Applied Soft Computing 111 (2021) 107708

Contents lists available at ScienceDirect

Applied Soft Computing

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

Forecasting fully vaccinated people against COVID-19 and examining future vaccination rate for herd immunity in the US, Asia, Europe, Africa, South America, and the World

Pınar Cihan

Department of Computer Engineering, Çorlu Faculty of Engineering, Tekirdağ, Namık Kemal University, Tekirdağ, Turkey

ARTICLE INFO

Article history:
Received 14 April 2021
Received in revised form 5 June 2021
Accepted 7 July 2021
Available online 14 July 2021

Keywords:
COVID-19
SARS-CoV-2
Vaccine
Herd immunity
ARIMA
Forecasting

ABSTRACT

Coronavirus disease 2019 (COVID-2019) has spread rapidly all over the world and it is known that the most effective way to eliminate the disease is vaccination. Although the traditional vaccine development process is quite long, more than ten COVID-19 vaccines have been approved for use in about a year. The COVID-19 vaccines that have been administered are highly effective enough, but achieving herd immunity is required to end the pandemic. The motivation of this study is to contribute to review the countries' vaccine policies and adjusting the manufacturing plans of the vaccine companies. In this study, the total number of people fully vaccinated against COVID-19 was forecasted in the US, Asia, Europe, Africa, South America, and the World with the Autoregressive Integrated Moving Average (ARIMA) model, which is a new approach in vaccination studies. Additionally, for herd immunity, the percentage of fully vaccinated people in these regions at the beginning of 2021 summer was determined. ARIMA results show that in the US, Asia, Europe, Africa, South America, and the World will reach 139 million, 109 million, 127 million, 8 million, 38 million, and 441 million people will be fully vaccinated on 1 June 2021, respectively. According to these results, 41.8% of the US, 2.3% of Asia, 17% of Europe, 0.6% of Africa, 8.8% of South America, and 5.6% of the World population will be fully vaccinated people against the COVID-19. Results show that countries are far from the herd immunity threshold level desired to reach for safely slow or stop the COVID-19 epidemic.

1. Introduction

The coronavirus was reported as pneumonia of unknown etiology in Wuhan, China in December 2019. In January 2020, it was determined that the disease agent was a new coronavirus (2019-nCoV) that was not previously detected in humans. Because of its high similarity to SARS coronavirus (SARS-CoV), it was named SARS-CoV-2, and the disease that occurred was later named coronavirus disease 19 (COVID-19) [1].

The virus spread rapidly around the world, as there is no treatment and vaccine for COVID-19 disease. In the first three months, 114 countries were affected by the virus, and 4291 deaths occurred. As a result of the rapid spread of the disease, it was declared as a global pandemic by the World Health Organization (WHO) on March 11, 2020 [2].

SARS-CoV-2 can be transmitted from an infected person without symptoms and can lead to a pandemic disease within a week. This situation shows how important and mandatory vaccination is in controlling SARS-CoV-2 [3]. Therefore, a lot of effort has been made recently to develop vaccines against human coronavirus infections such as MERS and SARS. However, the fact that antiviral agents or vaccines have not been developed against MERS and SARS viruses until today has made COVID-19 a global threat [4–6].

Developed SARS-CoV-2 vaccines are based on mRNA, viral vector, inactive and protein. mRNA type vaccines provide the genetic code (DNA or RNA) for our cells to produce viral proteins. Once the proteins (non-disease causing) are produced, the body develops immunity against the virüs [7]. The first COVID-19 vaccine is Pfizer–BioNTech and is of the mRNA type. In viral vector-based vaccines, the genetic material of the virus is inserted into other viruses that do not cause genetic disease and applied to humans. Inactivated or weakened virus vaccine technique is the classical method. Here, the weakened or dead virus is used. In protein-based vaccine, virus proteins are used either directly isolated from the virus or artificially produced. Developed SARS-CoV-2 vaccines and properties of these vaccines are given in Table 1.

Vaccines have proven to be the most effective and economical way to prevent and control infectious diseases [32]. Although COVID-19 vaccines are highly efficacy enough, herd immunity is required to end the pandemic. Herd immunity can be achieved by vaccination or natural immunization. The fastest way to achieve

E-mail address: pkaya@nku.edu.tr.

https://doi.org/10.1016/j.asoc.2021.107708

Table 1
SARS-CoV-2 vaccines and properties of vaccines.

Vaccine	Efficacy	Type	Storage	Doses
Pfizer–BioNTech [8]	95%	mRNA	−70 °C	2
Moderna [9]	94%	mRNA	−20 °C	2
SputnikV [10]	91%	Viral vector	<=−18 °C	2
AstraZeneca [9]	70%	Viral vector	2–8 °C	2
Convidecia [11]	66%	Viral vector	2–8 °C	1
Janssen [12]	76.7–85.4%	Viral vector	2–8 °C	1
CoviVac [12]	–	Inactivate	2–8 °C	2
Sinopharm [13]	79%	Inactivate	2–8 °C	2
CoronaVac/Sinovac [12]	50%–91%	Inactivate	2–8 °C	2
Covaxin [12]	81%	Inactivate	2–8 °C	2
EpiVacCorona [12]	100%	Protein	2–8 °C	2
ZF2001/RBD-Dimer [14]	70%–95%	Protein	2–8 °C	3
Novavax [15]	89%	Protein	2–8 °C	2

herd immunity is vaccination. Herd immunity for the COVID-19 disease may vary from country to country because not every individual in the community can be vaccinated (for example, babies, pregnant women, those with medical problems, or those who do not want to be vaccinated). For this reason, reaching the minimum vaccination rate determined in order to achieve herd immunity is important in terms of eliminating the COVID-19 pandemic. The percentage of people who need to be immune to ensure herd immunity varies with each disease. Because this may vary depending on the vaccine, the population, the populations prioritized for vaccination, and other factors [33]. By forecasting the total number of fully vaccinated people by country, it can be observed that approximately how much of the population will be vaccinated and how close the threshold level for herd immunity has been reached.

Estimating the future values of an observed time series plays an important role in many sciences and engineering fields. The main idea of the ARIMA model is to identify mathematical and statistical theory and then used it to forecast time-series data. Integrated Moving Average (ARIMA) model is frequently used to predict future values due to its simple structure, flexibility, fast applicability and captures many different patterns [18]. In the past epidemics and health studies, the ARIMA time series model was successfully used [34–38]. ARIMA models are also widely and successfully used for current COVID-19 epidemic research. Some studies using the ARIMA model for the COVID-19 epidemic are given in Table 2.

In the literature, there are many studies on estimating the number of COVID-19 cases and/or deaths using the ARIMA models. Nevertheless, there is a gap in the literature in estimating the prevalence trend of the COVID-19 vaccinations by using the ARIMA time series model.

Table 3
Date of first fully vaccinated people of the regions and the size of the dataset.

Region	Date	Data size
US	January 14, 2021	129
Asia	January 04, 2021	139
Europe	December 27, 2020	147
Africa	February 03, 2021	109
South America	January 13, 2021	130
World	December 27, 2020	147

Table 4
Results of the augmented Dickey–Fuller test of the time series data.

Region	Data	Dickey–Fuller	p-value
Us	Original	−1.048	0.735
	First-ordered	−1.348	0.607
	Second-ordered	−6.249	0.000*
Asia	Original	0.132	0.968
	First-ordered	−1.194	0.676
	Second-ordered	−10.130	0.000*
Europe	Original	3.790	1.000
	First-ordered	2.169	0.998
	Second-ordered	−5.538	0.000*
Africa	Original	0.199	0.972
	First-ordered	−2.786	0.060
	Second-ordered	−9.139	0.000*
South America	Original	3.790	1.000
	First-ordered	2.169	0.998
	Second-ordered	−5.537	0.000*
World	Original	−1.182	0.681
	First-ordered	−0.655	0.858
	Second-ordered	−7.234	0.000*

*P<0.05 and thus stationary at that level.

The main purpose of this study is to forecast the number of people fully vaccinated against COVID-19 in the US, Asia, Europe, Africa, South America, and the World and to analyze whether it has reached a sufficient rate for herd immunity. ARIMA model was used to forecast the future number of fully vaccinated people for selected regions. Thus, the number of vaccinations and vaccine supply plans needed in these regions in the future for herd immunization can be helped. The contributions of this paper are as follows;

- Forecasting the number of people fully vaccinated against COVID-19 in the near future with ARIMA, which is the new approach in the COVID-19 vaccination studies,
- To identify the most successful ARIMA models in estimating the number of fully vaccinated people against COVID-19 in the US, Asia, Europe, Africa, South America, and the World,

Table 2
Various studies on COVID-19 epidemic using ARIMA.

Authos(s)	Forecasting	Study area
Alzahrani et al. [16]	Case	Sudi Arabistan
Dehesh et al. [17]	Case	China, Italy, South Korea, Iran, Thailand
Ceylan [18]	Case	Italy, Spain, France
Anne [19]	Case	India
Kumar et al. [20]	Case and death	Fifteen countries
Tandon et al. [21]	Case	Eight countries and Asia regions
Ilie et al. [22]	Case	Nine countries
Kufel [23]	Case	Selected European
Sharma [24]	Case	India
Chaurasia [25]	Death	World
Argawu [26]	Case	Ten countries
Lukman et al. [27]	Case	SA, Nigeria, Ghana, Egypt
Perone [28]	Case	Italy, Russia, USA
Maheshwari et al. [29]	Case and death	India
Katoch et al. [30]	Case	India
Kırbaş et al. [31]	Case	Eight countries

Table 5
Performance comparison of candidate ARIMA models for the test dataset.

Region	Candidate models	AIC	RMSE	MAPE	Selected model
US	ARIMA (2,2,2)	3603	3330748	2.38	ARIMA (5,2,2)
	ARIMA (1,2,0)	3626	3151914	2.26	
	ARIMA (0,2,1)	3617	3651831	2.62	
	ARIMA (1,2,2)	3602	3272585	2.32	
	ARIMA (0,2,2)	3603	3349863	2.38	
	ARIMA (1,2,1)	3601	3299527	2.35	
	ARIMA (2,2,1)	3600	3193971	2.25	
	ARIMA (2,2,0)	3628	3060394	2.20	
	ARIMA (3,2,1)	3560	2678414	1.83	
	ARIMA (3,2,0)	3594	3373309	2.39	
	ARIMA (4,2,1)	3558	2536176	1.72	
	ARIMA (4,2,0)	3575	3265123	2.28	
	ARIMA (5,2,1)	3554	2234625	1.48	
	ARIMA (5,2,0)	3559	2398972	1.60	
	ARIMA (5,2,2)	3552	1964099	1.27	
	ARIMA (4,2,1)	3553	2536176	1.72	
	ARIMA (5,2,3)	3554	1983322	1.29	
	ARIMA (4,2,3)	3555	1996855	1.30	
Asia	ARIMA (2,2,2)	3966	4508909	4.81	ARIMA (1,2,3)
	ARIMA (0,2,0)	4010	13619654	14.08	
	ARIMA (1,2,0)	3992	12518718	12.97	
	ARIMA (0,2,1)	3964	4480895	4.77	
	ARIMA (1,2,1)	3963	4540281	4.84	
	ARIMA (2,2,1)	3964	4513366	4.81	
	ARIMA (1,2,2)	3964	4475124	4.77	
	ARIMA (0,2,2)	3963	4521102	4.82	
	ARIMA (0,2,3)	3965	4516004	4.81	
	ARIMA (1,2,3)	3966	4503660	4.80	
Europe	ARIMA (2,2,2)	4041	6372484	5.78	ARIMA (5,2,0)
	ARIMA (0,2,0)	4087	7442914	6.77	
	ARIMA (1,2,0)	4085	8084573	7.37	
	ARIMA (0,2,1)	4080	6017086	5.45	
	ARIMA (1,2,2)	4062	6364820	5.77	
	ARIMA (2,2,1)	4060	6221668	5.64	
	ARIMA (3,2,2)	4036	4816791	4.29	
	ARIMA (3,2,1)	4052	5759615	5.19	
	ARIMA (4,2,1)	4034	5219494	4.66	
	ARIMA (4,2,0)	4055	5219494	4.66	
	ARIMA (5,2,1)	4030	5075843	4.52	
	ARIMA (5,2,0)	3468	4783587	4.24	
Africa	ARIMA (2,2,2)	2161	824239	11.33	ARIMA (2,2,1)
	ARIMA (0,2,0)	2188	950504	13.17	
	ARIMA (1,2,0)	2170	932713	12.91	
	ARIMA (0,2,1)	2155	823725	11.32	
	ARIMA (1,2,1)	2157	823566	11.32	
	ARIMA (2,2,1)	2159	823488	11.32	
	ARIMA (1,2,2)	2159	823568	11.32	
	ARIMA (0,2,2)	2457	823448	11.32	
South America	ARIMA (2,2,2)	3465	322142	0.91	ARIMA (1,2,1)
	ARIMA (0,2,0)	3529	4085583	11.56	
	ARIMA (1,2,0)	3528	3878743	10.97	
	ARIMA (0,2,1)	3493	471956	1.22	
	ARIMA (1,2,2)	3479	400317	1.07	
	ARIMA (2,2,1)	3470	771082	2.06	
	ARIMA (2,2,2)	3467	322142	0.91	
	ARIMA (1,2,1)	3488	327430	0.90	
	ARIMA (1,2,3)	3472	533316	1.51	
	ARIMA (3,2,1)	3470	851619	2.36	
	ARIMA (3,2,3)	3468	364552	1.02	
World	ARIMA (0,2,0)	4352	26246314	6.57	ARIMA (5,2,1)
	ARIMA (1,2,0)	4344	22964805	5.76	
	ARIMA (0,2,1)	4332	9017266	2.32	
	ARIMA (1,2,1)	4321	9178706	2.37	
	ARIMA (2,2,1)	4322	9163123	2.36	
	ARIMA (1,2,2)	4323	9172276	2.37	
	ARIMA (0,2,2)	4322	8869982	2.29	
	ARIMA (2,2,0)	4345	20828062	5.23	
	ARIMA (4,2,0)	3070	9321756	2.37	
	ARIMA (4,2,1)	3696	5685208	1.46	
	ARIMA (5,2,0)	3698	6805696	1.74	
	ARIMA (5,2,1)	3693	5412285	1.39	
	ARIMA (5,2,2)	3690	5724002	1.46	

Table 6
Ljung–Box statistic of selected ARIMA models.

	Model	χ^2	df	p-value
US	ARIMA(5,2,2)	10.55	18	0.91
Asia	ARIMA(1,2,3)	24.15	18	0.15
Europe	ARIMA(5,2,0)	22.26	18	0.22
Africa	ARIMA(2,2,1)	21.09	18	0.27
South America	ARIMA(1,2,1)	27.54	18	0.07
World	ARIMA(5,2,1)	24.64	19	0.17

- To forecast the total number of fully vaccinated people in the US, Asia, Europe, Africa, South America, and the World in the near future,
- To determine the people fully vaccinated against COVID-19 rate of the population at the begging of the summer in the US, Asia, Europe, Africa, South America, and the World to achieve herd immunity.

2. Material and method

2.1. Data collection

In this study, the number of people fully vaccinated against COVID-19 in the US, Asia, Europe, Africa, South America and the World were used to model and validate the ARIMA. The dataset were obtained from the Our World in Data [39]. The datasets used in the study start from the first date of fully vaccinated people against COVID-19 and ends on May 22, 2021. So the number of data varies for each region. The date of the first fully vaccinated people of the investigated regions and data size are given in Table 3.

2.2. ARIMA model

ARIMA, also known as the Box-Jenkins model, is shown as ARIMA (p, d, q). Parameter p in the model is the order of autoregression, parameter d is the degree of difference, and parameter q is the order of the moving average. The ARIMA model consists of Autoregressive (AR), Moving Average (MA) and "I" stands for integration sections. Three basic ARIMA models for a stationary time series are mathematically representing as follows [40];

Autoregressive model of order p or AR(p) model:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{1}$$

Moving-average model of order q or MA(q):

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{2}$$

Autoregressive moving average model of order p and q or ARMA(p,q):

$$y_t = \delta + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{3}$$

Where, ϕ is the autoregression and θ is the moving average parameters. y_t is the actual value at a time t. δ is the constant. The random disturbance term ε_t is assumed to be white noise which means it is independently identically distributed with mean 0 and a common variance for all terms.

2.3. Model selection criteria

Akaike Information Criterion (AIC), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) criteria were used to select the most successful ARIMA models for the dataset. The AIC shows how well the model fits the observed series. The most suitable model for the dataset is that have smallest

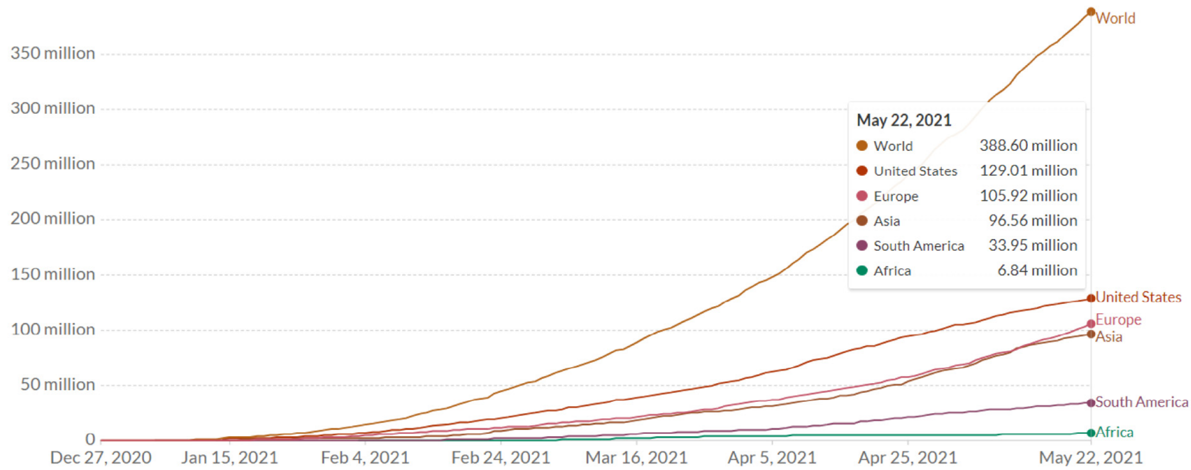

Fig. 1. Number of people fully vaccinated against COVID-19 in the examined regions [39].

Table 7
Forecasting of the number of people fully vaccinated against COVID-19 according to ARIMA models.

Date	US	Asia	Europe	Africa	South America	World
	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast
23.05.2021	130.033 M	97.761 M	107.840 M	6.953 M	34.325 M	393.781 M
24.05.2021	130.923 M	99.004 M	109.760 M	7.081 M	34.729 M	399.043 M
25.05.2021	131.749 M	100.246 M	111.856 M	7.193 M	35.140 M	404.305 M
26.05.2021	132.680 M	101.488 M	114.175 M	7.313 M	35.553 M	409.568 M
27.05.2021	133.755 M	102.730 M	116.558 M	7.423 M	35.967 M	414.830 M
28.05.2021	134.941 M	103.973 M	118.901 M	7.537 M	36.381 M	420.093 M
29.05.2021	136.099 M	105.215 M	121.062 M	7.644 M	36.794 M	425.355 M
30.05.2021	137.144 M	106.457 M	123.110 M	7.753 M	37.208 M	430.617 M
31.05.2021	138.059 M	107.700 M	125.167 M	7.856 M	37.622 M	435.880 M
1.06.2021	138.939 M	108.942 M	127.323 M	7.961 M	38.036 M	441.142 M

Table 8
Populations and COVID-19 information of regions.

Region	2021 population	People fully vaccinated (22.05.2021)	People fully vaccinated % (22.05.2021)	Forecasted people fully vaccinated (01.06.2021)
US	332.77 M	129.01 M	38.76	138.94 M
Asia	4.68 B	96.56 M	2.06	108.94 M
Europe	748.05 M	105.92 M	14.16	127.32 M
Africa	1.37 B	6.84 M	0.50	7.96 M
South America	433.99 M	33.95 M	7.82	38.04 M
World	7.87 B	388.60 M	4.93	441.14 M

AIC, RMSE and MAPE value [41]. These performance criteria can be calculated by using Eqs. (4)–(6), respectively.

$$\mathrm{AIC}\,(\mathrm{m}) = -2\log\mathrm{L} + 2m \tag{4}$$

Where, L is the likelihood function, m is the total number of parameters in the model.

$$\mathrm{RMSE} = \sqrt{\frac{(a_1 - p_1)^2 + (a_2 - p_2)^2 + \cdots + (a_n - p_n)^2}{n}} \tag{5}$$

$$\mathrm{MAPE} = 100\% \times \frac{|a_1 - p_1| + |a_1 - p_1| + \cdots + |a_n - p_n|}{n} \tag{6}$$

Where, p is the predicted value, a is the actual value. It is desirable for the above three error measurement criteria to be lower. Error is zero indicates that it is a statistically perfect model.

3. Results

Vaccination has started to be administered in countries on different dates. Until May 22, 2021, 129.01 million in the US, 96.56 million in Asia, 105.92 million in Europe, 33.95 million in South America, 6.84 million in Africa, and 388.60 million people fully vaccinated against COVID-19 in the World (Fig. 1).

The process of ARIMA modeling consists of four consecutive steps: identification of model, parameters estimation, diagnostic checking, and forecasting. In the first step, it is identified whether the time series variable is stationary and non-stationary. If the series is non-stationary, it is converted to stationary.

Time-series graphs are given in Fig. 2 to observe the stationarity of the dataset. Original time-series graphs are shown in Fig. 2 A. It is seen that there is an upward trend in all regions examined, that is, the original series is non-stationary. Therefore, the first-order difference was applied to the original data to stabilize the mean of the people fully vaccinated against COVID-19. The first-order difference of the series is shown in Fig. 3B. When the first-difference series are examined, trends of all series still observed, so the second-order difference was applied. As seen in Fig. 2C, after the second-order difference was taken, all series became stationary. Augmented Dickey–Fuller (ADF) test was applied to confirm the stationarity of the time series and the results are given in Table 4. ADF test results also show that the time series stabilize after the second difference was taken. This shows us that the parameter d is 2 in the ARIMA model.

Candidate ARIMA models' test results are given in Table 5. ARIMA models with minimum AIC, RMSE and MAPE criteria were chosen as the best models. Accordingly, the ARIMA(5,2,2),

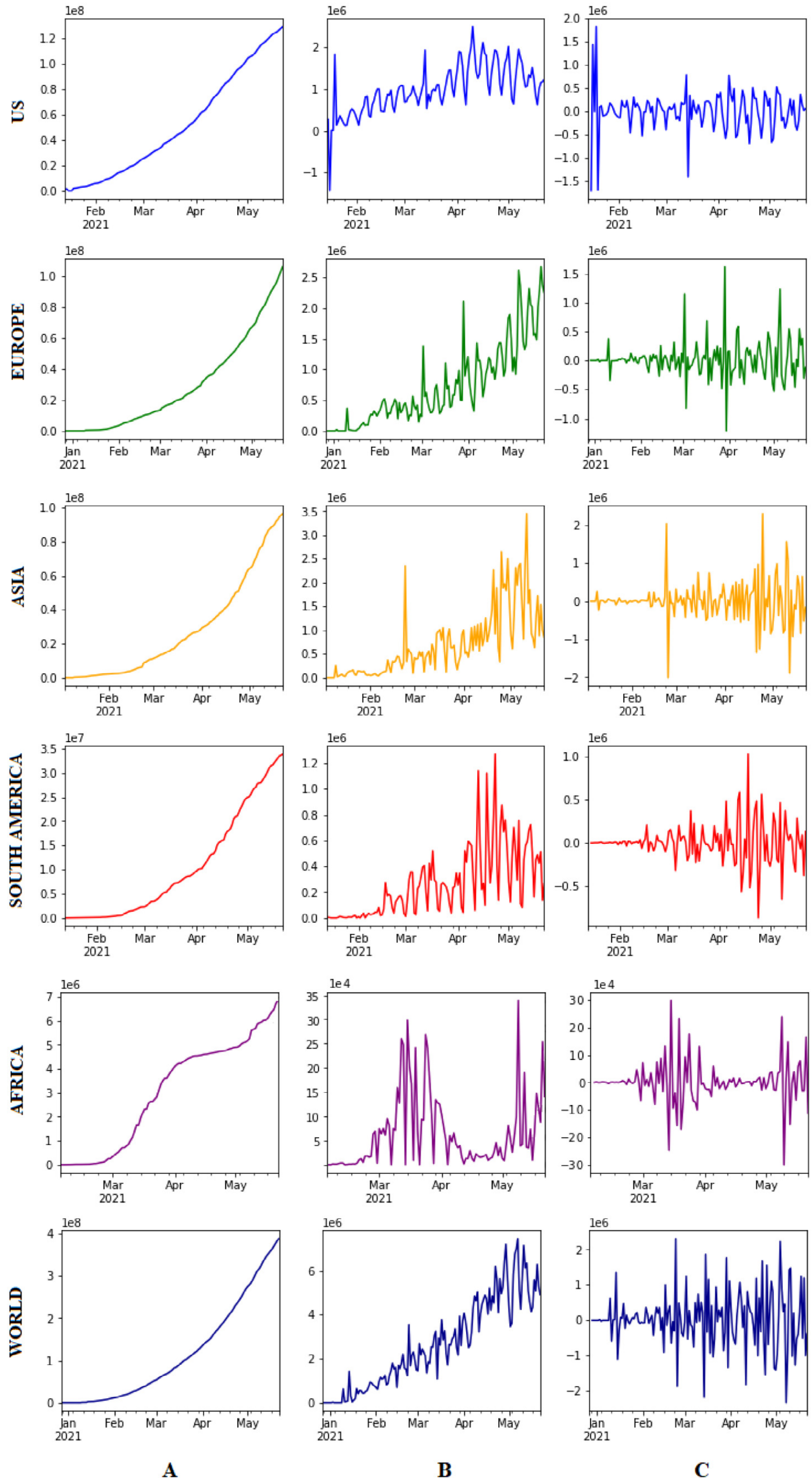

Fig. 2. (A) Original series, (B) first-order difference series, and (C) second-order difference series.

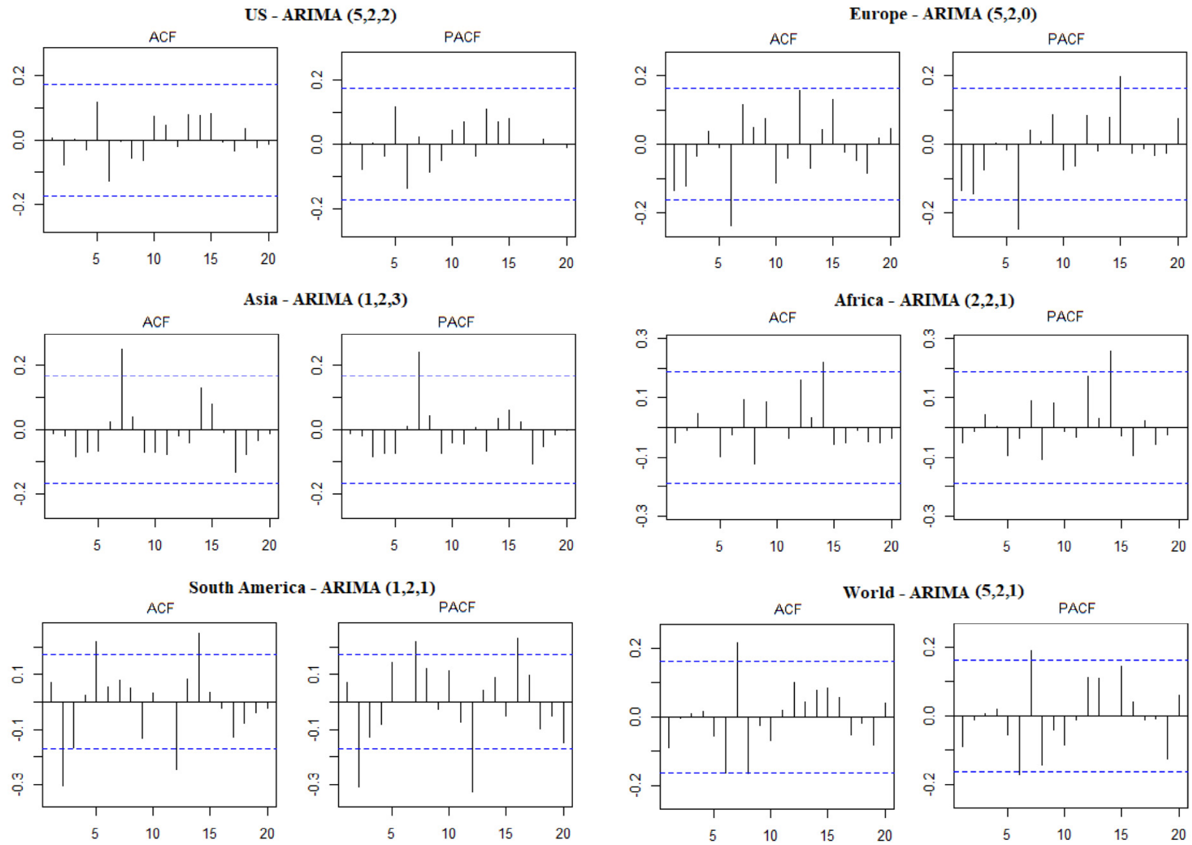

Fig. 3. Residuals of ARIMA models.

ARIMA(1,2,3), ARIMA(5,2,0), ARIMA(2,1,2), ARIMA(2,2,1) and ARIMA(0,2,2) models were selected for the US, Asia, Europe, Africa, South America and the World, respectively. The selected ARIMA models fitted the data reasonably well with a minimum $MAPE_{US}$ = 1.27, $MAPE_{Asia}$ = 4.80, $MAPE_{Europe}$ = 4.24, $MAPE_{Africa}$= 11.32, $MAPE_{SouthAmerica}$ = 0.90, and $MAPE_{World}$ = 1.39 values.

The ACF and PACF graphs of the residuals for the best fitted ARIMA models are presented in Fig. 3. The ACF determine whether the previous value in the series is related to the following value. The PACF shows the amount of correlation between a variable and a lag of itself. When ACF and PACF graphs are examined, it is seen that the residuals generally do not exceed significant boundaries. Box–Ljung test was used to check the residuals are white noise or not. The null hypothesis (H_0) for a Box–Ljung test is the residuals are independently distributed. Therefore, it is desirable to reject the null hypothesis. Large p-value indicate that the residuals have no remaining autocorrelations, i.e., they resemble white noise. Box–Ljung test results of models are given in Table 6.

Fig. 3 shows that the residuals are white noise since all the autocorrelation and partial autocorrelation coefficients are small and within two standard deviations at a 5% level of significance. Also, by Ljung–Box statistic results correlations are not significant and hence the residuals are white noise.

With these best models determined, the number of people fully vaccinated against COVID-19 was forecasted with 80%–95% confidence intervals (CI) (Fig. 4). Forecasting of the number of people fully vaccinated against COVID-19 for the next ten days according to ARIMA models is given in Table 7.

4. Discussions

The main purpose of this study is to determine the full vaccination rate in the US, Asia, Europe, Africa, South America, and the World on June 1, 2021. For this purpose, the number of fully vaccinated people in these regions was forecasted using the ARIMA time series model. Thus, they will be able to observe how far countries are from the threshold level required to achieve herd immunity.

With the determined ARIMA models, the number of people fully vaccinated against COVID-19 was estimated on June 1, 2021.

The populations, vaccination information, and the forecasted number of full vaccination people of the regions examined in this study are given in Table 8. Almost all countries desire to administer vaccines to at least 50% of their population until the beginning of the summer. This seems unlikely based on the total number of full vaccinated people estimated by ARIMA models. The ratio of the number of people fully vaccinated against COVID-19 to the population of the examined regions on June 1, 2021 is shown in Fig. 5.

According to the estimation results of ARIMA models, about 139 million, 109 million, 127 million, 8 million, 38 million, and 441 million people will be fully vaccinated in the US, Asia, Europe, Africa, South America, and the World, respectively (on June 1, 2021). The 2021 population of these regions are 333 million, 4.68 billion, 748 million, 1.37 billion, 434 million, and 7.87 billion, respectively. On June 1, 2021, it is predicted that fully vaccinated to 41.8% of the people in the US, 2.3% in Asia, 17% in Europe, 0.6% in Africa, 8.8% in South America, and 5.6% in the World. According to the findings obtained as a result of the study, the US reaches the highest level in the fully vaccinated rate on June 1, 2021. Nevertheless, none of the examined regions will reach the herd immunity threshold until June 1, 2021.

These days when the numbers of COVID-19 cases and deaths are at their peak again, the health systems of many countries have collapsed and they have started to impose curfews again.

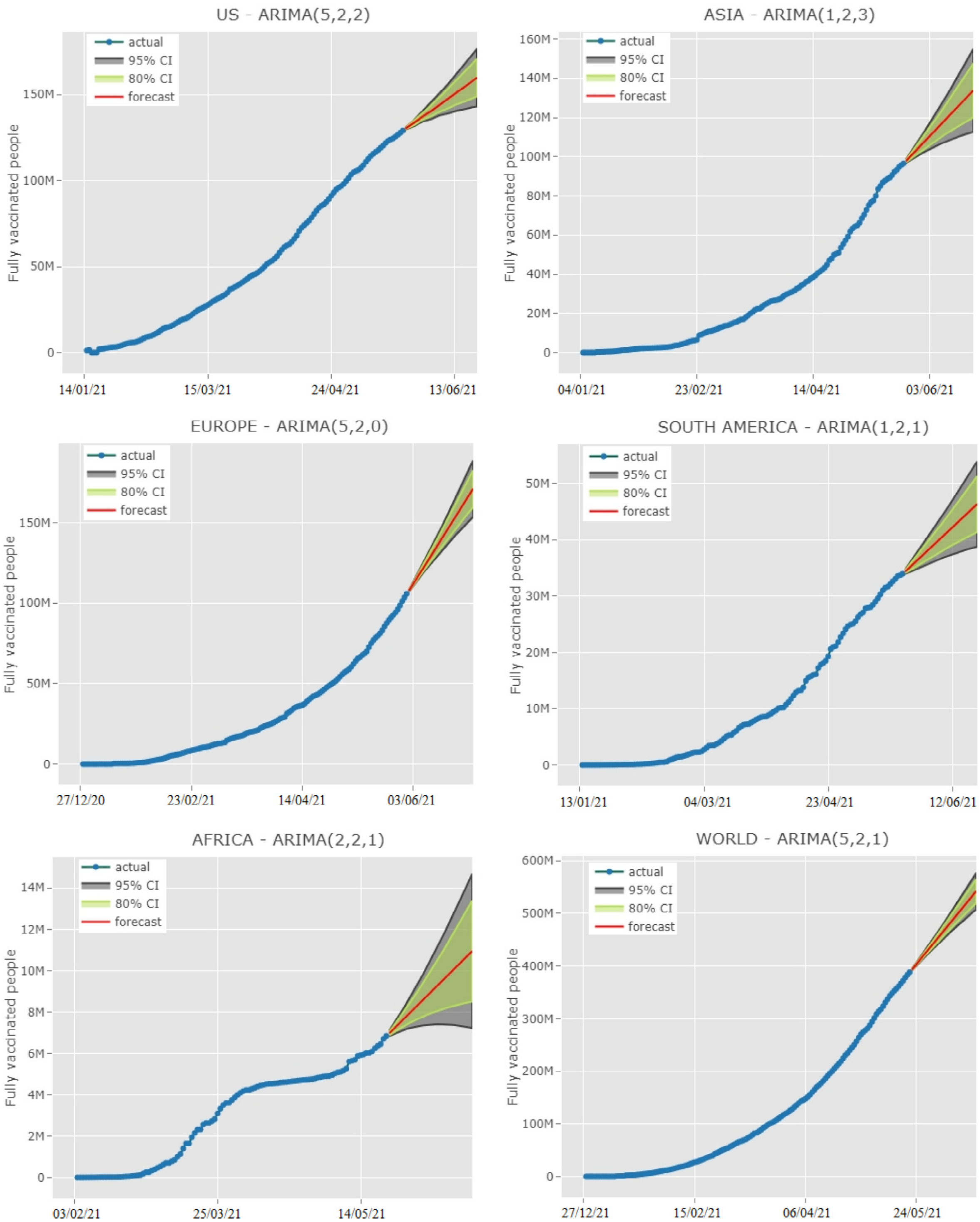

Fig. 4. Time series plots for the next 30 days according to best ARIMA models with 80%–95% CI.

In addition to the tragic deaths, prohibitions affect the country's economies negatively. To prevent all these negativities, it is thought that vaccination rates should reach quickly the threshold level required to provide herd immunity. Vaccination rates in Fig. 5 show that high-income countries should provide vaccines and healthcare support to low and middle-income countries. In addition, vaccine companies need to make plans to increase vaccine production rates, and countries to provide more vaccines and to apply vaccines quickly. Nevertheless, it has been reported that some mutant viruses are able to vaccine-escaped and antibody-resistant [42]. The rise of new variants may require the monitoring of immune escape variants, force second-generation production that addresses globally circulating variants, and already vaccinated people may be again vaccinated.

5. Conclusion

SARS-Cov-2 virus, which causes COVID-19 disease, is highly contagious. All countries quickly started vaccination development after the coronavirus was declared a global emergency

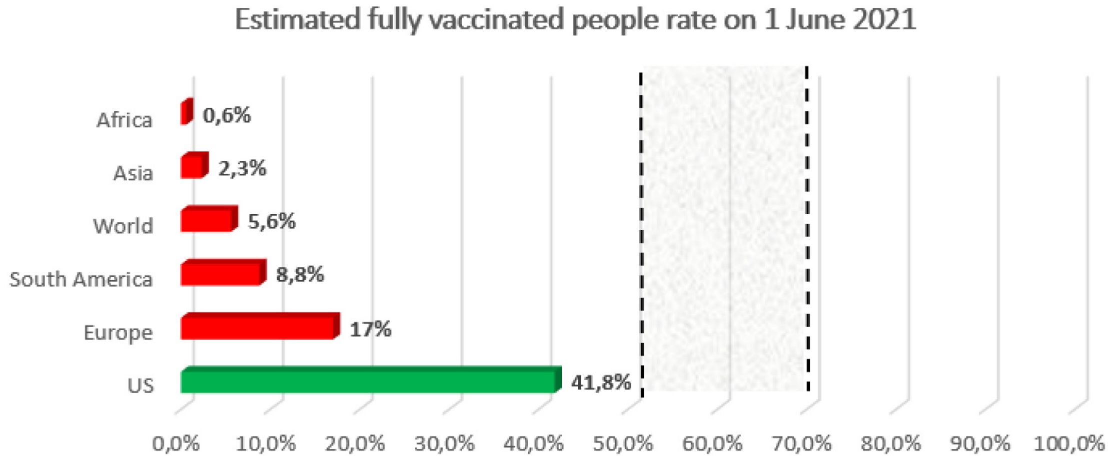

Fig. 5. The ratio of the estimated number of fully vaccinated people in examined regions.

by WHO. Up to now, more than ten COVID-19 vaccines have been approved for use. Although the efficacy of these vaccines is quite high, herd immunity is required to safely slow or stop the COVID-19 epidemic. The fastest way to achieve herd immunity is vaccination.

In this study, the total number of full vaccination people against COVID-19 in the US, Asia, Europe, Africa, South America, and the World on the next ten days was estimated using the ARIMA time series method. ARIMA models have been formulated with different parameters and the most successful models have been selected. AIC, RMSE, and MAPE criteria were used to compare model success. Models with the lowest values were selected to estimate the number of vaccines to be administered in the future. ARIMA(5,2,2), ARIMA(1,2,3), ARIMA(5,2,0), ARIMA(2,2,1), ARIMA(1,2,1) and ARIMA(5,2,1) models were selected for the US, Asia, Europe, Africa, South America, and the World, respectively. The selected ARIMA models fitted the data reasonably well with a minimum $MAPE_{US} = 1.27$, $MAPE_{Asia} = 4.80$, $MAPE_{Europe} = 4.24$, $MAPE_{Africa} = 11.32$, $MAPE_{SouthAmerica} = 0.90$, and $MAPE_{World} = 1.39$ values. According to the results of this study, on June 1, 2021, in the US, Asia, Europe, Africa, South America, and, the World of 139 million, 109 million, 127 million, 8 million, 38 million, and 441 million people will be fully vaccinated, respectively. This result shows that beginning of the summer, 41.8%, 2.3%, 17%, 0.6%, 8.8%, and 5.6% of people will be fully vaccinated in the US, Asia, Europe, Africa, South America, and the World, respectively. It has been observed that the other regions except the US are quite far from the herd immunity threshold level. The future goal of our study is to forecast the number of fully vaccinated people in the future with deep learning time series models. Thus, the performances of the ARIMA model can be compared with the deep learning methods.

CRediT authorship contribution statement

Pınar Cihan: Writing – original draft, methodology, software, validation, visualization, Writing – review and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

No funding to declare.

References

[1] World Health Organization (WHO), Novel Coronavirus (2019-NCoV), Situation Report, 2021, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200207-sitrep-18-ncov.pdf?sfvrsn=fa644293_2. (Accessed 9 February 2021).
[2] F. Di Gennaro, D. Pizzol, C. Marotta, M. Antunes, V. Racalbuto, N. Veronese, L. Smith, Coronavirus diseases (COVID-19) current status and future perspectives: a narrative review, Int. J. Environ. Res. Public Health 17 (2020) 2690, https://doi.org/10.3390/ijerph17082690.
[3] W.H. Chen, U. Strych, P.J. Hotez, M.E. Bottazzi, The SARS-CoV-2 vaccine pipeline: an overview, Curr. Trop. Med. Rep. (2020) 1–4, https://doi.org/10.1007/s40475-020-00201-6.
[4] H. Lu, Drug treatment options for the 2019-new coronavirus (2019-nCoV), Biosci. Trends 14 (2020) 69–71, https://doi.org/10.5582/bst.2020.01020.
[5] T. Pillaiyar, S. Meenakshisundaram, M. Manickam, Recent discovery and development of inhibitors targeting coronaviruses, Drug Discov. Today 25 (2020) 668–688, https://doi.org/10.1016/j.drudis.2020.01.015.
[6] Y. Jin, H. Yang, W. Ji, W. Wu, S. Chen, W. Zhang, G. Duan, Virology, epidemiology, pathogenesis, and control of COVID-19, Viruses 12 (2020) 372, https://doi.org/10.3390/v12040372.
[7] F. Kowalzik, D. Schreiner, C. Jensen, D. Teschner, S. Gehring, F. Zepp, mRNA-based vaccines, Vaccines 9 (2021) 390, https://doi.org/10.3390/vaccines9040390.
[8] F.P. Polack, S.J. Thomas, N. Kitchin, J.Absalon, A. Gurtman, S. Lockhart, et al., Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine, N. Engl. J. Med. 383 (2020) 2603–2615, https://doi.org/10.1056/NEJMoa2034577.
[9] M. Voysey, S.A.C. Clemens, S.A. Madhi, L.Y. Weckx, P.M. Folegatti, P.K. Aley, B. Angus, V.L. Baillie, S.L. Barnabas, Q.E. Bhorat, et al., Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: An interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK, Lancet 397 (2021) 99–111, https://doi.org/10.1016/S0140-6736(20)32661-1.
[10] P. Olliaro, E. Torreele, M. Vaillant, COVID-19 vaccine efficacy and effectiveness—the elephant (not) in the room, Lancet Microbe (2021) https://doi.org/10.1016/S2666-5247(21)00069-0.
[11] G.N. Peshimam, U. Farooq, CanSinoBIO's COVID-19 vaccine 65.7% effective in global trials, Pakistan official says, Reuters 8 (2021).
[12] Y. Yan, Y. Pang, Z. Lyu, R. Wang, X. Wu, C. You, et al., The COVID-19 vaccines: Recent development, challenges and prospects, Vaccines 9 (2021) 349, https://doi.org/10.3390/vaccines9040349.
[13] J. Cohen, Dosing debates, transparency issues roil vaccine rollouts, Science 371 (2021) 109–110, https://doi.org/10.1126/science.371.6525.109.
[14] E.E. Walsh, R.W. Frenck Jr., A.R. Falsey, N. Kitchin, J. Absalon, A. Gurtman, et al., Safety and immunogenicity of two RNA-based Covid-19 vaccine candidates, N. Engl. J. Med. 383 (2020) 2439–2450, https://doi.org/10.1056/NEJMoa2027906.
[15] C.D. Funk, C. Laferrière, A. Ardakani, Target product profile analysis of COVID-19 vaccines in phase iii clinical trials and beyond: an early 2021 perspective, Viruses 13 (2021) 418, https://doi.org/10.3390/v13030418.
[16] S.I. Alzahrani, I.A. Aljamaan, E.A. Al-Fakih, Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions, J. Infect. Publ. Health 13 (2020) 914–919, https://doi.org/10.1016/j.jiph.2020.06.001.
[17] T. Dehesh, H. Mardani-Fard, P. Dehesh, Forecasting of covid-19 confirmed cases in different countries with arima models, 2020, MedRxiv, https://doi.org/10.1101/2020.03.13.20035345.

[18] Z. Ceylan, Estimation of COVID-19 prevalence in Italy, Spain, and France, Sci. Total Environ. 729 (2020) 138817, https://doi.org/10.1016/j.scitotenv.2020.138817.
[19] R. Anne, ARIMA modelling of predicting COVID-19 infections, 2020, medRxiv, https://doi.org/10.1101/2020.04.18.20070631.
[20] P. Kumar, H. Kalita, S. Patairiya, Y.D. Sharma, C. Nanda, M. Rani, J. Rahmani, A.S. Bhagavathula, Forecasting the dynamics of COVID-19 pandemic in top 15 countries in 2020: ARIMA model with machine learning approach, 2020, MedRxiv, https://doi.org/10.1101/2020.03.30.20046227.
[21] H. Tandon, P. Ranjan, T. Chakraborty, V. Suhag, Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future, 2020, arXiv preprint arXiv:2004.07859.
[22] O.-D. Ilie, R.-O. Cojocariu, A. Ciobica, S.-I. Timofte, I. Mavroudis, B. Doroftei, Forecasting the spreading of COVID-19 across nine countries from Europe, Asia, and the American continents using the ARIMA models, Microorganisms 8 (2020) 1158, https://doi.org/10.3390/microorganisms8081158.
[23] T. Kufel, ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries, equilibrium, Q. J. Econ. Econ. Policy 15 (2020) 181–204.
[24] V.K. Sharma, U. Nigam, Modeling and forecasting of Covid-19 growth curve in India, Trans. Indian Natl. Acad. Eng. 5 (2020) 697–710, https://doi.org/10.1007/s41403-020-00165-z.
[25] V. Chaurasia, S. Pal, COVID-19 pandemic: ARIMA and regression model-based worldwide death cases predictions, SN Comput. Sci. 1 (2020) 1–12, https://doi.org/10.1007/s42979-020-00298-6.
[26] A. Argawu, Modeling and forecasting of COVID-19 new cases in the top 10 infected African countries using regression and time series models, 2020, https://doi.org/10.1101/2020.09.23.20200113, medRxiv.
[27] A.F. Lukman, R.I. Rauf, O. Abiodun, O. Oludoun, K. Ayinde, R.O. Ogundokun, COVID-19 prevalence estimation: Four most affected African countries, Infect. Dis. Model. 5 (2020) 827–838, https://doi.org/10.1016/j.idm.2020.10.002.
[28] G. Perone, ARIMA forecasting of COVID-19 incidence in Italy, Russia, and the USA, Russia, and the USA, 2020, https://doi.org/10.2139/ssrn.3612402, Available at SSRN.
[29] H. Maheshwari, D. Yadav, U. Chandra, D. Rai, Forecasting epidemic spread of COVID-19 in India using arima model and effectiveness of lockdown, Adv. Math. Sci. J. 9 (2020) 3419–3432, https://doi.org/10.37418/amsj.9.6.22.
[30] R. Katoch, A. Sidhu, An application of ARIMA model to forecast the dynamics of COVID-19 epidemic in India, Global Bus. Rev. (2021) 0972150920988653, https://doi.org/10.1177/0972150920988653.
[31] İ. Kırbaş, A. Sözen, A.D. Tuncer, F.Ş. Kazancıoğlu, Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches, Chaos Solitons Fractals 138 (2020) 110015, https://doi.org/10.1016/j.chaos.2020.110015.
[32] V. Rémy, N. Largeron, S. Quilici, S. Carroll, The economic value of vaccination: why prevention is wealth, J. Mark. Access Health Policy 3 (2015) 29284, https://doi.org/10.3402/jmahp.v3.29284.
[33] S.B. Omer, I. Yildirim, H.P. Forman, Herd immunity and implications for SARS-CoV-2 control, JAMA 324 (2020) 2095–2096, https://doi.org/10.1001/jama.2020.20892.
[34] L.-t. Cao, H.-h. Liu, J. Li, X.-d. Yin, Y. Duan, J. Wang, Relationship of meteorological factors and human brucellosis in hebei province, China, Sci. Total Environ. 703 (2020) 135491, https://doi.org/10.1016/j.scitotenv.2019.135491.
[35] A. Earnest, M.I. Chen, D. Ng, L.Y. Sin, Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore, BMC Health Serv. Res. 5 (2005) 1–8, https://doi.org/10.1186/1472-6963-5-36.
[36] Z. He, H. Tao, Epidemiology and ARIMA model of positive-rate of influenza viruses among children in wuhan, China: A nine-year retrospective study, Int. J. Infect. Dis. 74 (2018) 61–70, https://doi.org/10.1016/j.ijid.2018.07.003.
[37] J. Gaudart, O. Touré, N. Dessay, A. lassane Dicko, S. Ranque, L. Forest, J. Demongeot, O.K. Doumbo, Modelling malaria incidence with environmental dependency in a locality of Sudanese savannah area, Mali, Malar. J. 8 (2009) 1–12, https://doi.org/10.1186/1475-2875-8-61.
[38] P. Guan, D.-S. Huang, B.-S. Zhou, Forecasting model for the incidence of hepatitis A based on artificial neural network, World J. Gastroenterol.: WJG 10 (2004) 3579, https://doi.org/10.3748/wjg.v10.i24.3579.
[39] H. Ritchie, Coronavirus Source Data, European Center for Disease Prevention and Control, 2021, https://ourworldindata.org/coronavirus-source-data. (Accessed 22 May 2021).
[40] P.J. Brockwell, P.J. Brockwell, R.A. Davis, R.A. Davis, Introduction to Time Series and Forecasting, Springer, 2016.
[41] G. Jain, B. Mallick, A study of time series models ARIMA and ETS, 2017, Available at SSRN 2898968, https://doi.org/10.2139/ssrn.2898968.
[42] E.C. Thomson, L.E. Rosen, J.G. Shepherd, R. Spreafico, A. da Silva Filipe, J.A. Wojcechowskyj, et al., Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity, Cell 184 (2021) 1171–1187, https://doi.org/10.1016/j.cell.2021.01.037.