Research Article

The Analysis of the Incidence Rate of the COVID-19 Pandemic Based on Segmented Regression for Kuwait and Saudi Arabia

Muhammad Haroon Shaukat,[1] Naif Alotaibi [ID],[2] Ijaz Hussain [ID],[1] and Mansour Shrahili[3]

[1]*Department of Statistics, Quaid-I-Azam University, Islamabad, Pakistan*
[2]*Department of Mathematics and Statistics, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia*
[3]*Department of Statistics and Operations Research, King Saud University, Riyadh 11451, Saudi Arabia*

Correspondence should be addressed to Naif Alotaibi; nmaalotaibi@imamu.edu.sa and Ijaz Hussain; ijaz@qau.edu.pk

Since the initial detection of the novel coronavirus in Wuhan, China, at the end of 2019, the virus has spread rapidly worldwide and has become a global health threat. Due to rising infections, it has a significant impact on society as well as the economy. Although vaccines and treatment are available now, there is a need to control the pandemic's spread by appropriate strategies and policies. This study evaluates incidence such as positive rates and mortality rates through breakpoints, which were not undertaken in Kuwait and Saudi Arabia. In this regard, we have two-fold objectives: (1) to forecast the cumulative confirmed cases and death cases and (2) to compute the incidence rate within two consecutive days and evaluate the forecasted periods. The autoregressive integrated moving average model is used to forecast the cumulative confirmed cases and death cases for two months. The segmented regression model is used to split the pandemic time series into six periods and compute the incidence rate. Our results show that cumulative confirmed cases will reach 335733 in Kuwait and 445805 in Saudi Arabia by the beginning of June 2021. The cumulative death cases will reach 1830 and 7283 in Kuwait and Saudi Arabia, respectively. However, the positive rate will increase during the forecasted period in both countries, while the death rate will decrease in Kuwait and increase in Saudi Arabia. The study results can help public health organizations and decision-makers to control the spread of infectious diseases.

1. Introduction

During the past two decades, several severe outbreaks such as the severe acute respiratory syndrome coronavirus (SARS-CoV), H1N1 influenza, and Middle East respiratory syndrome coronavirus (MERS-CoV) have been reported in Middle East countries [1]. Most recently, a pandemic of unknown respiratory infections was reported on December 31, 2019, in Wuhan, China [1, 2]. Initially, the pandemic expeditiously spread from Hubei province to all provinces of China and then other countries. 22,112 confirmed cases were reported in Wuhan and 31,481 in mainland China on February 6, 2020 [3]. The World Health Organization (WHO) announced that the pandemic is named coronavirus disease 2019 (COVID-19) [4] and it was also declared as a global pandemic on March 11, 2020 [5]. The first case of

COVID-19 was directly linked to Huanan Seafood Wholesale market of Wuhan, where the transmission from the animal to human was considered as the main source. This virus can cross barriers and cause severe diseases in the society. Later, it was concluded that the COVID-19 outbreak is transmitted from person to person and that symptomatic people are the main source of spreading the infection [1]. Shortness of breath, fever, and cough are the main symptoms of COVID-19 [6]. The new virus is a highly infectious disease and has spread quickly worldwide [2]; the first case was diagnosed on February 24, 2020, in Kuwait [7] and on March 2, 2020, in Saudi Arabia [8]. Infections are on the rise around the world and pose a serious threat to human lives. The impact of the pandemic was not limited to the health of people's lives, but it has a huge impact on the world economy [5]. Many governments worldwide have introduced strict

curfew, partial lockdown, and quarantine and called for social distance and work from home options to reduce the spread of the outbreak [2]. This pandemic raised many questions. Among these questions are the following: (i) What will be the duration of the pandemic's peak? (ii) How many people will be affected during the peak of the pandemic? (iii) What will be the infection rate in the future? Many researchers around the world have performed several statistical analyses. At a time when the number of reported pandemics is growing rapidly, forecasting is vital to curb further spread [9]. The study in [8] has collected the daily confirmed, death, and recovery cases from March 2 to May 15, 2020. The aim of the study is to forecast the confirmed cases across Saudi Arabia by logistic growth (LG) and susceptible infected recovered (SIR) model. The authors have predicted a total of 69,979 to 79,000 cases and stated that the pandemic would reach its final stage by the end of June 2020. It is very important to know the spread rate of any pandemic at a specific time. The authors in [10] have presented a statistical framework and analyzed the growth and double rate of COVID-19 for six European countries by breakpoints. The work in [11] analyzed the propagation characteristic of COVID-19 with SARS and MERS by the growth model. The daily time series of COVID-19 is collected from January 21 to March 18, 2020, of four regions of China, while SARS and MERS data are collected from a different region of the world where there was a significant influence. The results indicate that the growth rate of a pandemic is twice as compared to SARS and MERS. It means that this outbreak can spread without intervention and cases can double within two to three days. The authors in [12] have introduced a new class of statistical model to provide the best description of COVID-19. They have used the daily confirmed, death, and recovery cases up to April 18, 2020, of Asian countries and studied the submodel of the proposed method known as extended Weibull distribution in detail. Their results indicate that the proposed model may provide a good fit to the death cases of COVID-19. Another study in [13] has presented the proposed NG-Weibull distribution to describe the best description of daily death cases of Iran and China from January 23 to April 2020. They suggest that the proposed model fits the daily death cases very closely to the actual situation. The authors in [14] have used the daily infected and recovered cases from February 2 to April 15, 2020. The purpose of their study is to evaluate the impact of intervention policy on the transmission rate of South Korea. The SIR model was utilized with and without joining points. They have observed that the spread rate has fallen to 0.23 after the breakpoint estimate on March 7, 2020. The researchers in [7] have used the deterministic and stochastic models to estimate the size of COVID-19 before and after repatriation. They have found that the estimated reproductive number (R_0) is 2.2 in Kuwait before the repatriation and until April 10, 2020. Even after the repatriation plan, it confirms the effectiveness of containment measures to control the transmission in Kuwait's state. The authors in [15] have used the daily confirmed cases from February 23 to May 7, 2020, to understand the temporal and spatial pattern for the two socioeconomic communities' citizen residents

and migrant workers. They have found that the outbreak is continuously growing ($R_0 > 2$), and it is an indication of significant transmission. Also, it is stated that the migrant worker's area has found a significant spread. However, the government's intervention measures have significantly reduced the infection rates in migrant workers' areas.

Due to increasing cases of COVID-19, the spread is not slowing down in Kuwait and Saudi Arabia. It is essential to know the outbreaks' trend, duration, peak, and incidence rate within a specific period, which have not been reviewed for both countries. In this paper, we will forecast the pandemic cases and then give insight into the incidence rate within two days as well as across the periods. The analysis of incidence rate and forecasting during the outbreak becomes essential and useful for preventing further spread, which can help in making political decisions such as staying at home and locking down. The rest of the article is arranged as follows: Section 2 includes the study area and methodology. Sections 3 and 4 cover the results and discussion of forecasting and incidence rates. Section 5 contains the conclusion and recommendations.

2. Methods

2.1. Study Area. All daily cumulative time series of COVID-19 in Kuwait and Saudi Arabia have been obtained from the following source: https://data.humdata.org/dataset/5dff64bc-a671-48da-aa87-2ca40d7abf02. This website updates the global data of the outbreak countrywise. The study area is represented in Figure 1. The daily cumulative time series consists of two variables, namely, confirmed and death cases. Therefore, the daily cumulative time series of confirmed and instances of death is included from February 24, 2020, to April 11, 2021, for Kuwait. Likewise, the size of the daily cumulative time series of confirmed and death cases is included from March 2, 2020, to April 11, 2021, for Saudi Arabia.

2.2. Autoregressive Integrated Moving Average Model (ARIMA). The study in [16] presented an autoregressive integrated moving average (ARIMA) model as the modification of the autoregressive moving average (ARMA) model. The ARMA model is only applicable when the series is stationary, while the ARIMA model can deal with both stationary and nonstationary series. The ARIMA model has three parameters, such as p, d, and q. The orders p and q are associated with autoregressive (AR) and moving average (MA) models, whereas d represents the differencing parameter and is used to make a stationary series. According to [17, 18], the general ARIMA model can be defined as

$$\phi(\beta)\nabla^d z_t = \varphi(\beta)e_t, \tag{1}$$

where $\phi(\beta)$ is associated with AR part of order p, while $\varphi(\beta)$ is associated with MA part of an order q. Moreover, the notation ∇^d indicates a different parameter. ARIMA's construction was based on three stages: identification,

FIGURE 1: Location of the study area.

parameter estimation, and model diagnostic. The time series analysis has been done on R version 3.6.1 by the "forecast" package.

2.3. Generalized Linear Model (GLM).

The generalized linear model (GLM) is a generalization of the classical linear regression model as it permits a response variable with error distribution except for a normal distribution. The GLM consists of three components: random component, systematic component, and link function. The GLM model permits the systematic component to connect with random components through link functions. According to [19], the general form of GLM model can be written as

$$g(\mu) = a + \beta_1 z_{i1} + \beta_2 z_{i2} + \cdots + \beta_k z_{ik}, \qquad (2)$$

where $g(\mu_i)$ represents the link function. The notations a and β_i $(i = 1, 2, 3, \ldots, k)$ indicate the intercept as well as slope coefficients. There are commonly used link functions, i.e.,

identity, log, power, square root, and logit. When the response variable is count, then the Poisson model is frequently used as the response distribution. In the Poisson model, the logarithm link function, i.e., $\log(\mu)$, is utilized to explain the response mean μ in terms of an explanatory variable z. The general form of the Poisson link function is as follows:

$$
\begin{aligned}
y_0 &\sim \text{Pois}(\mu), \\
\log(\mu) &= a + \beta_1 z_{i1} + \beta_2 z_{i2} + \cdots + \beta_k z_{ik}.
\end{aligned}
\qquad (3)
$$

2.3.1. Segmented Regression Model (SRM).

The nonlinear function estimates connected through two, three, or more straight lines at unknown points are referred to as breakpoints, join points, or change points. A relationship between the mean response $E[y_0] = \mu$ and the explanatory variable z is explained by adding the linear term of the model. In Muggeo et al. [20], the general segmented relationship can be described as

$$g(\mu) = \beta_0 + \beta_1 z + \phi_1 (z - \varphi_1) + \phi_2 (z - \varphi_2) + \cdots + \phi_q(z - \varphi_q), \quad (4)$$

where φ_i $(i = 1, 2, 3, \ldots, q)$ represent the breakpoints. Additionally, the notations β_1 and ϕ_i $(i = 1, 2, 3, \ldots, q)$ define the slope and difference in slope. The segmented relationship described that the time axis is split into $q + 1$ intervals based on event occurrence patterns.

$$\log(\mu) = \beta_0 + \beta_1 z + \phi_1 (z - \varphi_1) + \phi_2 (z - \varphi_2) + \cdots + \phi_q(z - \varphi_q). \quad (5)$$

There are $q + 1$ regimes with slope β_1, $\beta_2 = \beta_1 + \phi_1$, and $\beta_{q+1} = \beta_1 + \sum_{i=1}^{q} \phi_q$. The growth rate can be obtained by $r_k = (e^{(\beta_k)} - 1)$ for each regime. The analysis has been done on R-GUI version 3.6.1 by using the "segmented package."

2.4. Model Assessment Tools

2.4.1. Akaike Information Criterion (AIC).
Several tools are available in the literature to measure the goodness of fit of models. One of the best commonly used tools is the Akaike Information Criterion (AIC) that estimates models' quality. According to [21], the AIC can be described as

$$AIC = -2LnL' + 2P, \quad (6)$$

where L' represents the likelihood and P is the parameter of the model. The minimum value of AIC among the several models is an indication of the preferred model.

2.4.2. Accuracy Measure Tools.
Several tools measure the accuracy of the stochastic model, but the most frequently used tools are mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). According to [22], these tools are as follows:

$$MAE = \frac{\sum_{i=1}^{m} \left| y_{0,\text{obs},i} - y_{0,\text{pred},i} \right|}{m},$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} \left(y_{0,\text{obs},i} - y_{0,\text{pred},i} \right)^2}{m}}, \quad (7)$$

$$MAPE = \frac{1}{m} \sum_{1=1}^{m} \left| \frac{y_{0,\text{obs},i} - y_{0,\text{pred},i}}{y_{0,\text{obs},i}} \right| * 100,$$

where $y_{0,\text{obs},i}$ and $y_{0,\text{pred},i}$ are the observed and predicted time series. Besides, the notation m indicates the length of the time series. The lowest value of these accuracy measure tools is an indication of an accurate forecast.

3. Results

In the present study, we have used a daily time series of cumulative confirmed and death cases of Kuwait and Saudi Arabia to describe the incidence rate within two consecutive days. Initially, the ARIMA model is applied to both cumulative time series data for two months from April 12 to June 11, 2021. In this regard, the time series plot is

constructed to visualize the cumulative case trend and can be observed in Figure 2.

From Figure 2, (a) since May 24, the number of confirmed cases in Saudi Arabia is increasing compared to that in Kuwait and (b) similarly, the number of death cases in Saudi Arabia is rising, while Kuwait has slightly dropped.

In the identification phase, the Anderson–Darling (AD) test has been applied to verify the daily cumulative stationary assumption of time series. Initially, the parameters of ARIMA models are estimated by visualization of autocorrelation function (ACF) and partial autocorrelation function (PACF) plot as described in [17]. Subsequently, several combinations of parameters were used to test the suitability of the ARIMA models. Their effectiveness has been verified by the lowest values of MAE, RMSE, and MAPE. Of the suitable ARIMA models, the best model was selected based on the minimum value of AIC. Besides, the parameters of the best-fitted model are estimated by maximum likelihood estimation (MLE) and details are available in Table 1. The next step is to inspect the residuals of the model. The residuals should be uncorrelated and normally distributed. Therefore, the autocorrelation in the residuals was tested by the Ljung–Box test and visualization of ACF. Furthermore, the normality assumption of residuals is confirmed by the Shapiro–Wilk test and visualization of a histogram. The estimation detail of the best-fitted ARIMA models for each cumulative time series is described in Table 2.

After this, all required conditions have been confirmed by residuals of the best-fitted ARIMA model. Therefore, the two-month ahead forecast, such as April 12 to June 11, 2021, is done for each cumulative time series. The forecast, along with the 95% confidence level, is shown in Figures 3 and 4.

From Figure 3, in Kuwait, if the current strategy is continuous, then the cumulative cases might reach 292139 in May and 335733 in June, respectively, while in Saudi Arabia, the forecast for cumulative confirmed cases might reach 422520 and 445805 in May and June 2021, respectively.

From Figure 4, the cumulative death cases in Kuwait might increase to reach 1612 in May 2021 and 1830 in June, respectively, while in Saudi Arabia, the cumulative cases might reach 7022 and 7283 in May and June 2021, respectively.

From April 12 to June 11, 2020, the forecasted cumulative time series is merged with the historical time series to construct a new time series to analyze the incidence rate. The incidence may refer to positive or confirmed cases of COVID-19 and death cases. Therefore, each country's latest cumulative time series has been used to fit the GLM with the Poisson link function. The aim of using a Poisson link function is to deal with count data. The complete estimation detail of GLM is described in Table 3.

The next step is splitting the time series into intervals. The intervals may refer to breakpoints, change points, and join points. The SRM is used to partition the new time series of confirmed and death cases and fit separate line segments at each interval to analyze the incidence rate within two days. The splitting aims to examine the incidence rate that can be observed with a specified period. Finally, the SRM has split
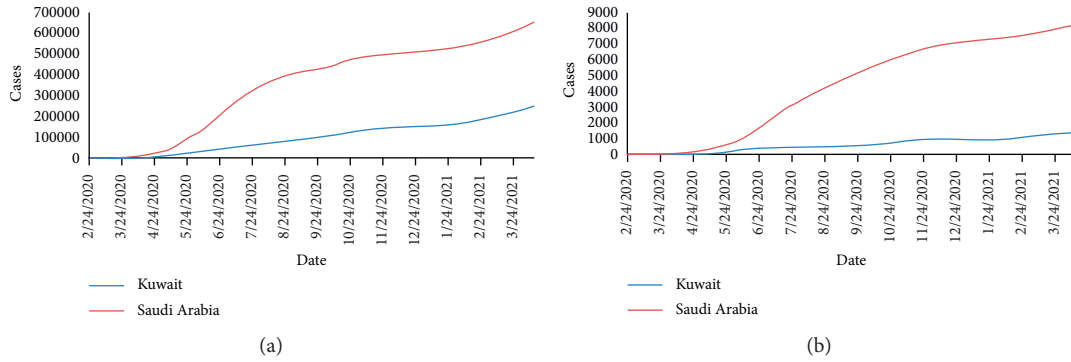
(a)   (b)

FIGURE 2: Time series plot of daily cumulative (a) confirmed COVID-19 cases and (b) death cases.

TABLE 1: Detail of parameters of ARIMA models.

Country	Case	Model	Parameter	Estimate	Standard error	Lower 95% CI	Upper 95% CI
Kuwait	Confirmed	(6,2,4)	ϕ_1	−0.7192	0.1096	−0.9340	−0.5045
			ϕ_2	−0.3652	0.0764	−0.5149	−0.2155
			ϕ_3	−0.6795	0.0618	−0.8006	−0.5583
			ϕ_4	−0.9080	0.0892	−1.0829	−0.7331
			ϕ_5	−0.5405	0.0637	−0.6654	−0.4155
			ϕ_6	−0.3110	0.0529	−0.4147	−0.2072
			φ_1	0.2315	0.1108	0.0144	0.4486
			φ_2	0.0348	0.0819	−0.1256	0.1953
			φ_3	0.4176	0.0700	0.2804	0.5547
			φ_4	0.6443	0.0926	0.4628	0.8259
Kuwait	Death	(7, 2, 2)	ϕ_1	0.0567	0.5089	−0.9408	1.0542
			ϕ_2	−0.7607	0.2708	−1.2915	−0.2298
			ϕ_3	−0.4490	0.2438	−0.9269	0.0289
			ϕ_4	−0.4795	0.1904	−0.8527	−0.1064
			ϕ_5	−0.2636	0.2024	−0.6603	0.1331
			ϕ_6	−0.2032	0.1247	−0.4476	0.0411
			ϕ_7	−0.1540	0.1062	−0.3622	0.0542
			φ_1	−0.6608	0.5147	−1.6697	0.3480
			φ_2	0.6337	0.1896	0.2620	1.0054
Saudi Arabia	Confirmed	(7,2,7)	ϕ_1	−0.1724	0.2641	−0.6901	0.3452
			ϕ_2	0.4314	0.1120	0.2119	0.6509
			ϕ_3	0.4799	0.1045	0.2751	0.6847
			ϕ_4	−0.0195	0.1470	−0.3076	0.2686
			ϕ_5	0.0337	0.0916	−0.1458	0.2132
			ϕ_6	−0.5144	0.0845	−0.6801	−0.3487
			ϕ_7	−0.2401	0.1814	−0.5957	0.1154
			φ_1	0.0901	0.2533	−0.4064	0.5865
			φ_2	−0.4385	0.1226	−0.6787	−0.1983
			φ_3	−0.2369	0.0884	−0.4102	−0.0637
			φ_4	0.0980	0.0986	−0.0953	0.2913
			φ_5	−0.3379	0.0834	−0.5013	−0.1746
			φ_6	0.5636	0.1175	0.3332	0.7940
			φ_7	0.4446	0.2126	0.0278	0.8614

TABLE 1: Continued.

Country	Case	Model	Parameter	Estimate	Standard error	Lower 95% CI	Upper 95% CI
Saudi Arabia	Death	(12, 2, 6)	ϕ_1	0.1861	0.1144	−0.0381	0.4102
			ϕ_2	−0.6983	0.1577	−1.0075	−0.3891
			ϕ_3	1.0515	0.1577	0.7424	1.3605
			ϕ_4	0.1685	0.1392	−0.1043	0.4413
			ϕ_5	0.1844	0.1199	−0.0505	0.4193
			ϕ_6	−0.1589	0.1057	−0.3661	0.0483
			ϕ_7	−0.4966	0.1070	−0.7064	−0.2869
			ϕ_8	−0.0602	0.0841	−0.2250	0.1046
			ϕ_9	−0.2952	0.0813	−0.4546	−0.1358
			ϕ_{10}	0.1658	0.0644	0.0397	0.2920
			ϕ_{11}	−0.0145	0.0538	−0.1200	0.0909
			ϕ_{12}	0.2560	0.0531	0.1519	0.3601
			φ_1	−0.7338	0.1115	−0.9524	−0.5151
			φ_2	1.0456	0.1999	0.6537	1.4374
			φ_3	−1.7246	0.2444	−2.2036	−1.2455
			φ_4	0.8096	0.2409	0.3375	1.2816
			φ_5	−0.5787	0.2090	−0.9883	−0.1690
			φ_6	0.6832	0.1174	0.4531	0.9132

TABLE 2: Estimation detail of best-fitted ARIMA models.

| Country	Cases	Parameter	AIC	MAE	RMSE	MAPE	Ljung–Box test		Shapiro–Wilk test	
							Statistic (Q)	P value	Statistic (W)	P value
Kuwait	Confirmed	(6, 2, 4)	4991.1900	73.1736	101.5751	0.9026	69.9440	0.5797	0.9660	3.40E-08
Kuwait	Death	(7, 2, 2)	1525.7400	1.3833	0.3766	0.9178	56.6190	0.7881	0.9876	0.002929
Saudi Arabia	Confirmed	(7, 2, 7)	5171.0600	74.1968	139.0189	1.3841	70.6900	0.3555	0.7433	2.20E-16
Saudi Arabia	Death	(12, 2, 6)	1911.1900	1.9343	2.7787	0.6540	71.5330	0.1269	0.9375	1.28E-11

(a)

(b)
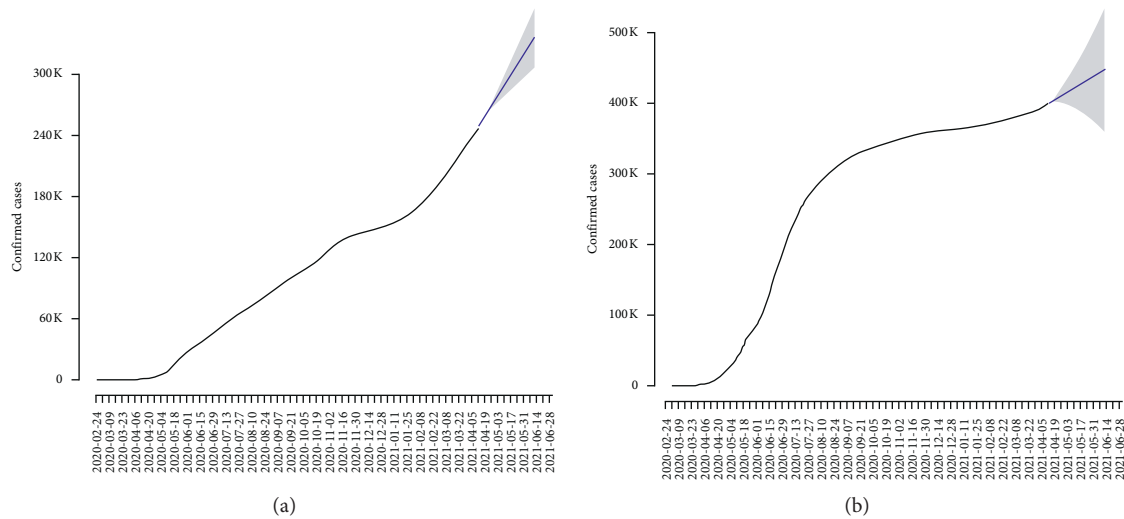
FIGURE 3: Forecast of cumulative confirmed cases with a 95% confidence level. (a) Kuwait. (b) Saudi Arabia.

the new time series into five breakpoints based on BIC's lowest value, and details can be seen in Table 4.

We have obtained five breakpoints that are connecting points between the lines. We consider these lines as periods where the trends of cases are similar. Moreover, the different periods are evident in different colors in the figures. Therefore, the incidence rate with a 95% confidence interval is calculated for each period and observed in Figures 5 and 6.

It is important to note that the vertical black dash line distinguishes between historical and forecast observations in both figures.

From Figure 5, it is found that the positive rate had increased, as period-1 (February 24 to May 19) 7.9032% of positive rate, and it is gradually decreased to 3.1656%, 1.3929%, 0.6649%, and 0.2167% in period-2, period-3, pe-riod-4, and period-5, respectively. The positive rate
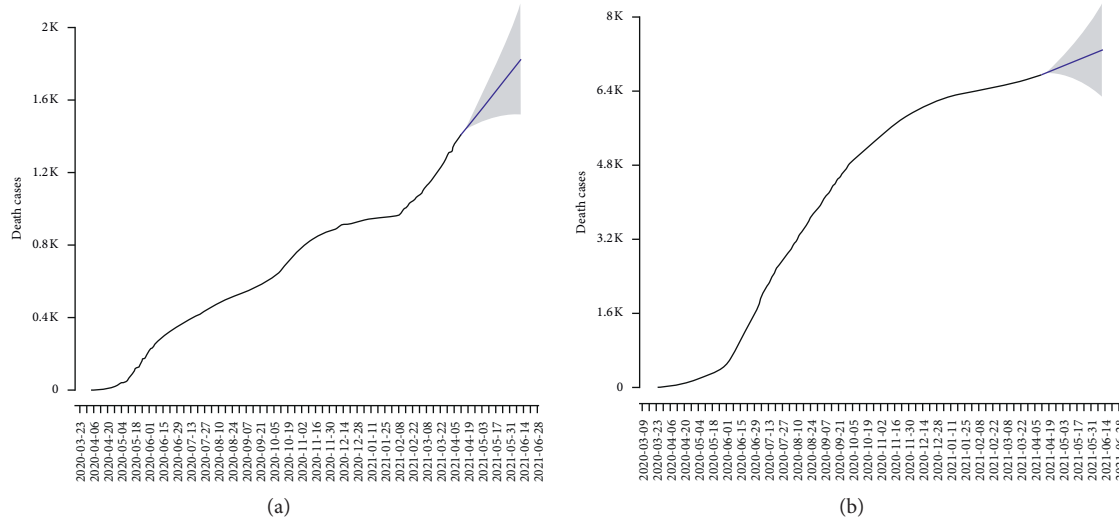
FIGURE 4: Forecast of cumulative death cases with 95% confidence level. (a) Kuwait. (b) Saudi Arabia.

TABLE 3: Parameter estimation of Kuwait and Saudi Arabia confirmed and death cases by the generalized linear model.

Country	Cases	AIC	Parameter	Estimate	Standard error	Z value	Pr (>|z|)
Kuwait	Confirmed	13724	Intercept	3.20 E+00	1.18E-02	271.07	2.00E-16
			Slope	7.61E-02	1.57E-04	486.17	2.00E-16
Kuwait	Death	3670	Intercept	4.74E-01	1.15E-01	4.14	3.47E-05
			Slope	9.83E-02	3.18E-03	30.88	2.00E-16
Saudi Arabia	Confirmed	30974	Intercept	4.57 E+00	8.31E-03	550.10	2.00E-16
			Slope	9.09E-02	1.49E-04	608.70	2.00E-16
Saudi Arabia	Death	4980	Intercept	3.29 E+00	2.16E-02	152.25	2.00E-16
			Slope	4.32E-02	3.01E-04	143.56	2.00E-16

TABLE 4: BIC values of breakpoints for Kuwait and Saudi Arabia confirmed and death cases.

Country	Cases	Break-1	Break-2	Break-3	Break-4	Break-5
Kuwait	Confirmed	345064	124058	57346	17205	13774
Kuwait	Death	5219	4577	3890	3711	3867
Saudi Arabia	Confirmed	740083	180503	173486	47434	31023
Saudi Arabia	Death	20976	6242	5306	5301	5029

increased to 0.5559% from January 25 to April 11, 2021, and then stayed stable in the forecast period (period-6). In Saudi Arabia, the positive rate was high in period-1, 9.5177% from March 2 to May 5, 2020, and then gradually changed. In period-2, the positive rate reached 3.1886% and then 1.6106%, 0.429%, and 0.0759% for period-2, period-3, period-4, and period-5, respectively. Moreover, the positive rate was 0.186% and stable for the forecast period and gradually increased after March 24, 2021 (period-6).

Figures 5(a) and 5(b) show that the positive rate is high in Saudi Arabia as compared to Kuwait during period-1. After that, the positive rate is decreased in the remaining periods for both countries. But Saudi Arabia has the lowest positive rate in comparison to Kuwait during period-6, which includes the forecast period. Thus, as we have observed in all historical periods, significant changes have become apparent.

As seen in Figure 6, in Kuwait, the mortality rate (10.329%) was high in period-1; then it is gradually decreased in period-2, period-3, and period-4 and then increased to reach 0.6136% during period-5. We can observe that the mortality rate is reduced up to 0.42289% during period-6. In Saudi Arabia, the mortality rate (4.4165%) is found to be the highest in period-1, and after that, it had a gradual decrease to reach 0.0835% till period-5, which then slightly increased to reach 0.1241% during the forecast period.

Figures 6(a) and 6(b) demonstrate that the mortality rate has decreased until period-4 in Kuwait and till period-5 in Saudi Arabia. But the abrupt changes have been found after the aforementioned periods in both countries. Overall, the mortality rate is lower in Saudi Arabia than in Kuwait during period-6.
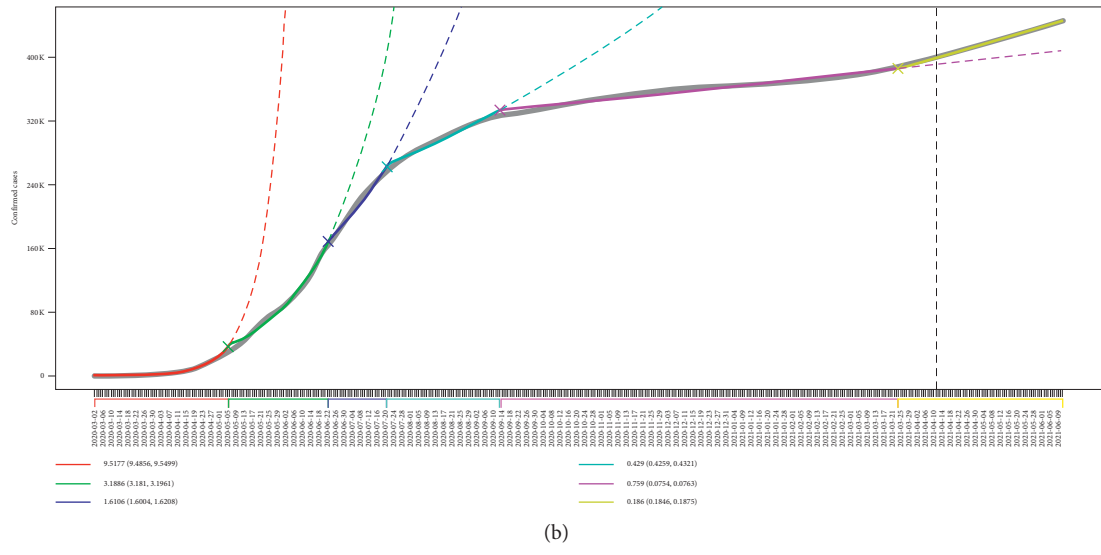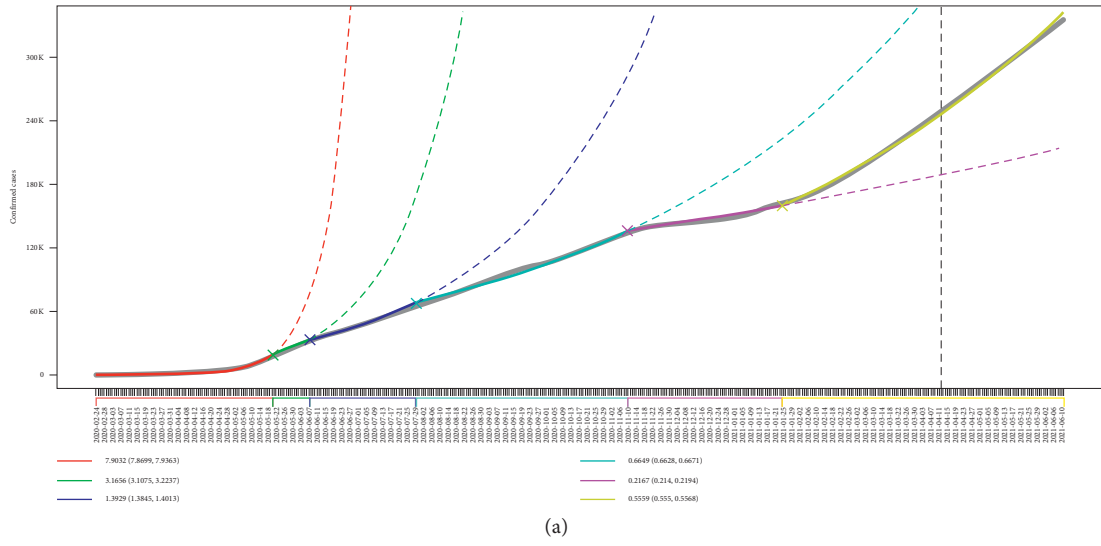
(a)

(b)

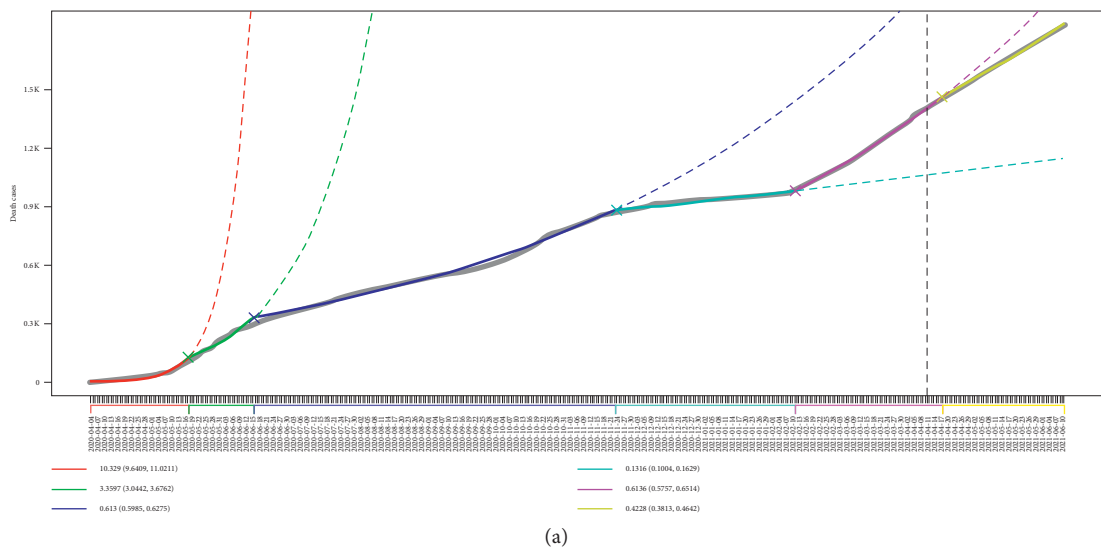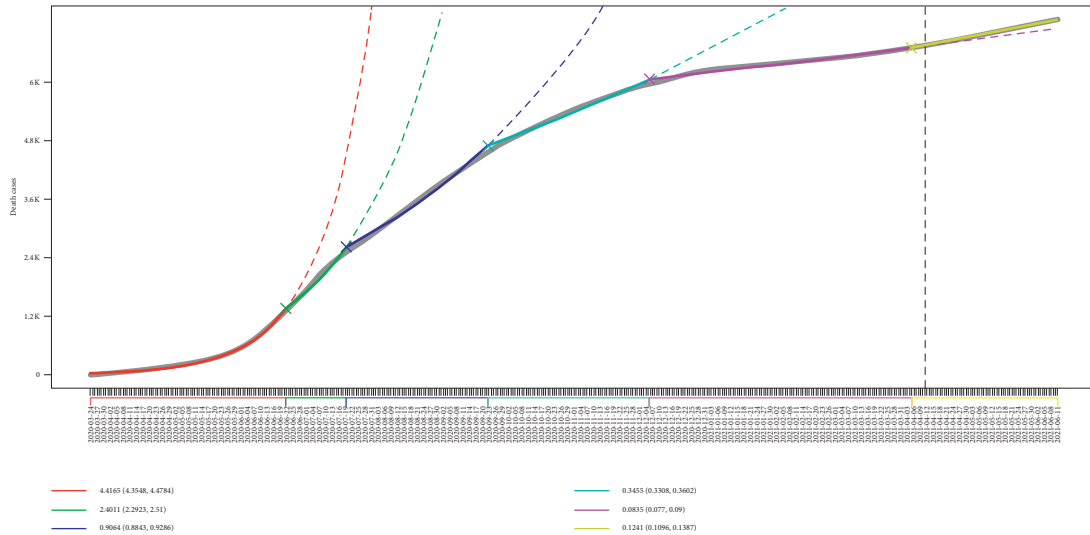FIGURE 5: Plot of positive rate for (a) Kuwait and (b) Saudi Arabia.

(a)

FIGURE 6: Continued.

Figure 6: Plot of mortality rate for (a) Kuwait and (b) Saudi Arabia.

4. Discussion

The world is currently facing the severe COVID-19 pandemic, which originated in Wuhan, China, and then spread to other countries. The confirmed cases were imported from Iran and Iraq by travelers in Kuwait and Saudi Arabia, respectively [23]. It is observed that the cumulative confirmed cases are increasing day by day in both countries. Still, there is a massive difference between the cumulative cases in Kuwait and Saudi Arabia (Figure 2(a)). Similarly, there is a lot of difference in cumulative death cases between both countries, but Kuwait has a slight decrease in mortality rate during September 2020, followed by a gradual increase (Figure 2(b)). Many factors play an essential role in the spread of COVID-19, such as population and community lifestyle [24]. However, the primary factor in the rising of cases in Saudi Arabia is the larger population than Kuwait. A previous study shows that the pandemic will be entered in the final phase at the end of June 2020 [8]. Unfortunately, the cases are gradually increasing until now. Before the situation worsens, we need to make an accurate prediction to take serious steps to control the pandemic's spread. The authors in [25] suggest that the ARIMA model is best to forecast daily cases, so we have utilized the ARIMA model to predict cumulative confirmed and death cases for both countries. Our forecast results indicate that if the current strategy of both governments remains the same, there will be a further increase in cumulative confirmed and death cases by June 12, 2021 (Figures 3 and 4). Furthermore, the SRM was used to explore the incidence rate within two days across the periods. Figure 5 shows that the positive rate is high in both countries during period-1 and then decreases to one-third during period-2 in Kuwait and Saudi Arabia.

In Kuwait, a strict curfew imposed from May 10 to May 30, 2020 [7], is the primary reason for reducing positive cases. Before the curfew, the Kuwait government has implemented a partial lockdown, such as 5 pm to 4 am, which was not as effective as strict curfew in the whole country. As likely to confirmed cases, the mortality rate was higher in period-1 than in other periods (Figure 6(a)). Similarly, the mortality rate is reduced to one-third in period-2 and then continuously decreased, and an abrupt rise was found in period-5. Therefore, we can say that the strict lockdown has reduced the peak of the pandemic in Kuwait, and it has proved to be very useful. Like Kuwait, Saudi Arabia had a higher positive rate during period-1 (Figure 5(b)), and the government imposed a partial lockdown on March 23 from 7 pm to 6 am. But, a high number of cases have been diagnosed during period-1 and the government had announced a complete lockdown on April 6 [26]. The lockdown significantly reduced the pandemic's spread, as we saw in period-2, where the positive rate of one-third had been identified. Figure 6(b) shows that the mortality rate was high in period-1 and then gradually decreased till period-5, and a slight increase was found in period-6.

Our forecast results indicate that if the current strategy continues, the number of cumulative confirmed and fatalities cases in both countries will increase by June 12, 2021. If we discuss the incidence rate, the positive rate will be increased during the forecast period (period-6) in both countries (Figure 5). But, the mortality rate is expected to decrease in Kuwait as compared to the previous period-5 and increase in Saudi Arabia during the forecast period (Figure 6). Overall, Kuwait's positive and mortality rates are expected to be higher than in Saudi Arabia during the forecast period (Figures 5 and 6). Our analysis suggests that the incidence rate may decrease further in the future but may require strict lockdown in both countries.

There are some limitations as the ARIMA model did not have automatic updates when the new observation was merged with the previous observation. Also, a long historical record may be required for obtaining prediction accuracy.

The ARIMA model is based on a linear structure. If the series has high nonlinearity, then the accuracy may decrease. There is no limit to the number of segmented variables and breakpoints in SRM. Because the segmented relationship assumes a linear pattern over each segment, response can follow a nonlinear trend.

5. Conclusion

This study granted more profound insight into the COVID-19 outbreak in Saudi Arabia and Kuwait during the intervention activities. According to our analysis, the pandemic's peak was during period-1, as it spreads rapidly in both countries after the first case is diagnosed, and then the transmission gradually decreases. Similarly, the mortality rate peaked during period-1 and has declined since then. Our forecast results reveal that the cumulative confirmed and death cases are expected to increase by the beginning of June in Kuwait and Saudi Arabia. Still, the positive cases may increase in both countries if the current strategy continues. At the same time, the mortality rates may decrease in Kuwait and increase in Saudi Arabia. Both governments can implement appropriate strategies and policies to avoid further transmission in the future. It is suggested that this platform provide valuable information about the incidence rate and provide guidance for future precautions.

Data Availability

All daily cumulative time series of COVID-19 of Kuwait and Saudi Arabia have been obtained from the following source: https://data.humdata.org/dataset/5dff64bc-a671-48da-aa87-2ca40d7abf02. This website updates the global data of the outbreak countrywise.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Di Napoli, "Features, evaluation, and treatment of coronavirus (COVID-19)," in *StatPearls [Internet]*Treasure Island (FL): StatPearls Publishing, Treasure island, FL, USA, PMID: 32150360, 2020.

[2] H. M. Singer, "Short-term Predictions of Country-specific Covid-19 Infection Rates Based on Power Law Scaling Exponents," 2020, http://arxiv.org/abs/2003.11997.

[3] D. Kang, H. Choi, J.-H. Kim, and J. Choi, "Spatial epidemic dynamics of the COVID-19 outbreak in China," *International Journal of Infectious Diseases*, vol. 94, pp. 96–102, 2020.

[4] C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, and P.-R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges," *International Journal of Antimicrobial Agents*, vol. 55, no. 3, Article ID 105924, 2020.

[5] X. Zhang, R. Ma, and L. Wang, "Predicting turning point, duration and attack rate of COVID-19 outbreaks in major Western countries," *Chaos, Solitons & Fractals*, vol. 135, p. 109829, 2020.

[6] S. Yang, P. Cao, P. Cao et al., "Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis," *Annals of Translational Medicine*, vol. 8, no. 4, p. 128, 2020.

[7] A. M. Almeshal, A. I. Almazrouee, M. R. Alenizi, and S. N. Alhajeri, "Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models," *Applied Sciences*, vol. 10, no. 10, 2020.

[8] D. Alboaneen, B. Pranggono, D. Alshammari, N. Alqahtani, and R. Alyaffer, "Predicting the epidemiological outbreak of the coronavirus disease 2019 (COVID-19) in Saudi Arabia," *International Journal of Environmental Research and Public Health*, vol. 17, no. 12, pp. 1–10, 2020.

[9] A. Cori, C. A. Donnelly, I. Dorigatti et al., "Key data for outbreak evaluation: building on the ebola experience," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 372, no. 1721, 2017.

[10] V. Muggeo, G. Sottile, M. Porcu, and V. M. R. Muggeo, "Modelling COVID-19 Outbreak: Segmented Regression to Assess Lockdown Effectiveness," April, 1–4, https://www.researchgate.net/publication/340664370, 2020.

[11] K. Liang, "Mathematical model of infection kinetics and its analysis for COVID-19, SARS and MERS," *Infection, Genetics and Evolution*, vol. 82, no. March, Article ID 104306, 2020.

[12] M. Zuo, S. K. Khosa, Z. Ahmad, and Z. Almaspoor, "Comparison of COVID-19 pandemic dynamics in Asian countries with statistical modeling," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 4296806, 16 pages, 2020.

[13] J. Zhao, Z. Ahmad, and Z. Almaspoor, "Modeling the COVID-19 pandemic dynamics in Iran and China," *CMC-Computers Materials & Continua*, vol. 67, no. 2, pp. 2111–2122, 2021.

[14] Y. J. Kim, M. H. Seo, and H. E. Yeom, "Estimating a breakpoint in the pattern of spread of COVID-19 in South Korea," *International Journal of Infectious Diseases*, vol. 97, pp. 360–364, 2020.

[15] M. A. Alkhamis, S. Al Youha, M. M. Khajah et al., "Spatio-temporal dynamics of the COVID-19 pandemic in the State of Kuwait," *International Journal of Infectious Diseases*, vol. 98, pp. 153–160, 2020.

[16] M. Geurts, G. E. P. Box, and G. M. Jenkins, "Time series analysis: forecasting and control," *Journal of Marketing Research*, vol. 14, no. 2, p. 269, 1977.

[17] A. K. Mishra and V. R. Desai, "Drought forecasting using stochastic models," *Stochastic Environmental Research and Risk Assessment*, vol. 19, no. 5, pp. 326–339, 2005.

[18] M. H. Shaukat, I. Hussain, M. Faisal et al., "Monthly drought prediction based on ensemble models," *PeerJ*, vol. 8, Article ID e9853, 2020.

[19] P. De Jong and G. Z. Heller, *Generalized Linear Models for Insurance Data*, Cambridge Books, Cambridge, UK, 2008.

[20] V. M. R. Muggeo, M. Attanasio, and M. Porcu, "A segmented regression model for event history data: an application to the fertility patterns in Italy," *Journal of Applied Statistics*, vol. 36, no. 9, pp. 973–988, 2009.

[21] H. Bozdogan, "Akaikes information criterion and recent developments in information complexity," *Journal of Mathematical Psychology*, vol. 44, no. 1, pp. 62–91, 2000, http://www.ingentaconnect.com/content/ap/mp/2000/00000044/00000001/art01277.

[22] K. S. Parmar and R. Bhardwaj, "Water quality management using statistical analysis and time-series prediction model," *Applied Water Science*, vol. 4, no. 4, pp. 425–434, 2014.

[23] M. E. Wilson and L. H. Chen, "Travellers give wings to novel coronavirus (2019-nCoV)," *Journal of Travel Medicine*, vol. 27, no. 2, pp. 1–3, 2020.

[24] S. Alrashed, N. Min-Allah, A. Saxena, I. Ali, and R. Mehmood, "Impact of lockdowns on the spread of COVID-19 in Saudi Arabia," *Informatics in Medicine Unlocked*, vol. 20, Article ID 100420, 2020.

[25] S. I. Alzahrani, I. A. Aljamaan, and E. A. Al-Fakih, "Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions," *Journal of Infection and Public Health*, vol. 13, no. 7, pp. 914–919, 2020.

[26] I. Anil and O. Alagha, "The impact of COVID-19 lockdown on the air quality of Eastern Province, Saudi Arabia," *Air Quality, Atmosphere, and Health*, vol. 14, 2020.