




## Article

# Models for COVID-19 Daily Confirmed Cases in Different Countries

Hamdy M. Ahmed <sup>1</sup>, Reda A. Elbarkouky <sup>2</sup>, Othman A. M. Omar <sup>2</sup> and Maria Alessandra Ragusa <sup>3,4,\*</sup>

<sup>1</sup> Higher Institute of Engineering, El-Shorouk Academy, El-Shorouk City, Cairo 11837, Egypt; hamdy\_17eg@yahoo.com

<sup>2</sup> Physics and Engineering Mathematics Department, Faculty of Engineering, Ain Shams University, Cairo 11517, Egypt; reda.elbarkouky@eng.asu.edu.eg (R.A.E.); osman\_ahmed@eng.asu.edu.eg (O.A.M.O.)

<sup>3</sup> Dipartimento di Matematica e Informatica, Università di Catania, Viale Andrea Doria 6, 95125 Catania, Italy

<sup>4</sup> RUDN University, 6 Miklukho, Maklay St, 117198 Moscow, Russia

\* Correspondence: mariaalessandra.ragusa@unict.it

**Abstract:** In this paper, daily confirmed cases of COVID-19 in different countries are modelled using different mathematical regression models. The curve fitting is used as a prediction tool for modeling both past and upcoming coronavirus waves. According to virus spreading and average annual temperatures, countries under study are classified into three main categories. First category, the first wave of the coronavirus takes about two-year seasons (about 180 days) to complete a viral cycle. Second category, the first wave of the coronavirus takes about one-year season (about 90 days) to complete the first viral cycle with higher virus spreading rate. These countries take stopping periods with low virus spreading rate. Third category, countries that take the highest virus spreading rate and the viral cycle complete without stopping periods. Finally, predictions of different upcoming scenarios are made and compared with actual current smoothed daily confirmed cases in these countries.



**Citation:** Ahmed, H.M.; Elbarkouky, R.A.; Omar, O.A.M.; Ragusa, M.A. Models for COVID-19 Daily Confirmed Cases in Different Countries. *Mathematics* **2021**, *9*, 659. <https://doi.org/10.3390/math9060659>

Academic Editor: Simeon Reich

Received: 24 February 2021

Accepted: 14 March 2021

Published: 19 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** COVID-19; Fourier fitting; curve fitting; optimization algorithm

## 1. Introduction

The epidemic coronavirus disease (COVID-19) was first detected in December 2019 in Wuhan city, China. Then it started to transmit to countries all over the world till the WHO declared it as a pandemic disease [1] on 11 March 2020. The new virus has different characteristics than the family it came from (coronaviruses). It has higher infection rates and average incubation period of 5.1 days. The maximum incubation period for the virus is 14 days [2]. The epidemic infection occurs through physical contacts with individuals or contaminated surfaces. New public safety steps have been associated to mitigate the impact of this pandemic disease like cancellation of public events, assuring online education, closing clubs and other places opened for social gatherings like concerts and sporting events. The total number of confirmed cases worldwide reached 62,844 million on 1 December 2020 with 1465 million deaths [3]. The main challenge during studying epidemics is how to predict the disease behavior, how many people will be infected in the future, determining the pandemic peak, second wave of the disease time of action, and the total deaths after the pandemic ends. Different researchers from different majors such as applied mathematics, data science, and epidemiology have been working on studying these trends of predictions. Based on this analysis, governments can take proper actions to limit the human and economic losses. Coronavirus not only represents a health crisis but also represents an economic crisis where people are losing their jobs without knowing when normality will return. The International Labor Organization estimates that 25 million people could lose their jobs [4]. The prediction of the peak date from one country to another and simulation of the variations could depend on the behaviors of people toward social distancing, hygiene measures, temperature, relative humidity, and wind speed.

It is known that the human immune system suffers from depression during the winter season, as the cold, dry air works to dry out the mucus in our noses, which acts as a first line of defense, to prevent viruses. Studies have proven that if the temperature increases, the infected person will spread the COVID-19 to fewer people. Some studies also indicate that the virus cannot survive at 86° Fahrenheit (F) so in this paper, the countries are categorized into three main categories and these will have a reflection on human cells infections. The first category of countries, with low transmission rates, have an average annual temperature between (60 and 100.4) °F. The second category of countries, with medium transmission rates, have an average annual temperature between (37 and 89) °F and finally the third category of countries, with the highest transmission rates, have an average annual temperature between (36.14 and 66.3) °F. The higher the spread rates, the higher the rates of infection of human cells and also their exposure to the virus will be higher, and thus the possibility of human cells to sustain the virus will depend on the location of the country and also in which category it belongs.

There are a few published papers, studying the analysis of COVID-19. In [5], authors used numerical approaches and logistic modelling technique to make a complete analysis for COVID-19. In [6], authors made a fractional-order-modified SEIR model of pandemic diseases and applied it on COVID-19 to predict the virus spreading behavior in Pakistan and Malaysia. In [7], authors made predictions about coronavirus transmission dynamics in the African countries. The model parameters (protection rate, infection rate, average incubation time, average quarantined time, cure rate, and mortality rate) are selected using Metropolis-Hasting (MH) parameter optimization method [8]. Authors modelled COVID-19 daily confirmed cases in Egypt and Iraq by using Gaussian fitting model and logistic model [9–11]. The virus dynamic behavior is modelled with a new SEIR model. One of the important quantities should be calculated during modelling virus dynamics is the basic reproduction number. This value helps eliminate the disease and expected number of secondary infections produced by an infected individual in a population when all individuals are susceptible to infection. In [12], authors made a multi-strain-modified SEIR epidemic model for COVID-19 and wrote a complete analysis on how to control the value of the reproduction number to control the next upcoming virus peaks.

In this paper as a way to predict the upcoming coronavirus waves, a new state-of-art of regression models are used to model daily confirmed cases in different countries. In this study, countries are classified based on the time of the full viral wave and the average annual temperatures, where if time of the viral wave is lower, the virus transmission rate is higher. Fourier model and sum of sine-waves model are used to fit the daily confirmed data and predict the upcoming wave peaks for these countries. In this paper, the actual data from [3,13] are used for generating the predictive regression models without making any statistical modifications on them. The mathematical regression models fit the data from 1 March 2020 (day 0) to 15 November 2020 (day 260) and hence, the models can predict different scenarios for each country in the period from 16 November 2020 (day 261) to 10 April 2021 (day 400). The real smoothed data from [14] was used to verify the accuracy of the prediction models. All used datasets used in this paper have been attached with a separate supplementary material file. The countries under study are categorized into three main categories:

- First category: countries in which coronavirus wave takes low transmission about two-year seasons (about 180 days) to make a complete virus cycle: (case study of Saudi Arabia and Egypt).
- Second category: countries with higher transmission rates with one-year season (about 90 days) to make complete virus cycle. These countries take offline periods with low spreading rate before entering the next wave cycle: (case study of United Kingdom, Germany and Italy).
- The third category: countries with the highest transmission rates with one-year season (about 90 days) to make complete cycle and without offline periods (consecutive waves): (case study of United States of America and Russia).

The remainder of this paper is organized as follows: Section 2 describes the mathematical background of the proposed predictive modelling. Section 3 elaborates the used optimization algorithm to fit the data. Section 4 includes the results of applying the predictive models on the available data of the three category countries. In the end the conclusions are given.

## 2. Predictive Mathematical Modelling

Mathematical curve fitting is a key for getting a mathematical relation between measured values and their dependent input parameters named as regression models [15] like Polynomial Model, Exponential Model, Power Model, and Fourier Model [16–18]. Choosing the suitable model to represent coronavirus data can be measured by the mean of determination coefficient ( $R^2$ ) which takes values from 0 to 1. Higher value of  $R^2$  means higher model accuracy [18].  $R^2$  can be calculated from Equation (1) with N data points.

$$R^2 = \frac{\sum_{i=1}^N (y_i - y_{avg})^2 - \sum_{i=1}^N (y_i - y_{model})^2}{\sum_{i=1}^N (y_i - y_{avg})^2} \quad (1)$$

where  $y_{avg}$  is the average value of measured data,  $y_i$  is the measured data value at time  $t_i$  and  $y_{model}$  is the corresponding value using the fitting model equation.

The used mathematical curve fitting techniques here for modeling daily confirmed cases of COVID-19 and predict upcoming scenarios are Fourier fitting models and sum of sine-waves fitting models.

### 2.1. Fourier Fitting Models

Fourier models are used to model periodic functions with three main parts which are constant term, cosine-wave terms and sine-wave terms. As coronavirus is spreading like a wave and has a possibility of repetitions so this leads to near periodic wave assumptions. So, modelling using Fourier fitting model here is a good choice of fitting. Fourier fitting model equation with n-terms is presented in Equation (2) [19] and  $w$  is the fundamental frequency of sine and cosine terms.

$$y_{model} = a_0 + \sum_{i=1}^n a_n \cos(nwx) + b_n \sin(nwx) \quad (2)$$

$(a_0, a_1, a_2, b_1, b_2, \dots, a_n, b_n, w)$  are unknown coefficients and selected using numerical optimization techniques to satisfy the least square regression criteria between data and fitting model.

### 2.2. Sum of Sine-Waves Fitting Models

Sine-waves are always used to model periodic functions. Using the sum of sine waves with different frequencies can be used to model near periodic functions. These models are used here to fit data of coronavirus daily confirmed cases in the countries under study. Also modelling by this kind of mathematical equations can result in predicting next upcoming coronavirus waves in the countries under study. Equation (3) represents the sum of sine-waves equation with n-terms.

$$y_{model} = \sum_{i=1}^n a_n \sin(b_n x + c_n) \quad (3)$$

$(a_1, b_1, c_1, a_2, b_2, c_2, \dots, a_n, b_n, c_n)$  are unknown coefficients and optimized using numerical optimization techniques to minimize the error between model and measured daily confirmed cases.

The used optimization algorithm for optimal selections of each model coefficients is invasive weed optimization algorithm (IWO) [20]. The optimization process leads to increasing model accuracy and decreasing root mean square of errors (RMSE) between actual number (N) of data points and the fitting model. The objective function will be minimized using the optimization algorithm when modelling data by Fourier model with n-terms is as indicated in Equation (4). Similarly, the objective function used when modelling with the sum of sine-waves is as indicated in Equation (5).

Minimize:

$$RMSE = \sqrt{\frac{1}{N} (y_i - (a_0 + \sum_{i=1}^n a_n \cos(nwx) + b_n \sin(nwx)))^2} \tag{4}$$

Subject to:

$$\begin{aligned} a_{0_{\min}} &\leq a_0 \leq a_{0_{\max}} \\ a_{i_{\min}} &\leq a_1, a_2, \dots, a_n \leq a_{i_{\max}}, i = 1, 2, 3, \dots, n \\ b_{i_{\min}} &\leq b_1, b_2, \dots, b_n \leq b_{i_{\max}}, i = 1, 2, 3, \dots, n \end{aligned}$$

Minimize:

$$RMSE = \sqrt{\frac{1}{N} (y_i - \sum_{i=1}^n a_n \sin(b_n x + c_n))^2} \tag{5}$$

Subject to:

$$\begin{aligned} a_{i_{\min}} &\leq a_1, a_2, \dots, a_n \leq a_{i_{\max}}, i = 1, 2, 3, \dots, n \\ b_{i_{\min}} &\leq b_1, b_2, \dots, b_n \leq b_{i_{\max}}, i = 1, 2, 3, \dots, n \\ c_{i_{\min}} &\leq c_1, c_2, \dots, c_n \leq c_{i_{\max}}, i = 1, 2, 3, \dots, n \end{aligned}$$

### 3. Invasive Weed Optimization Algorithm

Invasive weed optimization algorithm will be used to select the optimal values of each fitting model coefficients. Steps of invasive weed optimization technique [20]:

- Populations Initialization: populations (seeds) with finite numbers are being scattered over d-dimensional searching space at random positions. The variable d is the number of the optimization process variables.
- Seeds Reproduction: seeds grow to form plants and produce newer number of seeds.
- Spatial Dispersal: The new seeds are being randomly scattered over the search space using standard normal distribution functions with variable variance. The standard deviation ( $\sigma$ ) of the normal random functions will be produced with initial value ( $\sigma_{initial}$ ) and final value ( $\sigma_{final}$ ) in all steps. During simulation, a nonlinear alteration-modulation index ( $s$ ) is selected to reach certain satisfactory performance. Standard deviation ( $\sigma_{iter}$ ) at each iteration ( $iter$ ) is calculated, where  $iter_{max}$  is the maximum number of iterations.

$$\sigma_{iter} = \frac{(iter_{max} - iter)^s}{(iter_{max})^s} (\sigma_{initial} - \sigma_{final}) + \sigma_{final} \tag{6}$$

- Competitive Exclusion: In this process and after maximum number of plants is reached, only the plants with lower fitness can pull out and produce seeds, others are being tossed out. The process continues in each iteration till maximum iterations is reached. Flow chart of the algorithm is as indicated in Figure 1.

Before starting the optimization process to get each fitting model coefficients and reach the different scenarios for each country, the following mathematical assumptions are made and taken into consideration during the optimization process:

- First category countries: these countries have low virus transmission rates and large wave time period, so the next upcoming wave peak will not exceed the first wave peak.

- Second category countries: these countries have medium transmission rates and will have two consecutive waves and with peaks multiples of the first wave peak but not out of control.
- Third category: these countries have the high transmission rates, so the next wave will have increased consecutive peaks and virus spread will be out of control without any limitations.

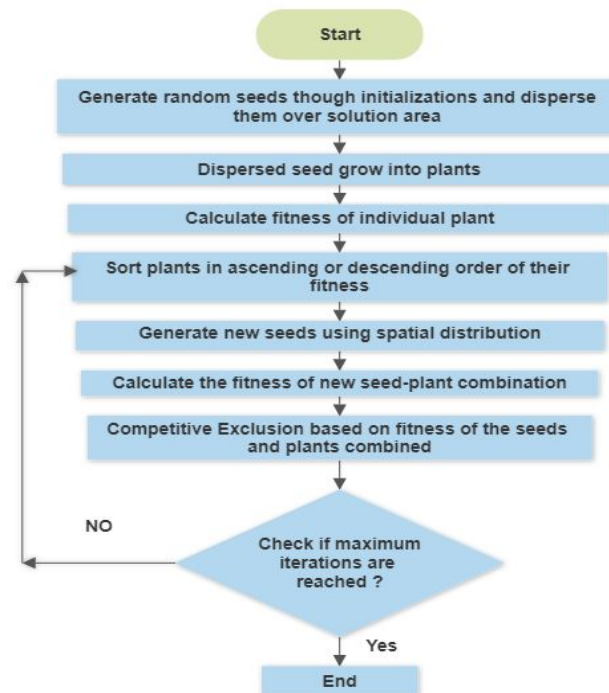


Figure 1. Invasive weed optimization (IWO) algorithm flow chart.

#### 4. Results and Discussion

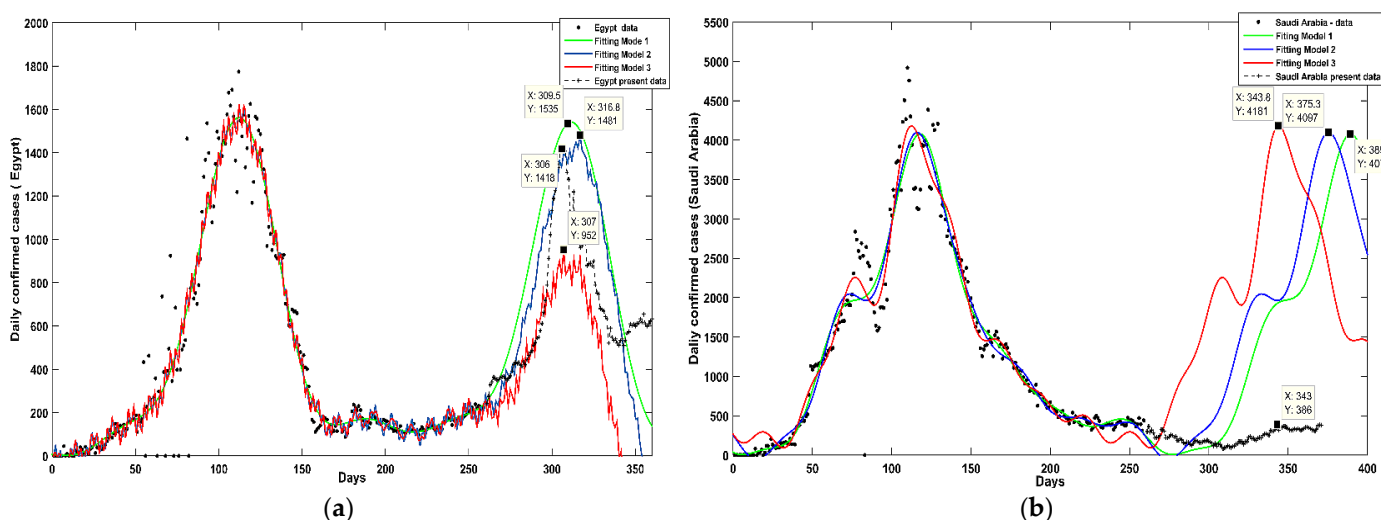
Results in this part are segmented according to the three main categories (first category countries, second category countries, and third category countries). Each country modelled with three fitting models to fit coronavirus data from 1 March 2020 (day 0) to 15 November 2020 (day 260) with least square error regression and predict different next upcoming wave scenarios. Optimization process is handled with 100 iterations for each model. Finally, all predictions of the models are compared with the real smoothed current data of daily confirmed cases in these countries [14].

##### 4.1. First Category Countries

The two countries under study here are Egypt and Saudi Arabia. Starting with Egypt, coronavirus daily confirmed cases are modelled using sum of sine-waves model with number of terms equals four, seven, and eight. The accuracy of modelling with these models is described with model determination coefficient  $R^2$ , sum square of error (SSE), and root mean square of errors (RMSE). Table 1 shows the accuracy of each model. Optimized coefficients of each model reached by the optimization technique is also presented. According accuracy of fitting the three models have very near accuracy of fitting ( $R_1^2 = 0.9538$ ,  $R_2^2 = 0.94$ ,  $R_3^2 = 0.9418$ ). In case of next upcoming wave peak in Egypt, the three models give different scenarios. As shown in Figure 2a, Model one scenario for Egypt has a peak of daily confirmed cases of 1535 person in day 310 (4 January 2021). Model two predicts the wave peak in day 317 (11 January 2021) with daily confirmed cases 1481. Model three predicts the best scenario of Egypt with the peak of daily confirmed cases of 952 person in day 307 (1 January 2021). Also, these three models predict how will be the daily confirmed cases numbers after the wave peak till day 360 (23 February 2021).

**Table 1.** Used mathematical fitting models for Egypt data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$\sum_{i=1}^{i=4} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=7} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=8} a_i \sin(b_i x + c_i)$		
$a_1 = 543.8$	$b_1 = 0.00356$	$c_1 = 0.8723$	$a_1 = 663.1$	$b_1 = 0.00854$	$c_1 = 0.0381$	$a_1 = 673.9$	$b_1 = 0.0101$	$c_1 = -0.1441$
$a_2 = 306.8$	$b_2 = 0.06305$	$c_2 = 0.8165$	$a_2 = 369.2$	$b_2 = 0.03404$	$c_2 = 3.668$	$a_2 = 288.3$	$b_2 = 0.03923$	$c_2 = 3.149$
$a_3 = 631.8$	$b_3 = 0.03121$	$c_3 = -1.853$	$a_3 = 385.4$	$b_3 = 0.02492$	$c_3 = -0.2774$	$a_3 = 431.8$	$b_3 = 0.02371$	$c_3 = -0.3358$
$a_4 = 106.8$	$b_4 = 0.09438$	$c_4 = -2.892$	$a_4 = 321.5$	$b_4 = 0.06116$	$c_4 = 1.082$	$a_4 = 293.1$	$b_4 = 0.06203$	$c_4 = 1.015$
			$a_5 = 114.4$	$b_5 = 0.09294$	$c_5 = 3.516$	$a_5 = 113.4$	$b_5 = 0.09291$	$c_5 = 3.497$
			$a_6 = 33.39$	$b_6 = 0.5596$	$c_6 = 0.2357$	$a_6 = 33.35$	$b_6 = 0.5598$	$c_6 = 0.2219$
			$a_7 = 31.15$	$b_7 = 1.994$	$c_7 = -2.128$	$a_7 = 32.54$	$b_7 = 1.998$	$c_7 = -2.509$
						$a_8 = 30.74$	$b_8 = 1.969$	$c_8 = 0.1529$
SSE = $4.003 \times 10^6$ , RMSE = 127.3, $R^2 = 0.9358$			SSE = $3.742 \times 10^6$ , RMSE = 125.4, $R^2 = 0.94$			SSE = $3.629 \times 10^6$ , RMSE = 124.3, $R^2 = 0.9418$		



**Figure 2.** Curve fitting models for daily confirmed cases for first category countries: (a) curve fitting models for Egypt data; (b) curve fitting models for Saudi Arabia data.

Saudi Arabia daily confirmed cases data are modelled using Fourier model with number of terms equals six, seven, and eight. Each model optimized coefficients are as indicated in Table 2. The accuracy of each model is described through  $R^2$ , SSE and RMSE. From Figure 2b, Saudi Arabia takes higher time about 200 days to complete the first virus wave cycle while Egypt takes only 150 days. This means that the next wave peak of Saudi Arabia will be delayed than Egypt. Model one predicts that next peak of Saudi Arabia will be in day 389 (29 March 2021) with daily confirmed cases 4078 person. In case of the model number two, the wave peak will be in day 376 (16 March 2021) with daily confirmed cases of 4097 person. Model three predicts that next peak will be in day 344 (16 February 2021) with daily confirmed cases 4181 person. The different scenarios of Saudi Arabia have near peak values but with a shift in time of action. For all scenarios the second wave cycle will not end before day 400 (10 April 2021).

The accuracy of the proposed models of each country is measured through sketching current real smoothed data for Egypt and Saudi Arabia [14] on same sketch with their predictive models as indicated in Figure 2. In case of Egypt, the actual current smoothed data are very near in shape with the forecasting models. The most accurate model of predictions is model three with peak value equals 1481 people in day 317 (11 January 2021) in comparison with the actual second virus peak which equals 1418 people in day 306 (31 December 2021). In case of Saudi Arabia, model one is the most accurate model to make

predictions till day 307 (1 January 2021). After that date, Saudi Arabia starts to vaccinate its population which results in over damping the second virus wave with peak equals 386 people in day 343 (15 February 2021). The actual vaccinated people in both Egypt and Saudi Arabia are as indicated in Table 3. Egypt starts to vaccinate thousands of people from its population after 24 February 2021 (day 352). This vaccination could damp the next wave of Egypt based on the number of actual vaccinated people per cumulative cases.

**Table 2.** Used mathematical fitting models for Saudi Arabia data with optimized coefficients.

Fitting Model 1		Fitting Model 2		Fitting Model 3	
$a_0 + \sum_{i=1}^{i=6} a_i \cos(iwx) + b_i \sin(iwx)$		$a_0 + \sum_{i=1}^{i=7} a_i \cos(iwx) + b_i \sin(iwx)$		$a_0 + \sum_{i=1}^{i=8} a_i \cos(iwx) + b_i \sin(iwx)$	
$a_0 = 1305$	$w = 0.02314$	$a_0 = 1395$	$w = 0.02422$	$a_0 = 1497$	$w = 0.02714$
$a_1 = -1405$	$b_1 = 578.3$	$a_1 = -1473$	$b_1 = 415$	$a_1 = -1558$	$b_1 = -61.53$
$a_2 = 247.7$	$b_2 = -545.8$	$a_2 = 350.8$	$b_2 = -440.2$	$a_2 = 430.1$	$b_2 = -93.8$
$a_3 = -90.21$	$b_3 = 128.1$	$a_3 = -140.8$	$b_3 = 72.92$	$a_3 = -183.3$	$b_3 = -129$
$a_4 = 46.91$	$b_4 = -243.9$	$a_4 = 175.7$	$b_4 = -208.7$	$a_4 = 239.6$	$b_4 = 117.2$
$a_5 = 50.73$	$b_5 = 176.5$	$a_5 = -35.92$	$b_5 = 159.7$	$a_5 = -140.5$	$b_5 = -6.673$
$a_6 = -118.4$	$b_6 = -37.12$	$a_6 = -62.3$	$b_6 = -48.45$	$a_6 = 42.76$	$b_6 = -82.72$
		$a_7 = 80.02$	$b_7 = 578.3$	$a_7 = -34.32$	$b_7 = 108.2$
				$a_8 = -25.43$	$b_8 = -146.6$
SSE = $2.213 \times 10^7$ , RMSE = 300.5, $R^2 = 0.9425$		SSE = $2.115 \times 10^7$ , RMSE = 295, $R^2 = 0.945$		SSE = $1.903 \times 10^7$ , RMSE = 281, $R^2 = 0.9505$	

**Table 3.** Actual vaccinated people in first category countries [14].

Country	Date	Cumulative Cases	Actual Vaccinated	Population	Vaccinated per Cumulative Cases
Egypt	24 February 2021	185,334	0	102,334,403	0
Saudi Arabia	5 March 2021	379,092	1,182,943	34,813,867	3.120464162

4.2. Second Category Countries

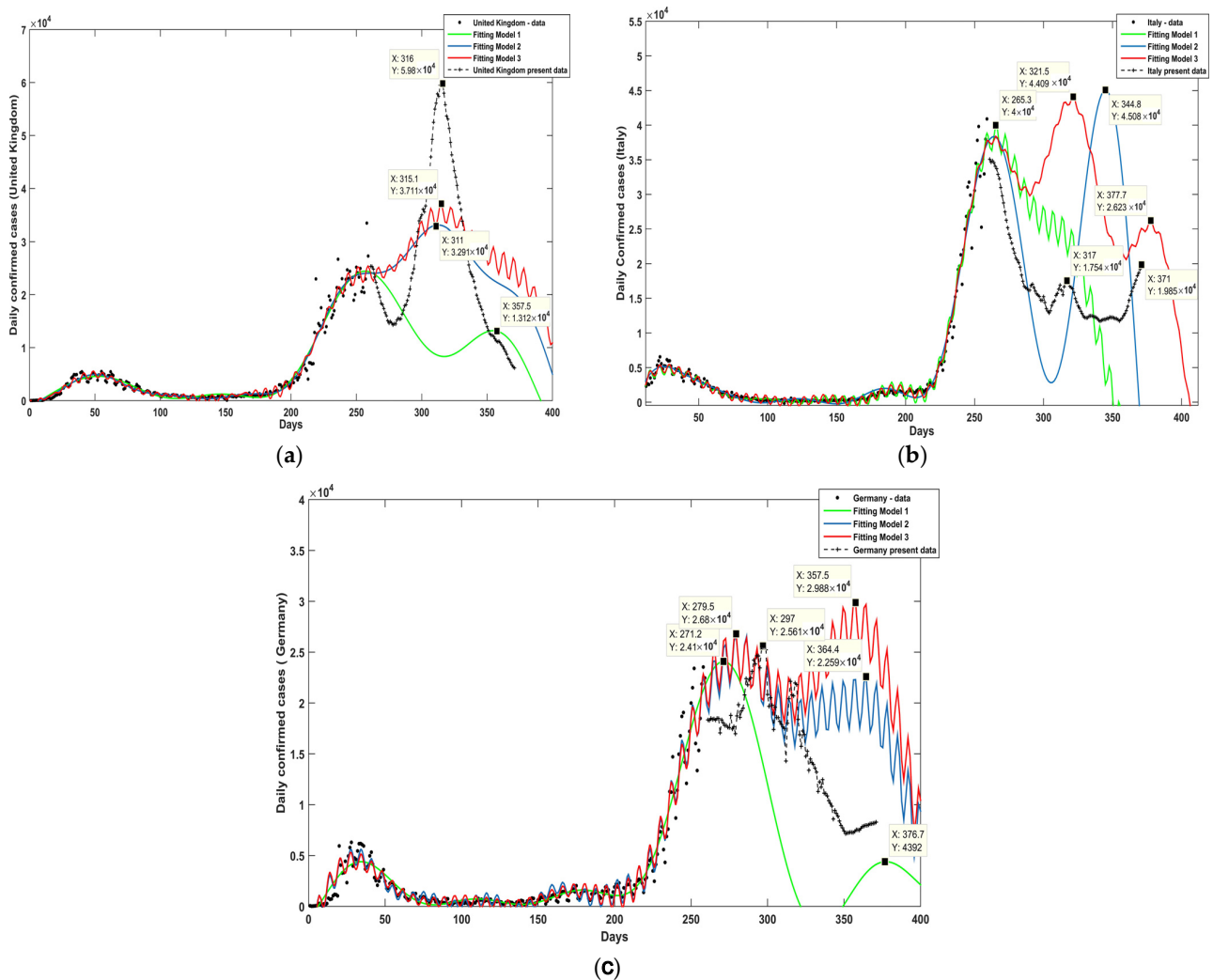
Countries under study in this category are United Kingdom, Italy, and Germany. These three countries first virus wave cycle takes only one-year season (within 90 days) which means that second wave started in these countries before the first category countries. Starting with United Kingdom, daily confirmed cases are modelled using the sum of sine-waves model with number of terms four, five, and six. The optimized coefficients of each model using IWO algorithm are as indicated in Table 4. Accuracy of each model can also be shown through  $R^2$ , SSE, and RMSE. All models predict that countries in this category will be exposed to consecutive waves. As shown Figure 3a, Model one predicts the best-case scenario and indicating that United Kingdom reached the peak of second wave and will have next peak in day 357 (26 February 2021) with peak value equals 13,180 persons which is lower than the current peak. Model two predicts that next wave peak will be with daily confirmed cases equals 32,910 persons in day 311 (11 January 2021). Model three predicts that next peak will reach 37,110 daily confirmed cases in day 316 (17 January 2021).

Italy daily confirmed cases are modelled using the sum of sine-waves model with five, six, and seven terms. Each model accuracy of fitting and optimized coefficients is as described in Table 5. Model one makes the best scenario as it predicts that Italy has reached the peak of the second wave in day 265 (20 November 2020) with daily confirmed cases equals 40,000 and it will not have a third wave before day 400 (10 April 2021). Model two same as model one shows that Italy reached second wave peak and will have the next wave peak on day 345 (14 February 2021) with number of daily confirmed cases equals 45,080. Model three predicts that Italy will have two next upcoming peaks (third and fourth peak).

The third peak will be in day 322(22 January 2021) with daily confirmed cases equals 40,900 and fourth peak will be in day 378 (18 March 2021). The reached results of the three models are as sketched in Figure 3b.

**Table 4.** Used mathematical fitting models for United Kingdom data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$\sum_{i=1}^{i=4} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=5} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=8} a_i \sin(b_i x + c_i)$		
$a_1 = 12290$	$b_1 = 0.00747$	$c_1 = -0.542$	$a_1 = 21270$	$b_1 = 0.00953$	$c_1 = -0.8266$	$a_1 = 22040$	$b_1 = 0.00874$	$c_1 = -0.7489$
$a_2 = 7628$	$b_2 = 0.02684$	$c_2 = 0.613$	$a_2 = 14400$	$b_2 = 0.02033$	$c_2 = 1.183$	$a_2 = 13930$	$b_2 = 0.01997$	$c_2 = 1.223$
$a_3 = 33830$	$b_3 = 0.0563$	$c_3 = 1.454$	$a_3 = 1385$	$b_3 = 0.05471$	$c_3 = 0.6467$	$a_3 = 1303$	$b_3 = 0.05599$	$c_3 = 0.5582$
$a_4 = 3337$	$b_4 = 0.05711$	$c_4 = 4.554$	$a_4 = 1689$	$b_4 = 0.07899$	$c_4 = 2.187$	$a_4 = 1700$	$b_4 = 0.07788$	$c_4 = 2.315$
			$a_5 = 1309$	$b_5 = 0.09304$	$c_5 = -2.453$	$a_5 = 1261$	$b_5 = 0.09313$	$c_5 = -2.439$
						$a_6 = 436.1$	$b_6 = 0.8894$	$c_6 = 1.947$
						$a_7 = 14580$	$b_7 = 0.7912$	$c_7 = -1.023$
						$a_8 = 14560$	$b_8 = 0.7917$	$c_8 = 2.076$
SSE = $5.684 \times 10^8$ , RMSE = 1517, $R^2 = 0.9545$			SSE = $5.194 \times 10^8$ , RMSE = 1459, $R^2 = 0.9602$			SSE = $4.368 \times 10^8$ , RMSE = 1363, $R^2 = 0.9666$		



**Figure 3.** Curve fitting models for daily confirmed cases for second category countries: (a) curve fitting models for United Kingdom data; (b) curve fitting models for Italy data; (c) curve fitting models for Germany data.



**Table 5.** Used mathematical fitting models for Italy data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$\sum_{i=1}^{i=5} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=6} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=7} a_i \sin(b_i x + c_i)$		
$a_1 = 20000$	$b_1 = 0.01263$	$c_1 = -1.266$	$a_1 = 24710$	$b_1 = 0.00923$	$c_1 = -0.9455$	$a_1 = 27960$	$b_1 = 0.011$	$c_1 = -1.127$
$a_2 = 16420$	$b_2 = 0.02194$	$c_2 = 0.851$	$a_2 = 16920$	$b_2 = 0.02204$	$c_2 = 0.7784$	$a_2 = 21440$	$b_2 = 0.02024$	$c_2 = 0.9949$
$a_3 = 10400$	$b_3 = 0.0635$	$c_3 = -1.005$	$a_3 = 4569$	$b_3 = 0.04886$	$c_3 = 0.6835$	$a_3 = 4201$	$b_3 = 0.05069$	$c_3 = 0.4756$
$a_4 = 38770$	$b_4 = 0.07715$	$c_4 = 0.6183$	$a_4 = 2740$	$b_4 = 0.08085$	$c_4 = 0.1128$	$a_4 = 3070$	$b_4 = 0.08125$	$c_4 = 0.09188$
$a_5 = 30890$	$b_5 = 0.07975$	$c_5 = 3.489$	$a_5 = 3000$	$b_5 = 0.09997$	$c_5 = 1.318$	$a_5 = 4634$	$b_5 = 0.1009$	$c_5 = 1.257$
			$a_6 = 1660$	$b_6 = 0.1117$	$c_6 = 3.37$	$a_6 = 3061$	$b_6 = 0.108$	$c_6 = 3.757$
						$a_7 = 632.2$	$b_7 = 0.9004$	$c_7 = 0.8979$
SSE = $4.577 \times 10^8$ , RMSE = 1370, $R^2 = 0.9737$			SSE = $4.166 \times 10^8$ , RMSE = 1315, $R^2 = 0.976$			SSE = $3.617 \times 10^8$ , RMSE = 1233, $R^2 = 0.9792$		

Germany daily confirmed coronavirus data are modelled using Fourier model with four terms and sum of sine-waves model with number of terms equals seven and eight. Each model fitting accuracy and optimized results is as indicated in Table 6. The accuracy is also compared here through  $R^2$ , SSE, and RMSE indices. As indicated in Figure 3c, the first model describes the best scenario for Germany with reached second wave peak in day 272 (27 November 2020) and the value equals 24,100 persons then daily confirmed cases starts to decrease again. The third wave of the virus will have a peak in day 377 (17 March 2021) but with smaller value than second wave peak where it only equals 4392 cases. Second and third fitting models admitted the same scenario about the second wave peak with same daily confirmed cases and in same day. Both models predict that it will be in day 279 (3 December 2021) with daily confirmed cases equals 26,800 persons. The third wave predicted peak time of action is different when comparing model two and model three. Model two predicts that it will be in day 365 (5 March 2021) with daily confirmed cases equals 22,590 persons.

**Table 6.** Used mathematical fitting models for Germany data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$a_0 + \sum_{i=1}^{i=4} a_i \cos(iwx) + b_i \sin(iwx)$			$\sum_{i=1}^{i=7} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=8} a_i \sin(b_i x + c_i)$		
$a_0 = 4948$	$w = 0.01832$	$a_1 = 18920$	$b_1 = 0.01002$	$c_1 = -0.9919$	$a_1 = 14690$	$b_1 = 0.008242$	$c_1 = -0.7996$	
$a_1 = 1713$	$b_1 = -6890$	$a_2 = 13300$	$b_2 = 0.01918$	$c_2 = 1.135$	$a_2 = 8897$	$b_2 = 0.02222$	$c_2 = 0.7905$	
$a_2 = -5237$	$b_2 = -1578$	$a_3 = 3520$	$b_3 = 0.05628$	$c_3 = -0.3835$	$a_3 = 3278$	$b_3 = 0.05336$	$c_3 = 0.01089$	
$a_3 = -3071$	$b_3 = 3513$	$a_4 = 2297$	$b_4 = 0.07041$	$c_4 = 1.044$	$a_4 = 1750$	$b_4 = 0.07153$	$c_4 = 0.9566$	
$a_4 = 403.9$	$b_4 = 2153$	$a_5 = 21500$	$b_5 = 0.8897$	$c_5 = 3.178$	$a_5 = 36350$	$b_5 = 0.8895$	$c_5 = 3.185$	
			$a_6 = 368.3$	$b_6 = 0.9251$	$c_6 = 1.678$	$a_6 = 366.9$	$b_6 = 0.9252$	$c_6 = 1.678$
			$a_7 = 20950$	$b_7 = 0.8892$	$c_7 = 0.07616$	$a_7 = 36150$	$b_7 = 0.8892$	$c_7 = 0.06693$
						$a_8 = 483.2$	$b_8 = 0.1485$	$c_8 = 2.758$
SSE = $3.535 \times 10^8$ , RMSE = 1192, $R^2 = 0.9394$			SSE = $2.321 \times 10^8$ , RMSE = 987.4, $R^2 = 0.9602$			SSE = $2.013 \times 10^8$ , RMSE = 925.5, $R^2 = 0.9655$		

Forecasting accuracy of the predictive models of each country is made through sketching current real smoothed data for United Kingdom, Italy, and Germany [14] on same sketch with their proposed predictive models as indicated in Figure 3. The predictive models are helpful tools for forecasting virus behavior till day 307 (1 January 2021) as they give curves very near in shape with actual current smoothed data. All countries in this category started to vaccinate their population to limit the virus spreading and each country’s actual vaccinated people per actual confirmed cases is as indicated in Table 7. In

case of United Kingdom, government starts to vaccinate its population with high rate to control virus transmission and damp the next virus waves. However, United Kingdom has the highest virus wave peak among countries in this category with peak equals 59,810 in day 316 (10 January 2021) but using vaccination with high rate results in over damping the virus next wave and makes daily today confirmed cases lower than both Italy and Germany. The actual vaccinated people per actual confirmed cases for United Kingdom reached 5.06 as indicated in Table 7. In case of Italy, the actual vaccinated people per actual confirmed cases is still low with value equals 1.179, so virus daily confirmed cases started to increase again as indicated in Figure 3b. From indicated results, Italian government should increase the number of actual vaccinated people per cumulative cases to more control virus behavior. In case of Germany, the actual number of vaccinated people per cumulative cases has an intermediate value. As, indicated in Figure 3c, today daily confirmed cases of Germany start to increase again with a lower rate so, number of actual vaccinated people per cumulative cases should slightly increase.

Table 7. Actual vaccinated people in second category countries [14].

Country	Date	Cumulative Cases	Actual Vaccinated	Population	Vaccinated per Cumulative Cases
United Kingdom	4 March 2021	4,213,764	21,358,815	67,886,004	5.068820893
Italy	5 March 2021	3,023,129	3,564,698	60,461,828	1.179141876
Germany	4 March 2021	2,484,306	4,736,174	83,783,945	1.906437452

4.3. Third Category Countries

Case studies in this category are United States of America (U.S.A.) and Russia. These countries have the highest transmission rates of virus spread and suffer from consecutive virus waves without offline periods and increasing virus peak wave value. U.S.A. daily confirmed cases data are modelled using sum of sine-waves model with five, six, and seven terms. Accuracy of each model is as described in Table 8. As in Figure 4a, first model represents the best-case scenario where it expects that next upcoming wave peak will be in day 350 (19 February 2021) with daily confirmed cases equals 405,900 persons. Model two expects the intermediate scenario with peak daily confirmed cases equals 529,800 in day 365 (6 March 2021). Model three is the worst-case scenario with the highest peak value as it expects that it will be 830,400 cases in day 365 (6 March 2021).

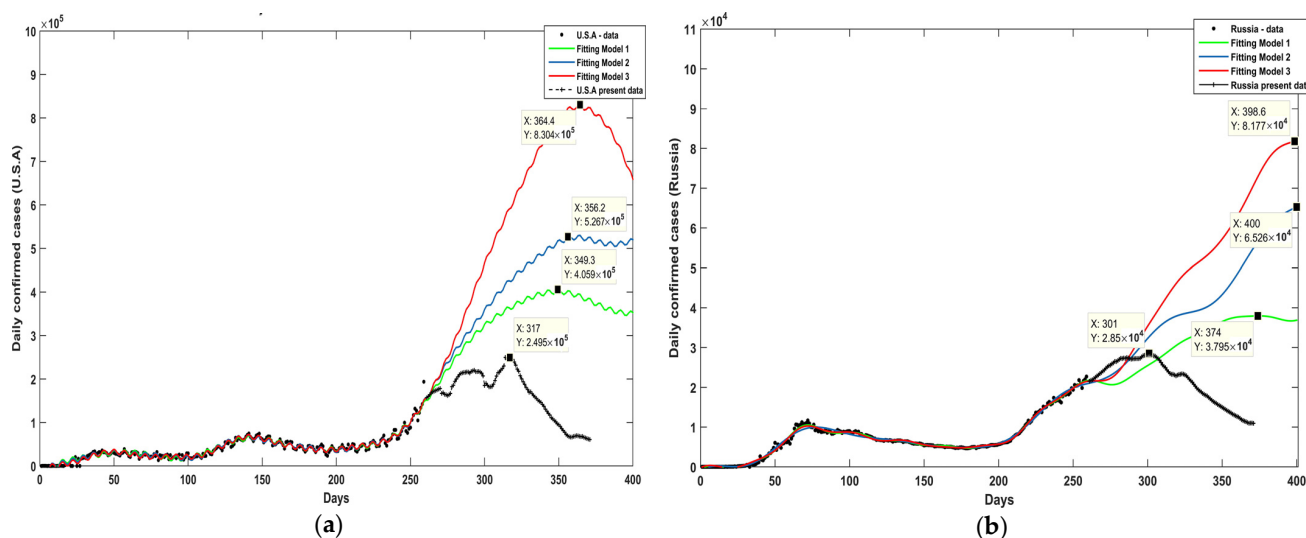


Figure 4. Curve fitting models for daily confirmed cases for third category countries: (a) Curve fitting models for U.S.A. data; (b) Curve fitting models for Russia data.

Daily confirmed cases data for Russia are modelled using five, seven, and eight terms of sum of sine-waves model. Each model optimized coefficients, satisfying least square regression criteria, is as indicated in Table 9. All models have accuracy of fitting and low SSE. In case of models' predictions, model one predicts the best-case scenario with the next upcoming wave peak which equals 37,950 in day 374 (15 March 2021). Model two predicts the intermediate scenario with next wave peak to be in day 400 (10 April 2021) with 65,260 daily confirmed cases. Model three expects the next peak in day 397 (7 March 2021) with daily confirmed cases equals 81,770 persons, so it represents the worst case with the highest value of peak daily confirmed cases.

**Table 8.** Used mathematical fitting models for U.S.A. data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$\sum_{i=1}^{i=5} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=6} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=7} a_i \sin(b_i x + c_i)$		
$a_1 = 400600$	$b_1 = 0.00321$	$c_1 = -0.2983$	$a_1 = 452100$	$b_1 = 0.00403$	$c_1 = -0.4072$	$a_1 = 498600$	$b_1 = 0.00941$	$c_1 = -1.017$
$a_2 = 96230$	$b_2 = 0.02034$	$c_2 = 0.6373$	$a_2 = 141900$	$b_2 = 0.0178$	$c_2 = 0.9664$	$a_2 = 40820$	$b_2 = 0.01767$	$c_2 = 1.073$
$a_3 = 37820$	$b_3 = 0.04033$	$c_3 = 1.48$	$a_3 = 92600$	$b_3 = 0.03939$	$c_3 = 1.604$	$a_3 = 107900$	$b_3 = 0.03055$	$c_3 = 2.696$
$a_4 = 7571$	$b_4 = 0.08621$	$c_4 = 2.28$	$a_4 = 8093$	$b_4 = 0.08655$	$c_4 = 2.23$	$a_4 = 8886$	$b_4 = 0.08118$	$c_4 = 2.858$
$a_5 = 5684$	$b_5 = 0.9012$	$c_5 = 0.4016$	$a_5 = 5648$	$b_5 = 0.9011$	$c_5 = 0.4179$	$a_5 = 5637$	$b_5 = 0.901$	$c_5 = 0.4269$
			$a_6 = 3327$	$b_6 = 0.1378$	$c_6 = 2.088$	$a_6 = 2297$	$b_6 = 0.1451$	$c_6 = 1.196$
						$a_7 = 2168$	$b_7 = 0.1231$	$c_7 = 2.613$
SSE = $1.022 \times 10^{10}$ , RMSE = 6473, $R^2 = 0.9457$			SSE = $8.741 \times 10^9$ , RMSE = 6023, $R^2 = 0.9536$			SSE = $8.248 \times 10^9$ , RMSE = 5887, $R^2 = 0.9562$		

**Table 9.** Used mathematical fitting models for Russia data with optimized coefficients.

Fitting Model 1			Fitting Model 2			Fitting Model 3		
$\sum_{i=1}^{i=5} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=7} a_i \sin(b_i x + c_i)$			$\sum_{i=1}^{i=8} a_i \sin(b_i x + c_i)$		
$a_1 = 83530$	$b_1 = 0.00223$	$c_1 = -0.1837$	$a_1 = 91990$	$b_1 = 0.00264$	$c_1 = 0.2512$	$a_1 = 35880$	$b_1 = 0.00313$	$c_1 = -0.1839$
$a_2 = 14410$	$b_2 = 0.01383$	$c_2 = 1.518$	$a_2 = 20040$	$b_2 = 0.01428$	$c_2 = 1.382$	$a_2 = 7460$	$b_2 = 0.01973$	$c_2 = 0.7906$
$a_3 = 2341$	$b_3 = 0.03704$	$c_3 = 4.395$	$a_3 = 1405$	$b_3 = 0.04701$	$c_3 = 3.023$	$a_3 = 2006$	$b_3 = 0.04475$	$c_3 = 3.259$
$a_4 = 1536$	$b_4 = 0.08227$	$c_4 = 1.251$	$a_4 = 1874$	$b_4 = 0.08093$	$c_4 = 1.425$	$a_4 = 1455$	$b_4 = 0.07198$	$c_4 = 2.66$
$a_5 = 939.5$	$b_5 = 0.09879$	$c_5 = 1.871$	$a_5 = 1085$	$b_5 = 0.09618$	$c_5 = 2.279$	$a_5 = 549.9$	$b_5 = 0.1076$	$c_5 = 0.3223$
			$a_6 = 362.7$	$b_6 = 0.1633$	$c_6 = 3.053$	$a_6 = 320.4$	$b_6 = 0.1634$	$c_6 = 2.919$
			$a_7 = 308.7$	$b_7 = 0.1987$	$c_7 = 0.3792$	$a_7 = 329.4$	$b_7 = 0.199$	$c_7 = 0.2491$
						$a_8 = 377.6$	$b_8 = 0.123$	$c_8 = 0.3394$
SSE = $8.804 \times 10^7$ , RMSE = 600.7, $R^2 = 0.9872$			SSE = $5.083 \times 10^7$ , RMSE = 462.2, $R^2 = 0.9926$			SSE = $4.492 \times 10^7$ , RMSE = 437.2, $R^2 = 0.9935$		

Current real smoothed data for U.S.A. and Russia [14] are sketched with the predictive modelling curves as in Figure 4. In case of U.S.A., forecasting models give good predictions when compared with current data till day 307 (1 January 2021). After that date, U.S.A. began to vaccinate its population so virus transmission rate starts to decrease and next virus wave is damped. The actual number of vaccinated people in U.S.A. till 5 March 2021 is as indicated in Table 10. In case of Russia, all predictive models give very accurate predictions till day 310 (4 January 2021). After that date, Russia began to vaccinate its population. Vaccination makes the next virus wave over damped. The actual vaccinated people per cumulative confirmed cases for both U.S.A. and Russia are as indicated in Table 10. As this value is higher, the vaccination effect will be stronger and virus transmission rate will be lower and virus break out is more controlled.

**Table 10.** Actual vaccinated people in third category countries [14].

Country	Date	Cumulative Cases	Actual Vaccinated	Population	Vaccinated per Cumulative Cases
U.S.A.	5 March 2021	28,894,541	55,547,697	331,002,647	1.922428773
Russia	5 March 2021	4,252,876	4,908,178	145,934,460	1.154084436

## 5. Conclusions

In this article, both Fourier models and sums of sine-waves models are used to predict the upcoming coronavirus peaks in the countries under study. In the first category countries (Egypt and Saudi Arabia), the models used gave different scenarios for each country in a form of three different scenarios with different wave peak and action time. From obtained results, Egypt and Saudi Arabia will be only exposed to a second wave till 10 April 2021. In case of second category countries (United Kingdom, Italy, and Germany), the models used gave different three scenarios of upcoming coronavirus wave peak. Most of the used models predict that these countries will suffer from two consecutive wave peaks and will suffer a third virus wave before 10 April 2021. In these countries, the spread of the virus will be controlled because the time for the second wave is limited and the daily confirmed cases decrease after the second wave reaches its peak. In the third category countries (The United States and Russia), the used models expect that these countries will reach the peak of the third wave of coronavirus before 10 April 2021, and these countries will suffer from consecutive increasing peaks. Finally, all predictive models for countries under study are compared to their current smoothed current data of daily confirmed cases to check their prediction accuracy. In the case of Egypt, the only country without vaccination effect, predictive models give very near curve shapes to the actual smoothed current data. For the remaining countries, with different values of actual vaccinated people per cumulative cases, the predictive models are helpful tools of forecasting virus behavior till day 307 (1 January 2021). After that date, the vaccination effect starts to limit the virus transmission rate and the next wave is damped in these countries. In our future work, we will make developments on the current predictive models considering how vaccination affects the virus spread rate.

**Author Contributions:** The authors declare that the study was realized in collaboration with the same responsibility. All authors have read and agreed to the published version of the manuscript.

**Funding:** RUDN University Strategic Academic Leadership Program; RUDN University Strategic Academic Leadership Program.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** All used data in this paper are valid online on WHO coronavirus (COVID-19) dashboard. DOI: <https://covid19.who.int/> (accessed on 15 November 2020) and Worldometers. Coronavirus Cases. 2020. DOI: <https://www.worldometers.info/Coronavirus/Coronavirus-cases> (accessed on 15 November 2020).

**Acknowledgments:** The research is partially supported by the Community University Research Engagement Centre (CURE) and the RUDN University Program 5-100.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

1. World Health Organization. Available online: <https://www.who.int/emergencies/diseases/novel-coronavirus2019> (accessed on 15 November 2020).
2. Lauer, S.; Grantz, K.; Bi, Q.; Jones, F.; Zheng, Q.; Meredith, H.; Azman, A.; Reich, N.; Lessler, J. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* **2020**, *172*, 577–582. [[CrossRef](#)]
3. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available online: <https://covid19.who.int/> (accessed on 15 November 2020).
4. Barua, S. Understanding Coronanomics: The economic implications of the Coronavirus (COVID-19) pandemic. *SSRN Electron. J.* **2020**. [[CrossRef](#)]
5. Ahmed, A.; Salam, B.; Mohammad, M.; Ali, A.; Sarbaz, A.; Khoshnaw, S. Analysis Coronavirus disease (COVID-19) model using numerical approaches & logistic model. *Int. Open Access J. Aims Bioeng.* **2020**, *7*, 130–146.
6. Alrabaiah, H.; Arfan, M.; Shah, K.; Mahariq, I.; Ullah, A. A comparative study of spreading of novel Coronavirus disease by using fractional order modified SEIR model. *Alex. Eng. J.* **2020**. [[CrossRef](#)]
7. Zhao, Z.; Li, X.; Liu, F.; Zhu, G.; Ma, C.; Wang, L. 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. *Sci. Total Environ.* **2020**, *729*, 138959. [[CrossRef](#)]
8. Zhu, G.; Li, X.; Su, Y.; Zhang, K.; Bai, Y.; Ma, J.; Li, C.; Hu, X.; He, J. Simultaneous parameterization of the two-source evapotranspiration model by Bayesian approach: Application to spring maize in an arid region of northwest China. *Geosci. Model Dev.* **2014**, *7*, 741–775. [[CrossRef](#)]
9. Amar, L.; Taha, A.; Mohamed, M. Prediction of the final size for COVID-19 epidemic using machine learning: A case study of Egypt. *Infect. Dis. Model.* **2020**, *5*, 622–634. [[CrossRef](#)] [[PubMed](#)]
10. Schüttler, J.; Schlickeiser, R.; Schlickeiser, F.; KrSchöttlerger, M. COVID-19 Predictions Using a Gauss Model, Based on Data from April 2. *Physics* **2020**, *2*, 197–212. [[CrossRef](#)]
11. Bacaër, N. Verhulst and the logistic equation (1838). In *A Short History of Mathematical Population Dynamics*; Springer Ltd.: London, UK, 2011; pp. 35–39.
12. Khyar, O.; Allali, K. Global dynamics of a multi-strain SEIR epidemic model with general incidence rates: Application to COVID-19 pandemic. *Nonlinear Dyn.* **2020**, *102*, 489–509. [[CrossRef](#)] [[PubMed](#)]
13. Worldometers. Coronavirus Cases. 2020. Available online: <https://www.worldometers.info/Coronavirus/Coronavirus-cases> (accessed on 15 November 2020).
14. Our World in Data. Available online: <https://ourworldindata.org/grapher/daily-cases-covid-region> (accessed on 6 March 2021).
15. Chapra, S.C.; Canale, R.P. *Numerical Methods for Engineers*; McGraw-Hill, Inc.: New York, NY, USA, 1988.
16. Süli, E.; Mayers, D.F. *An Introduction to Numerical Analysis*; Cambridge University Press: Cambridge, UK, 2003.
17. Guest, P.G.; Guest, P.G. *Numerical Methods of Curve Fitting*; Cambridge University Press: Cambridge, UK, 2012.
18. Abuelma'atti, M. A simple algorithm for fitting measured data to Fourier-series models. *Int. J. Math. Educ. Sci. Technol.* **2006**. [[CrossRef](#)]
19. Wang, F.; Zhang, K. The theoretical analysis and algorithm on a class of optimal curve fitting problems. *Appl. Math. Comput.* **2006**, *183*, 622–633. [[CrossRef](#)]
20. Mehrabian, A.; Lucas, C. A novel numerical optimization algorithm inspired from weed colonization. *Ecol. Inform.* **2006**, *1*, 355–366. [[CrossRef](#)]