.0984 | 0.8983   | 0.8628   |
| Smoothing Spline       | 2.5994 $\times 10^{-5}$ | 1 | 1 |

Table 10
The error of Covid-19 mortality rate modeling in the USA (2021).

| Name of Fitting Method | RMSE   | R-Square | Adj R-sq |
|------------------------|--------|----------|----------|
| Spring                 |        |          |          |
| Fourier                | 229.1288 | 0.8130   | 0.7700   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 213.0591 | 0.8514   | 0.8012   |
| Polynomial             | 370.7634 | 0.4574   | 0.3979   |
| Sum of Sine            | 212.4805 | 0.8522   | 0.8022   |
| Smoothing Spline       | 1.7996 $\times 10^{-5}$ | 1 | 1 |
| Summer                 |        |          |          |
| Fourier                | 316.4684 | 0.6057   | 0.5151   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 215.3381 | 0.8322   | 0.7755   |
| Polynomial             | 312.6551 | 0.5735   | 0.5267   |
| Sum of Sine            | 143.2135 | 0.9258   | 0.9007   |
| Smoothing Spline       | 1.4822 $\times 10^{-5}$ | 1 | 1 |
| Autumn                 |        |          |          |
| Fourier                | 334.8708 | 0.8885   | 0.8625   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 681.1454 | 0.5765   | 0.4311   |
| Polynomial             | 856.5714 | 0.1903   | 0.1003   |
| Sum of Sine            | 327.9060 | 0.9019   | 0.8682   |
| Smoothing Spline       | 3.9189 $\times 10^{-5}$ | 1 | 1 |
| Winter                 |        |          |          |
| Fourier                | 512.0617 | 0.8258   | 0.7847   |
| Interpolant            | NaN    | 1        | NaN      |
| Gaussian               | 996.4363 | 0.3953   | 0.1846   |
| Polynomial             | 1.0584 $\times 10^{-3}$ | 0.1731 | 0.0800   |
| Sum of Sine            | 508.1565 | 0.8427   | 0.7879   |
| Smoothing Spline       | 5.1312 $\times 10^{-5}$ | 1 | 1 |

Table 7 and Table 8 show the errors calculated during modeling by six different functions in Iran from March 1, 2020, to March 1, 2022, for the mortality rate of Covid-19 in this country in four seasons: spring, summer, autumn, and winter. As can
be seen and similar to the previous one, the smoothing spline function with the root mean square error of $7.9 \times 10^{-7}$ and $6.4245 \times 10^{-7}$ respectively 2020–2021 and 2021–2022, has the lowest error in the winter.

Fig. 6 shows the modeling of the Covid-19 mortality rate using the smoothing spline function in the four seasons of spring, summer, autumn, and winter during the two years 2020 and 2021 in Iran.

Table 9 and Table 10 are available for the United States, showing the prediction error in modeling the Covid-19 mortality rate by the six different functions Fourier, Interplant, Gaussian, polynomial, the sum of sine, and smoothing spline from March 1, 2020, to March 1, 2022. As you can see in the tables, the smoothing spline is the least error function with the root mean square error of $2.8422 \times 10^{-13}$ and $1.7996 \times 10^{-5}$ for the spring seasons of 2020 and 2021, respectively.

According to the above tables, the smoothing spline function was selected with the least error. Fig. 7 shows the modeling of the Covid-19 mortality rate using the smoothing spline function in the four seasons of spring, summer, fall, and winter during the two years 2020 and 2021 in the United States (see Fig. 8).

Tables 11 and 12 show the error calculated during modeling by six different functions: Fourier, Interpolant, Gaussian, polynomial, the sum of sine, and smoothing spline in South Korea from March 1, 2020, to March 1, 2022, for the country’s mortality rate of Covid-19 during four seasons: spring, summer, fall, and winter. As you can see, similar to the previous, the smoothing spline function is the least error with the root mean square error of $4.7791 \times 10^{-8}$ and $1.9858 \times 10^{-7}$ for the summer season of 2020 and 2021, respectively.

According to Table 11 and Table 12, the smoothing spline was selected as the least error and best function. Fig. 6 shows the modeling of the Covid-19 mortality rate using the smoothing spline function in the four seasons of spring, summer, autumn, and winter in 2020 and 2021 in South Korea.

Fig. 9 shows a block diagram of the different stages of implementing the method proposed in this paper. First, the input data is pre-processed and missing values are removed. After the data are cleaned, they are fitted to a curve by different mathematical models, and then according to the lowest amount of prediction error, the best model is determined in terms of prediction.

4. Discussion and conclusion

The spread of Covid-19 has seriously compromised public health around the world. Mathematical models have been used by many researchers in recent years to estimate disease occurrences, weather phenomena, oil prices, and stock market volatility. With the help of mathematical models, this paper attempts to predict the mortality rate of people affected by
Corona virus in 2020 and 2021. Using the results of this study, we can control the disease, reduce deaths, and alert people to follow health protocols such as using face masks, social distancing, isolating and washing hands regularly. Given that there is no definitive vaccine or treatment available at the time of writing, effective prediction of mortality trends, including peak

### Table 11
The error of Covid-19 mortality rate modeling in South Korea (2020).

| Name of Fitting Method | RMSE       | R-Square | Adj R-sq |
|------------------------|------------|----------|----------|
| Spring                 |            |          |          |
| Fourier                | 1.9901     | 0.5534   | 0.4507   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 1.7361     | 0.6877   | 0.5820   |
| Polynomial             | 1.9914     | 0.4984   | 0.4500   |
| Sum of Sine            | 1.7712     | 0.6749   | 0.5650   |
| Smoothing Spline       | 2.1836 × 10⁻⁷ | 1        | 1        |
| Summer                 |            |          |          |
| Fourier                | 0.6672     | 0.4727   | 0.3516   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 0.6201     | 0.5815   | 0.4399   |
| Polynomial             | 0.7434     | 0.2656   | 0.1948   |
| Sum of Sine            | 0.6218     | 0.5791   | 0.4368   |
| Smoothing Spline       | 4.7791 × 10⁻⁸ | 1        | 1        |
| Autumn                 |            |          |          |
| Fourier                | 1.5024     | 0.2286   | 0.0489   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 1.4220     | 0.3657   | 0.1480   |
| Polynomial             | 1.4685     | 0.5815   | 0.4399   |
| Sum of Sine            | 1.2166     | 0.5358   | 0.3764   |
| Smoothing Spline       | 1.4482 × 10⁻⁷ | 1        | 1        |
| Winter                 |            |          |          |
| Fourier                | 4.4340     | 0.7338   | 0.6709   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 3.5138     | 0.8467   | 0.7933   |
| Polynomial             | 4.3498     | 0.7118   | 0.6833   |
| Sum of Sine            | 3.8193     | 0.8189   | 0.7558   |
| Smoothing Spline       | 3.9470 × 10⁻⁷ | 1        | 1        |

### Table 12
The error of Covid-19 mortality rate modeling in South Korea (2021).

| Name of Fitting Method | RMSE       | R-Square | Adj R-sq |
|------------------------|------------|----------|----------|
| Spring                 |            |          |          |
| Fourier                | 2.0287     | 0.2695   | 0.1017   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 1.9744     | 0.3642   | 0.1491   |
| Polynomial             | 2.1161     | 0.1192   | 0.0226   |
| Sum of Sine            | 1.8507     | 0.4414   | 0.2524   |
| Smoothing Spline       | 1.9395 × 10⁻⁷ | 1        | 1        |
| Summer                 |            |          |          |
| Fourier                | 2.1608     | 0.6342   | 0.5501   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 1.8239     | 0.7605   | 0.6795   |
| Polynomial             | 2.2555     | 0.50583  | 0.5809   |
| Sum of Sine            | 2.0853     | 0.6869   | 0.5810   |
| Smoothing Spline       | 1.9858 × 10⁻⁷ | 1        | 1        |
| Autumn                 |            |          |          |
| Fourier                | 4.6356     | 0.8591   | 0.8283   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 4.1458     | 0.8966   | 0.8611   |
| Polynomial             | 4.6195     | 0.8447   | 0.8275   |
| Sum of Sine            | 4.4668     | 0.8799   | 0.8387   |
| Smoothing Spline       | 3.7021 × 10⁻⁷ | 1        | 1        |
| Winter                 |            |          |          |
| Fourier                | 18.4080    | 0.6099   | 0.5179   |
| Interpolant            | NaN        | 1        | NaN      |
| Gaussian               | 17.3109    | 0.6838   | 0.5736   |
| Polynomial             | 17.7491    | 0.5971   | 0.5518   |
| Sum of Sine            | 16.1716    | 0.7241   | 0.6279   |
| Smoothing Spline       | 1.4566 × 10⁻⁶ | 1        | 1        |
timing, is essential. In this paper, we presented a mathematical model of Covid-19 to predict mortality rates in several countries, the results of which can be generalized to other countries.

Studying and modeling the incidence and mortality rate of Covid-19 in different conditions can provide a good understanding of the progression of the disease and prediction of its future. In this paper, the modeling of the number of infected by Covid-19 and the mortality rate due to Covid-19 during the two years 2020 and 2021, four seasons of spring, summer, autumn, and winter, in three countries including Iran, the USA, and South Korea were studied. This study analyzed six models of Fourier, Interplant, Gaussian, Polynomial, Sum of Sine, and Smoothing Spline, which was selected as the least error and best model. The results show that in the cold seasons of the year, i.e., autumn and especially winter, the incidence of Covid-19 has jumped sharply, which is very noticeable in the USA and South Korea diagrams, Figs. 2 and 3, according to the Ministry of Health, the number of new cases on November 27, 2020, and December 3, 2020, has been registered respectively, 14,051 and 13,922. As you can see, the blue chart indicates winter and, although irregular, has a significant slope. It is noteworthy that in Fig. 3 of South Korea’s 2021 infection statistics, while in the three seasons of spring, summer and autumn the chart remains at a low level, but in mid-winter, we encounter a significant increase in the incidence of Covid-19 and the slope of the chart to it is easily visible that on February 2, 2021, the number of new cases was 171,448, while the number of cases on March 1, 2021, was only 341.

Availability of data and materials

The data used in this paper is cited throughout the paper.

Funding

No source of funding for this work.

Authors’ contributions

All authors evenly contributed to the whole work. All authors read and approved the final manuscript.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Acknowledgments

Not applicable.

Abbreviations

| Abbreviation | Full Form |
|--------------|-----------|
| WHO          | World Health Organization |
| RMSE         | Root Mean Square Error |
| ARIMA        | Autoregressive Integrated Moving Average |
| SEIR         | Susceptible Exposed Infectious Recovered |
| SIRD         | Susceptible Infected Recovered Deceased |
| PIBA         | Patient Information Based Algorithm |
| ANN          | Artificial Neural Network |
| GEP          | Gene Expression Programming |
