Research Article

Modeling and Simulation: A Study on Predicting the Outbreak of COVID-19 in Saudi Arabia

Ahmed Msmali, Mutum Zico, Idir Mechai, and Abdullah Ahmadini

Department of Mathematics, Faculty of Science, Jazan University, Jizan, Saudi Arabia

Correspondence should be addressed to Mutum Zico; mmeetei@jazanu.edu.sa

Received 9 January 2021; Revised 15 May 2021; Accepted 18 June 2021; Published 30 June 2021

Academic Editor: Li Li

Copyright © 2021 Ahmed Msmali et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The novel coronavirus disease (COVID-19) has resulted in an ongoing pandemic affecting the health system and economy of more than 200 countries worldwide. Mathematical models are used to predict the biological and epidemiological tendencies of an epidemic and to develop methods for controlling it. In this work, we use a mathematical model perspective to study the role of behavior change in slowing the spread of COVID-19 in Saudi Arabia. The real-time updated data from March 2, 2020, to January 8, 2021, were collected from the Saudi Ministry of Health, aiming to provide dynamic behaviors of the epidemic in Saudi Arabia. During this period, 363,692 people were infected, resulting in 6293 deaths, with a mortality rate of 1.73%. There was a weak positive relationship between the spread of infection and mortality \( R^2 = 0.459 \). We used the susceptible-exposed-infection-recovered (SEIR) model, a logistic growth model, with a special focus on the exposed, infected, and recovered individuals to simulate the final phase of the outbreak. The results indicate that social distancing, hygienic conditions, and travel limitations are crucial measures to prevent further spread of the epidemic.

1. Introduction

Coronaviruses are a large family of viruses that have spread among humans, animals, and livestock. The virus may spread from animals, such as bats, to humans through various intermediate hosts and cause severe respiratory syndrome [1]. They generally cause mild respiratory infections, dry cough, fever, and difficulty in breathing. In November 2019, the first COVID-19 outbreak occurred in Wuhan, the capital city of Hubei province and the seventh largest city of China [2–4]. In December 2019, the 25 first cases of infected people in Wuhan city were reported, including 7 critically ill cases. Since then, the epidemic outbreak has drawn serious concerns and global attention. At that time, China banned travel to and from Wuhan city and implemented a national emergency response. The suspected and confirmed cases were isolated, public transport was suspended, schools and entertainment venues were closed, and these control measures taken in China were taken as lessons to other countries around the world [5].

The Chinese authorities on January 7, 2020, identified the novel coronavirus as the causative agent, and on January 10, 2020, the World Health Organization (WHO) designated the novel coronavirus as 2019-nCoV. On the same day, the WHO released a wide range of interim guidance for all countries on how to prepare for and respond to the COVID-19 pandemic, which included a strategy and plan on how to monitor potentially infected people and collect and test samples and practical guidance on how to manage patients and control and mitigate the burden of infection in health centers. This guidance also highlighted the maintenance of the right drug supplies and effective communication with the public regarding the new virus [6]. On February 12, 2020, the WHO named the 2019-nCoV as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which was adopted as the official designation. The WHO officially declared the outbreak as a global pandemic because of the rapid spread of the virus to approximately 220 countries and territories around the world [7].
On March 2, 2020, the first case of COVID-19 infection in Saudi Arabia was identified and announced by the Ministry of Health. From the first week of March, the number of confirmed COVID-19 cases has gradually increased, reaching 2932 confirmed cases on April 9, 2020. A period of increasing infection cases was noticed in June and July. At present, new cases of infection are gradually decreasing. On January 8, 2021, 363,692 confirmed cases of COVID-19 infection with 355,382 recovered and 6293 deaths were recorded in Saudi Arabia (shown in Figure 1).

Mathematical modeling and simulations are used as important tools to predict the probability and severity of disease outbreaks and provide information to understand the dynamic behavior of the infection. Furthermore, mathematical models have been widely used as an essential tool for investigating the dynamics of the spread of infectious diseases such as foot-and-mouth disease, SARS, Ebola, and Zika [7–11]. Sun et al. [12] presented dynamical models for COVID-19 based on mathematical and data analyses to explore the effects of lockdown and medical resources in Wuhan. They showed that imposing lockdown measures in the city lowered the peak value of new infections and reduced the final scale. In [13], the effectiveness of lockdown was studied using the age-structured version of the susceptible-exposed-infection-recovered- (SEIR-) type for COVID-19. The authors suggested that the time of introducing lockdown for some fixed duration should be adjusted for each particular region to achieve maximum effect or should be maintained for a long time.

This study aims to provide a further understanding of the transmission dynamics of the virus infection and facilitate control measures using real data from March 2, 2020, to January 8, 2021. We studied the basic reproduction number of viruses during the current outbreak. The prediction of the pandemic tendency of COVID-19 in Saudi Arabia is important in the application of control and prevention measures by the Ministry of Health.

1.1. Related Works. The COVID-19 pandemic in all countries spread almost in the same scenario. During the first wave, the number of infected cases increased exponentially, and then, its growth declined. Mathematical models of different complexity levels are used to predict the dynamics of the spread of COVID-19. In this section, we present an overview of commonly used mathematical models to study the spread of COVID-19.

1.1.1. SIR Model. The classical model in epidemiology is the susceptible-infected-recovery (SIR) model which characterizes the spread of infectious disease through the majority of the population. The parameters used in this model helped to predict the spread of COVID-19 infection with the following presumptions: (a) The birth and death rates are ignored or remain equal to make a constant population size. (b) The effects of exterior factors on the parameters of the model are ignored. (c) The spread of the disease transmits only from humans to humans [14].

In SIR mode, the population is divided into three compartments, namely, S which is the number of susceptible people, I which is the number of infectious people, and R which is the number of recovered or deceased people. The system of ordinary differential equations (ODE) is stated in

\[
\frac{dS}{dt} = -\frac{\beta IS}{N},
\]

\[
\frac{dI}{dt} = -\frac{\beta IS}{N} - \gamma I,
\]

\[
\frac{dR}{dt} = \gamma I,
\]

where \(\beta\) and \(\gamma\) are the parameters representing the probability of infection and the average number of people recovered at time \(t\) (in days). Here, \(N = S + I + R\) is the total population.

Din et al. [15] used the SIR model to show that some factors such as exposure, death, and cure rates affect the susceptible, infected, and recovered population involving immigration. The simulated results for infection rate increase with involvement of the concern immigration rate. And with a smaller immigration rate, the infection rate declined when compared at higher order and vice versa.

The SIR model with stochastic approaches is likely to produce results different from deterministic models whenever the model is run for some parameters. Zhang et al. [16] study the effects of the environment on the spread of COVID-19 using stochastic modeling. The threshold of the stochastic COVID-19 model is determined according to the large and small noise.

1.1.2. SEIR Model. In SEIR, a new compartment named \(E\) for exposed people that are not yet infectious is added [14]. The classical SEIR model is given by the system of differential equations as follows:

\[
\frac{dS}{dt} = -\frac{\beta IS}{N},
\]

\[
\frac{dE}{dt} = -\frac{\beta IS}{N} - \gamma E,
\]

\[
\frac{dI}{dt} = \gamma E - \delta I,
\]

\[
\frac{dR}{dt} = \gamma I.
\]

All the parameters are similarly defined with the SIR model, with \(\delta\) being the incubation rate. And, \(N = S + I + R\) is the total population (constant) ignoring the birth and death rates.

Ashleigh et al. [17] proposed an aged-structured compartmental model of COVID-19 transmission in the population of Ontario, Canada. The model was stratified by age groups and used the modified susceptible-exposed-infectious-recovered model, incorporating additional compartments for public health, severities of clinical symptoms, and risk of hospital admission.
In [18], the SEIR model was used to study the dynamics of the spread and prevention of COVID-19. The uncertainty in case observation was also discussed in this model by modeling a Poisson-observed process of new symptomatic cases. The model transmission was considered as a geometric random walk process, and the transmission rate was simulated using the Monte Carlo simulation.

Adekunle et al. [19] analyzed various forms of mathematical models that are relevant for the containment, risk analysis, and features of COVID-19. The SEIR model seems the most reliable extension model because of its plausibility in explaining heterogeneous changes in behavior. And they concluded that mathematical models can significantly guide the implementation of public health decisions when properly expressed and estimated.

A modified SEIR model was proposed in [20], with six compartments, namely, susceptible-exposed-symptomatic-asymptomatic-hospitalized-recovered (SEYNHR). The differential equations for asymptomatic and hospitalized cases are given by

\[ \frac{dN}{dt} = \alpha N \left( \Phi_Y + K + \mu \right) N, \]

\[ \frac{dH}{dt} = \varepsilon Y - (r + \mu_2 + \mu) H, \]

where \( \Phi_N \) is the rate of immunity against the virus, \( K \) is the rate of transmission from compartment \( N \) to \( Y \), \( \varepsilon \) is the hospitalization rate, and \( \mu_2 \) is the death rate due to disease.

In [21], the modified SEIR under Atangana–Baleanu–Caputo (ABC) derivative with fractional order is used to show how the disease infection may be controlled and stabilized. Furthermore, simulation is performed using Adams–Bashforth numerical technique concerning results for two famous cities of China, Wuhan and Huanggang. Fractional models are more likely to be accurate in predicting the infection dynamics.

1.1.3. Logistic Growth Model. The logistic model is used to describe the rate of change in infectious persons depending on the rate of growth and capacity of the population. It has been used to model population growth in a region and bacterial growth in a broth and implemented in a binary decision-making process [22]. The logistic model equation with confirmed cases \( P \), growth rate \( r \), and the initial number of confirmed cases \( P_0 \) at time \( t = 0 \) is given as follows:

\[ P(t) = \frac{KP_0 e^{rt}}{K + P_0 \left( e^{rt} - 1 \right)}. \]

Here, \( P(t) \) is the cumulative confirmed cases of infection. The logistic model describes properly the growth rate of

![Figure 1: Number of (a) confirmed, (b) recovered, (c) death, and (d) cumulative confirmed cases due to COVID-19 in Saudi Arabia from 2nd March 2020 to 8th January 2021.](image-url)
infected people by COVID-19 and fitted with the available data. The use of a generalized logistic growth model improves the agreement with the real data; however, the prognostic number of the total number of cases is slightly higher than using the simple logistic model [23].

Batista [24] estimated the final size and peak time of COVID-19 in China, South Korea, and the rest of the world by using the logistical growth regression model. According to the model prediction, the daily epidemic size forecasts begin to converge and he concluded that the outbreak is under control. Wu et al. [25] employ the classical logistic growth model and generalized form of the model to compare the outbreak dynamics in the 29 provinces in China. The model estimate growth rate of infection decreases as fewer susceptible people were available with infection and better prediction at the later stage of the epidemic.

### 1.1.4. ARIMA Model

Autoregressive Integrated Moving Average (ARIMA) is used for the stochastic time-series model. It is assumed that the time series should be linear and follow the statistical distribution. The popular versions of the ARIMA model which are widely applied for the prediction approach are Autoregressive (AR), Moving Average (MA), and seasonal ARIMA (SARIMA) [26]. The mathematical formulation for the ARIMA $\alpha, \beta, \gamma$ model using lag polynomials $L$ is given in

\[ (1 - L)^{\phi} \varphi(L)(1 - L)^{\beta} \beta(L)y_t = \theta(L)\varepsilon_t, \quad (11) \]

\[ \left(1 - \sum_{i=1}^{\alpha} \varphi_i L^i\right) (1 - L)^{\phi} \varphi(L)(1 - L)^{\beta} \beta(L)y_t = \left(1 - \sum_{i=1}^{\gamma} \theta_i L^i\right)\varepsilon_t, \quad (12) \]

where $\alpha, \beta,$ and $\gamma$ are the autoregressive, integrated, and moving average parts of the model, respectively, and they are nonnegative integers. And, $\varphi_i$ and $\theta_i$ are the model coefficients.

The ARIMA model along with the wavelet-based forecasting model is used to generate short-term forecasting of the number of daily confirmed cases for the six countries [27]. Adiga et al. [28] used both ARIMA and logistic growth models to study the trend and to provide short- and long-term forecasting of the prevalence of COVID-19 cases and their dynamics. The main idea of the ARIMA model is to precisely estimate the short-term prevalence of recovered and death cases due to infection in Saudi Arabia. It has more advantage in short-term forecasting which gives more accurate projections of the pandemic. For long-term forecasting, they used a logistic growth model. The ARIMA model is widely used to predict the dynamics of COVID-19 disease in the most affected 15 countries and showed a realistic number very similar to the real data [29].

Chen et al. [30] developed a Bats-Hosts-Reservoir-People (BHRP) transmission network model to simulate the infection probability from bats to humans. Furthermore, they simplified the model as reservoir-people (RP) transmission. The estimated result of the basic reproduction number was 2.30. And, the simulation results showed that it was 3.58 from person to person. The results indicated that the infection rate for COVID-19 in some countries might be higher than for MERS.

Mao et al. [31] investigated the biological and epidemiological trends in the prevalence and mortality due to outbreaks of COVID-19 infection. The data of the outbreak were obtained from the World Health Organization. The study was done on gender-based analysis demonstrating that the identified cases consisted mostly of men with a median age range of 50–65 years and concluding that there was a large variation in the biological and epidemiological trends in both growth factors of the number of cases and the mortality cases.

Alrasheed et al. [32] built a network simulation model of the spread of COVID-19 in Saudi Arabia using the SIR model. A set of attributes such as age, gender, nationality, and location were defined for nodes in network simulation. They also computed the percentage of people that must be vaccinated to stop the epidemic.

We compare the ARIMA model with the compartmental models by highlighting some of their advantages and disadvantages. Table 1 shows the comparison between the ARIMA and compartmental models in long-term and short-term forecasting of the COVID-19 pandemic.

## 2. Materials and Methods

The study was conducted in the Department of Mathematics, College of Science, Jazan University, Saudi Arabia. Real-life data were collected based on the number of confirmed cases, recovered cases, and deaths due to COVID-19 infections in all regions of Saudi Arabia. Cumulative data were collected and obtained from the WHO and the official website of the Ministry of Health, Saudi Arabia (https://covid19.moh.gov.sa/). We also obtained data from search engines, including [7] and research articles published in [49]. We evaluated the reproduction number, which represents the average number of people to which an infected person can transmit the virus. Preliminary studies estimated the reproduction number to be between 1.5 and 3.5. The spread of the virus gradually stops with a reproduction number below 1.0. For comparison, the reproduction numbers for common flu and SARS are 1.3 and 2.0, respectively [7].

### 2.1. Ethical Consideration

In our study, the information on epidemiological tendencies and the current situation of COVID-19 infection was obtained from the WHO, Worldmeter-Coronavirus, and Saudi Ministry of Health Command and Control Center. Therefore, ethical approval was not required.

### 2.2. Statistical Data and Analysis

The updated data related to the cumulative number of infection and recovery cases were analyzed. Table 2 shows the number of confirmed and death cases in the seven major cities in Saudi Arabia. Regression analysis was performed using the data from Table 2. We determined the coefficient of determination $R^2$ as 0.34 and the $p$ value as 0.00014, which is considered significant. According to the population in each city, the city of Dhahran
had the highest percentage of confirmed (average 4.5%) and death cases (average 0.8%), whereas Tabuk had the lowest percentage of confirmed (average 0.89%) and death (average 0.018%) cases.

We collected data related to daily confirmed, recovered, and death cases due to COVID-19 released by the Saudi Ministry of Health from March 2, 2020, to January 8, 2021. The data were arranged in a matrix form with rows representing the date and columns representing the number of new confirmed cases, number of cumulative confirmed cases, and number of death cases. Table 3 shows the number of cases from the highest number of infections recorded on June 16, 2020. The monthly cumulative confirmed and death cases are listed in Table 4. The latest updated cases are presented in Table 5.

From the reported cases, the number of newly confirmed cases has been decreasing, with a higher recovery rate. This is due to the effective control and prevention measures imposed by the Saudi Arabian government. Figure 2 shows a statistical analysis used to interpret the real data. The analysis was performed by considering confirmed, recovered, and death cases. Figures 2(a) and 2(b) show the analysis with the daily confirmed cases as the independent variable and the death cases as the dependent variable. We obtained the following values: $R = 0.676$, $R^2 = 0.457$, adjusted $R^2 = 0.455$, standard error = 10.301, statistically significant $F < 0.001$, and $p < 0.001$.

Figures 2(c) and 2(d) show the analysis with the daily confirmed cases as the independent variable and the recovered cases as the dependent variable. We obtained the following values: $R = 0.754$, $R^2 = 0.568$, adjusted $R^2 = 0.567$, standard error = 825.3, statistically significant $F < 0.001$, and $p < 0.001$.

2.3. Models. We performed long-term forecasting of COVID-19 infection in real-time using SEIR and logistic growth models. We considered both continuous and discrete cases in the logistic model equation. These models assist the prevention measures provided by the Ministry of Health, Saudi Arabia, by predicting the number of infected people per day.

| Table 1: Comparison of research works for COVID-19 disease prediction using modeling. |
|---|---|---|
| Types of modeling | Advantages | Disadvantages |
| ARIMA | (i) The model prediction shows an acceptable accuracy for the daily number of cases in Italy, Spain, and France [33]. (ii) It achieves the smallest root mean square error, mean absolute percentage error, root mean square relative error, and highest $R^2$ value [35]. (iii) The model performance is more satisfactory than the SIR model as it results in minimum values of residuals [37]. (iv) It compensates residuals between the model and real data making the prediction value accurate [39]. (v) It is simple and easy to use. | (i) Variation in observed values and model specification leads to model unstable [34]. (ii) It is a pure learning algorithm and model parameters do not provide any information of epidemiological interest [36]. (iii) The model is not suitable for nonstationary data and hence fits for short-term forecasting [38]. (iv) The prediction accuracy has a direct relation with the number of observations [40]. (v) The related time-series data are assumed to be linear. |
| Compartmental models | (i) It predicts the dynamic nature of the disease spread [41]. (ii) Stochastic models successfully determine the transmission dynamics of COVID-19 through postpandemic period [43]. (iii) Prediction of cumulative confirmed cases to understand the unreported cases [45]. (iv) They present the effects of public health policies on the outcome of the pandemic [47]. (v) They perform better when modified to include more states and parameters. | (i) The models are mostly deterministic and work with large populations suitable for long-term forecasting [42]. (ii) The models have the drawback of having two unknown parameters (latent period and initial latent population) [44]. (iii) They cannot include contact tracing of the infected population unless vast assumptions are made [46]. (iv) The ODE-based compartmental models do not include environment factors [48]. (v) They are complex and require presumptions before modeling. |

| Table 2: Regression analysis for the confirmed and death cases due to COVID-19 in major cities of Saudi Arabia. |
|---|---|---|---|---|
| Major cities in Saudi Arabia | Number of confirmed cases (%) | Number of deaths (%) | $R^2$ |
| Riyadh | 60336 | 1.43 | 1237 | 0.030 | 0.676 |
| Makkah | 34999 | 2.64 | 2273 | 0.179 | 0.457 |
| Madinah | 22643 | 1.74 | 163 | 0.013 | 0.568 |
| Tabuk | 4036 | 0.89 | 83 | 0.018 | 0.455 |
| Dhahran | 4483 | 4.50 | 793 | 0.800 | 0.567 |
| Qassim | 14072 | 1.02 | 195 | 0.014 | 0.567 |
| Jazan | 3553 | 3.39 | 466 | 0.442 | 0.676 |
2.3.1. Susceptible-Exposed-Infected-Recovered Model. The simulation of the well-known SEIR model was performed to represent the dynamic behaviors of the spreading process of the COVID-19 epidemic in Saudi Arabia. The SEIR model framework is an epidemic spreading model inspired by [50]. In epidemiology, the SEIR model is a compartmental epidemic model widely used to characterize the outbreak of COVID-19. In this model, the spread of infection depends on the number of susceptible and infected populations. The incubation period is considered when individuals are infected but do not have symptoms. As the COVID-19 has a long incubation period, it is reasonable to model the epidemic with another compartment, which exposes humans that are infected but not virus spreaders.

We performed a simulation process to estimate the parameters of the model to obtain the best fit to the reported data of the COVID-19 outbreak in Saudi Arabia. In the SEIR model, we assume long time, no vital dynamics, and a population $N$ size constant. In this model, individuals are classified into four groups or compartments according to their infectious status. We classified the infection types as either with or without symptoms. The compartments of the model are as follows:

(i) Susceptible ($S_n$): number of humans that are not infected by the virus but may be infected
(ii) Exposed ($E_h$): number of humans that are infected but do not spread the virus
(iii) Infected ($I_h$): number of infected humans with symptoms
(iv) Infected ($I_n$): number of infected humans without symptoms
(v) Recovered ($R_h$): number of humans that are recovered from infection

The SEIR model is

<table>
<thead>
<tr>
<th>Dates</th>
<th>New cases</th>
<th>Cumulative confirmed cases</th>
<th>Recovered cases</th>
<th>Daily death cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 June 2020</td>
<td>4919∗</td>
<td>141235</td>
<td>2122</td>
<td>39</td>
</tr>
<tr>
<td>17 June 2020</td>
<td>4757</td>
<td>145992</td>
<td>2253</td>
<td>48</td>
</tr>
<tr>
<td>18 June 2020</td>
<td>4301</td>
<td>150293</td>
<td>1849</td>
<td>45</td>
</tr>
<tr>
<td>19 June 2020</td>
<td>3941</td>
<td>154234</td>
<td>3153</td>
<td>46</td>
</tr>
<tr>
<td>20 June 2020</td>
<td>3379</td>
<td>157613</td>
<td>2213</td>
<td>37</td>
</tr>
<tr>
<td>21 June 2020</td>
<td>3393</td>
<td>161006</td>
<td>4045</td>
<td>40</td>
</tr>
<tr>
<td>22 June 2020</td>
<td>3139</td>
<td>164145</td>
<td>4710</td>
<td>39</td>
</tr>
<tr>
<td>23 June 2020</td>
<td>3123</td>
<td>167268</td>
<td>2912</td>
<td>41</td>
</tr>
<tr>
<td>24 June 2020</td>
<td>3372</td>
<td>170640</td>
<td>5085</td>
<td>41</td>
</tr>
<tr>
<td>25 June 2020</td>
<td>3938</td>
<td>174578</td>
<td>2589</td>
<td>46</td>
</tr>
<tr>
<td>26 June 2020</td>
<td>3927</td>
<td>178505</td>
<td>1657</td>
<td>37</td>
</tr>
<tr>
<td>27 June 2020</td>
<td>3989</td>
<td>182494</td>
<td>2627</td>
<td>40</td>
</tr>
<tr>
<td>28 June 2020</td>
<td>3943</td>
<td>186437</td>
<td>2363</td>
<td>48</td>
</tr>
<tr>
<td>29 June 2020</td>
<td>4387</td>
<td>190824</td>
<td>3648</td>
<td>50</td>
</tr>
<tr>
<td>30 June 2020</td>
<td>3402</td>
<td>194226</td>
<td>1994</td>
<td>49</td>
</tr>
</tbody>
</table>

Table 3: Number of COVID-19 cases in Saudi Arabia from 16th June 2020 to 30th June 2020 (∗ the highest number of infection cases).

<table>
<thead>
<tr>
<th>Dates</th>
<th>New cases</th>
<th>Cumulative confirmed cases</th>
<th>Recovered cases</th>
<th>Daily death cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Jan.</td>
<td>101</td>
<td>362981</td>
<td>182</td>
<td>9</td>
</tr>
<tr>
<td>2 Jan.</td>
<td>82</td>
<td>363063</td>
<td>180</td>
<td>7</td>
</tr>
<tr>
<td>3 Jan.</td>
<td>94</td>
<td>263157</td>
<td>166</td>
<td>10</td>
</tr>
<tr>
<td>4 Jan.</td>
<td>103</td>
<td>363260</td>
<td>146</td>
<td>9</td>
</tr>
<tr>
<td>5 Jan.</td>
<td>117</td>
<td>363377</td>
<td>144</td>
<td>7</td>
</tr>
<tr>
<td>6 Jan.</td>
<td>108</td>
<td>363485</td>
<td>138</td>
<td>6</td>
</tr>
<tr>
<td>7 Jan.</td>
<td>97</td>
<td>363582</td>
<td>171</td>
<td>4</td>
</tr>
<tr>
<td>8 Jan.</td>
<td>110</td>
<td>363692</td>
<td>174</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5: Number of COVID-19 cases in Saudi Arabia from 1st January 2021 to 8th January 2021.

Table 4: Monthly number of COVID-19 cases in Saudi Arabia from July to December 2020.

Table 5: Number of COVID-19 cases in Saudi Arabia from 1st January 2021 to 8th January 2021.
\[
\frac{d}{dt}S_h = \mu_h N_h - \left( \frac{I_h S_h \omega c \beta}{N_h} + \mu_h \right) S_h, \quad (13)
\]
\[
\frac{d}{dt}E_h = \frac{I_h S_h \omega c \beta}{N_h} - \frac{I_h S_h \omega c \beta}{N_h} - \alpha E_h, \quad (14)
\]
\[
\frac{d}{dt}I_h = \alpha E_h - \frac{I_h \beta}{N_h} + (\delta_i + \mu_i + \mu_h) I_h, \quad (15)
\]
\[
\frac{d}{dt}I_n = (1 - \delta_i) E_h - \frac{I_h \beta}{N_h} - \mu_n, \quad (16)
\]
\[
\frac{d}{dt}R_h = (\mu_i + \mu_a) I_h - \mu_h R_h, \quad (17)
\]

where \( \beta \) is the probability of spread of the virus per contact, \( c \) is the rate of contact, \( \omega \) is the transition rate of infection from people without symptoms, \( \alpha \) is the transition rate of the virus of exposed individuals to the infected environment, \( \delta_i \) is the transition rate of symptomatic infected individuals, \( \mu_i \) and \( \mu_a \) are the recovery rates of symptomatic and asymptomatic infected individuals, respectively, and \( \mu_h \) is the disease-induced death rate.

The compartmental model diagram representing the SEIR is based on the dynamic behaviors of the disease, epidemiological status of the individuals, and intervention measures (shown in Figure 3). We estimated the parameters of the model using the data obtained from confirmed cases of COVID-19 in Saudi Arabia.

In this model, the basic reproduction number (\( R_0 \)) is an important parameter for understanding the basic characteristics of an epidemic. It is fixed, which indicates that the average number of secondary people in a completely susceptible population is infected by a single infected person at the beginning of the epidemic. This indicates the transmission potential of infectious diseases. The value of \( R_0 \) is affected by several factors:

(i) Period of infectiousness

(ii) Probability of infection being transmitted during contact between an infected person (with or without symptoms) and a susceptible person

(iii) Rate of contact between infected and susceptible people

Compared with the basic reproduction number, the effective reproduction number (\( R_t \)) is used to measure...
secondary people in a completely susceptible population infected by a single infected person in which some control measures have been implemented. The total population size is considerably larger than the infection size. The effective reproduction number threshold is studied as an indication of the transmissibility of the virus, representing the average number. V_\text{the effective transmissibility of the virus, representing the average number of secondary people in a completely susceptible population infected by a single infected person in which some control measures have been implemented.}

The infection type included symptomatic and asymptomatic types. Infection with the virus is considered to be spreading in the population when \( R_t > 1 \). When \( R_t < 1 \), the infection of the virus decreases due to the escape of susceptible and exposed individuals.

2.3.2. Logistic Growth Models. The logistic growth model has been widely used to predict population growth rates with limited resources and space. The model was originally developed by Haberman [51] and was used to predict the 2015 Ebola epidemic [10, 52]. In this model, the dynamics of the epidemic are expressed as a cumulative number of cases in predicting the infection of COVID-19 in Saudi Arabia, provided that all the control measures have been implemented.

Using the classical logistic equation in the continuous case, the epidemic can be defined as:

\[
\frac{dC}{dt} = rC \left( 1 - \frac{C}{K} \right),
\]

(19)

where \( C \) is the cumulative infection cases at time \( t \), \( r \) is the increasing infection rate, and \( K \) is the maximum infection size.

To investigate the relationship between the number of cases per day and the total number of maximum infection sizes, we obtained the following solution from (19):

\[
C(t) = \frac{K}{1 + A e^{rt}},
\]

(20)

\[
A = \frac{K - C_0}{C_0}
\]

At the initial stage of the epidemic, the solution is represented by an exponential function:

\[
C(t) = C_0 e^{rt},
\]

(21)

and if the graph of this equation approximates the increase in the number of cases at the initial stage, we can determine the growth rate \( r \). We can neglect this logistic model equation if the data do not fit the exponential dependence.

Predicting the maximum possible number of infected individuals is the most important characteristic. This can be predicted only at the stage of the noticeable difference between the data and the exponential curve, provided that the number of infected people is not small. As health centers operate considering cases per day, it is necessary to consider the difference logistic equation. The difference equation, which is a discrete case of the logistic equation governing the epidemic, is defined as follows:

\[
C(n) - C(n - 1) = rC(n - 1) \left( 1 - \frac{C(n - 1)}{K} \right),
\]

(22)

where \( n \) is the discrete time interval in days and \( C_0 \) is the initial number of infections.

From equation (22), we can obtain a simple equation representing the relationship between the number of cases per day and the total number of cases (maximum infection size).

3. Results

3.1. Simulation Results of the SEIR Model. To perform the simulation model parameters, we used empirical data on COVID-19 in Saudi Arabia, which was obtained from the official website of the Ministry of Health [53]. We illustrate the simulated results of susceptible, exposed, confirmed, and recovery cases by applying the proposed SEIR model starting from May 1, 2020, until January 8, 2021. The predicted cases from the model were compared with the actual data. MATLAB and MAPLE were used to simulate the SEIR model. The epidemic simulation graphs are shown in Figure 4.

Figures 4–6 show the simulation of the spread of COVID-19 in Saudi Arabia using our modified SEIR epidemic model, considering symptomatic and asymptomatic infections. According to the figures, the infection due to the asymptotic transmission is slightly higher than that of the symptomatic transmission. However, our model predicted that the rate of infection gradually decreased with time (months). The simulation parameters shown in Table 6 and obtained from the real-time data (empirical data) from Saudi Arabia were used in our model. The fitting curve between the

---

**Figure 3:** Diagram of the SEIR model adopted for simulating the COVID-19 infection. The infection type included symptomatic and asymptomatic types.

**Figure 4:** Simulation of the spread of COVID-19 in Saudi Arabia using our modified SEIR epidemic model, considering symptomatic and asymptomatic infections.
solution of our model and the real data with a 95% confidence interval is shown in Figure 7(a). The solid line in dark yellow corresponds to the solution of the model, and the blue and red color lines correspond to the lower and upper 95% confidence intervals, respectively. Figure 7(b) shows the curve fitting between the model solution and real data.

We used MATLAB to simulate our model equations using the above parameters to predict future outbreaks of COVID-19 in Saudi Arabia. The variations in the parameters and corresponding model behaviors are shown in Figure 8. Figures 8 and 9 show the dynamics of the spread of COVID-19 infection to humans with the parameters $\mu_i$ and $\beta_a$ estimated from the actual number of cases in Saudi Arabia. From the simulated results, our model predicted that the COVID-19 infection in the environment will gradually decrease with time. According to the SEIR model, the
epidemic will reach its final phase with a minimum infection rate at the end of February 2021.

Using the model equations, we considered the direction field in various regions of the phase space, as shown in Figure 10.

From Figure 10(a), the number of infected people both with and without symptoms heads toward a stable equilibrium point depending on the initial conditions. Further, susceptible people will reduce in size and reach an equilibrium point. Figure 10(b) shows the phase portal diagram for the number of infected people with symptoms and susceptible people. We obtained negative eigenvalues, and thus, the equilibrium point at the Sh axis was stable. Hence, susceptible people reached a stable stage, whereas infected people were at a minimum (almost zero infection). Figure 10(c) shows the three-dimensional phase portal for susceptibility, infection, and recovery; the recovery of humans increased with no bounds for the change in parameters related to infection and susceptibility.

### 3.2. Simulation Results of the Logistic Model

MATLAB was used to simulate both continuous and discrete cases of the logistic model equation. From the simulation results, we compared the prediction of the cumulative infection cases with the real data. Figure 11 shows the comparison of the actual and predicted cumulative infection cases in Saudi Arabia based on the logistic growth equation.

The actual data of cumulative confirmed cases from May 1, 2020, were used to compare the cumulative infection case with the logistic growth model simulation and predict the future scenarios of the epidemic. The initial number of infected individuals on this date was $C_0 = 1362$.

We used the exact and discrete case of logistics growth models to predict the cumulative infection cases in Saudi Arabia. Figures 11(a)–11(c) show the predicted cumulative infection cases with infection rates of 0.015 for 360, 400, and 460 days, respectively. The simulation was repeated with an infection rate of 0.016 for 360, 400, and 460 days, as shown in Figures 11(b), 11(d), and 11(f), respectively. These models, for discrete and continuous cases, tend to underestimate the maximum infection size $K$ and thus could serve as lower bounds for future prediction. The logistic growth model provided accurate results for estimating the final size of the epidemic in Saudi Arabia.

### 4. Discussion

Our proposed SEIR model, considering both the infection types with and without symptoms, is suitable for the slow transmission of COVID-19 in Saudi Arabia. The parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Dimension</th>
<th>Method/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.095</td>
<td>Day$^{-1}$</td>
<td>Markov Chain Monte Carlo (MCMC) [54]</td>
</tr>
<tr>
<td>$c$</td>
<td>14.5</td>
<td>Day$^{-1}$</td>
<td>Data fitted [53, 54]</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.009</td>
<td>Day$^{-1}$</td>
<td>Data fitted</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1</td>
<td>Day$^{-1}$</td>
<td>MCMC</td>
</tr>
<tr>
<td>$\delta_i$</td>
<td>0.13</td>
<td>Day$^{-1}$</td>
<td>Data fitted [53]</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>0.96</td>
<td>Day$^{-1}$</td>
<td>Data fitted [53]</td>
</tr>
<tr>
<td>$\mu_a$</td>
<td>0.94</td>
<td>Day$^{-1}$</td>
<td>Data fitted [53]</td>
</tr>
<tr>
<td>$\mu_h$</td>
<td>$1.7826 \times 10^{-5}$</td>
<td>Day$^{-1}$</td>
<td>MCMC</td>
</tr>
<tr>
<td>$N$</td>
<td>34,218,169</td>
<td>Individual</td>
<td>[54]</td>
</tr>
<tr>
<td>Birth rate</td>
<td>1252</td>
<td>Individual/day</td>
<td>[54]</td>
</tr>
</tbody>
</table>

**Table 6: Estimated values of parameters.**
were simulated from the real data and found to be different for different scenarios. Moreover, the logistic growth model was used to predict the future dynamics of the outbreak in Saudi Arabia.

COVID-19 infection is an important global health issue and has infected a considerable portion of the world population. In this study, we investigated the biological and epidemiological tendencies in the prevalence and mortality due to the outbreak of COVID-19 infection. The infection of COVID-19 has expanded to over 220 countries and territories. It has infected 89,858,494 people and has caused 1,930,829 (2.14%) deaths from December 2019 to January 8, 2021.

The infection of COVID-19 was rapidly transmitted during June, July, and August. The spreading phase of COVID-19 has gradually decelerated due to various control measures, such as lockdown and travel limitations. COVID-19 is highly contagious because of its biological characteristics [55]. Infected people can transmit the disease before they show clinical symptoms, which is one of the main reasons for the rapid spread of the virus from Wuhan city, China, to the rest of the world and Saudi Arabia. However,
the Ministry of Health, Saudi Arabia, has taken various precautionary measures to control the spread of COVID-19. Some of these preventive measures are as follows: (1) providing a COVID-19 center helpline, number 937, and suspending the entry into Saudi Arabia for Umrah; (2) ceasing the issue of visas to individuals coming from countries already affected by COVID-19; (3) impeding the entry of citizens of Gulf countries into two holy cities, Makkah and Almadinah; (4) addressing more than 25 hospitals all over Saudi Arabia to COVID-19 cases; and (5) providing massive campaigns for awareness in public places and educational organizations [56]. All these preventive measures were important in limiting the spread of COVID-19 in Saudi Arabia. In addition, the number of deaths compared to other countries is very low, which might be due to the good care of infected patients in hospitals in Saudi Arabia.

Recently, a new mutated coronavirus variant was identified in the United Kingdom (UK). The health secretary of the UK has mentioned that a new variant of coronavirus may be driving infection in the southeast region. One of the most significant mutations of this virus is in the spike protein, which may result in a more infectious virus with rapid spreading [57].
Figure 10: Phase portal for the SEIR model at the equilibrium points: (a) $I_n$ and $I_h$ plane; (b) $I_n$ and $S_h$ plane; (c) $R_h$ and $S_h$ plane.

Figure 11: Continued.
Table 7: Outbreak prediction of cumulative confirmed cases.

<table>
<thead>
<tr>
<th>Model</th>
<th>360 days</th>
<th>400 days</th>
<th>460 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted by the logistic model (r = 0.015)</td>
<td>discrete: 278338 exact: 293544</td>
<td>discrete: 494863 exact: 523390</td>
<td>discrete: 1136294 exact: 1203734</td>
</tr>
<tr>
<td>Predicted by the logistic model (r = 0.016)</td>
<td>discrete: 392238 exact: 415936</td>
<td>discrete: 717755 exact: 762952</td>
<td>discrete: 1687537 exact: 1792251</td>
</tr>
<tr>
<td>Real data (cumulative)</td>
<td>363692</td>
<td>363692</td>
<td>363692</td>
</tr>
</tbody>
</table>

Table 8: Outbreak prediction of confirmed cases per day.

<table>
<thead>
<tr>
<th>Model</th>
<th>360 days</th>
<th>400 days</th>
<th>460 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted by the SEIR model (r = 0.015)</td>
<td>Infection ($I_h$): 10358</td>
<td>Infection ($I_h$): 7075</td>
<td>Infection ($I_h$): 4064</td>
</tr>
<tr>
<td></td>
<td>infection ($I_h$): 7526</td>
<td>infection ($I_h$): 3668</td>
<td>infection ($I_h$): 111</td>
</tr>
<tr>
<td>Predicted by the SEIR model (r = 0.016)</td>
<td>Infection ($I_h$): 10262</td>
<td>Infection ($I_h$): 7074</td>
<td>Infection ($I_h$): 4063</td>
</tr>
<tr>
<td></td>
<td>infection ($I_h$): 7414</td>
<td>infection ($I_h$): 3669</td>
<td>infection ($I_h$): 112</td>
</tr>
<tr>
<td>Real data (active cases)</td>
<td>2772</td>
<td>2772</td>
<td>2772</td>
</tr>
</tbody>
</table>
Transition rate $\alpha = 0.015$ with fixed $\mu_i = 0.96$ and $\mu_a = 0.94$.

Figure 12: Continued.
The three main characteristics that enable easier spreading of the new coronavirus variant SARS-CoV-2 are as follows:

(i) Rapidly replacing another version of the virus
(ii) Mutations that affect part of the virus
(iii) Some of these mutations that increase the ability of the virus to infect cells [58]

Over time, as more mutations occur, the existing vaccines may need to be adjusted accordingly. However, SARS-CoV-2 does not mutate quickly as a flu virus, and the available vaccines have proved to be effective in trials. There is no detection of a new strain of coronavirus in Saudi Arabia. The Saudi health representative also stated that the mutated coronavirus variant did not change the way the virus affects humans [59].

Considering the present scenario, by adopting precaution measures and travel limitations with other countries, we predicted the possible number of cumulative infection cases in Saudi Arabia. Table 7 shows the estimated predictions for future epidemic tendencies using the logistic growth rate model. A graphical representation of the epidemic prediction is shown in Figure 8. The predictions using our models were less significantly different from the real data. The prediction number with the solutions of the logistic growth model was more accurate and very close to the cumulative COVID-19 cases around 360 days was close to the real number of cumulative
cases. Furthermore, we estimated the possible cumulative infection cases for 400 and 460 days.

In Table 8, the prediction of the possible number of daily infection cases (active cases) in Saudi Arabia is shown using the SEIR model. Simulation of the outbreak prediction was performed for three phases: 360, 400, and 460 days. The prediction of the daily infection cases using the SEIR model considering the type of infection with and without symptoms was simulated, and its graphical representation is shown in Figure 12. From the simulation process, the SEIR model with infection types had a higher prediction compared with the real data on daily active cases. However, the estimation of the nature of infection reduces over time. The infection rate of the virus in asymptomatic individuals seems to be higher than that of asymptomatic individuals. It is suggested that asymptomatic people should be identified to contain the spread of the disease.

Our future research directions will focus on the extension of the proposed models with the availability of vaccination and are more suitable for the mutated SARS-CoV-2.

5. Conclusions

The aim of this study was to model and simulate the spread of COVID-19 in Saudi Arabia using compartmental epidemic models. We studied the dynamic nature of the COVID-19 outbreak in Saudi Arabia for the reported cases from March 2, 2020, to January 8, 2021, using mathematical modeling and simulation. We used both SEIR and logistic growth models to simulate the spread of COVID-19. The SEIR model has been extended by considering both symptomatic and asymptomatic infection types. According to the model, asymptomatic infected individuals have a higher chance of spreading the virus. In addition, we adopted a logistic growth model for both discrete and continuous time to understand the transmission dynamics of COVID-19 in Saudi Arabia. The models predicted the cumulative confirmed cases and daily active cases through a simulation process. However, the logistic growth model provided a good prediction close to the actual number of cases. The deterministic logistic model was more appropriate for describing the daily active cases. Both models have their own advantages in predicting the future number of cases with the SEIR model being higher than the logistics model.

In conclusion, our models predicted the future nature of the outbreak and assessed the effectiveness of Saudi Health Ministry control measures on infection. Furthermore, according to our model predictions, the future outbreak under strict control measures is expected to be minimal from February 2021.

Data Availability

The data are available on the Saudi Ministry of Health website.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors’ Contributions

All authors have equally contributed to preparing this manuscript.

Acknowledgments

The authors would like to thank the Scientific Research Ethics Committee, Jazan University, under the grant number (Cov19-38), Ministry of Education, Saudi Arabia, for the financial support of this research.

References


[52] G. Chowell, L. Simonson, C. Viboud, and Y. Kuang, "Is West Africa approaching a catastrophic phase or is the 2014 Ebola
epidemic slowing down? different models yield different answers for Liberia,” *PLOS Currents Outbreaks*, vol. 6, 2014.


