

Research Article

COVID-19 Pandemic Data Modeling in Pakistan Using Time-Series SIR

Muhammad Taimoor,¹ Sajid Ali ,¹ Ismail Shah ,¹ and Fred Roland Muwanika ²

¹Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan

²School of Statistics, Makerere University, P.O. Box 7062, Kampala, Uganda

Correspondence should be addressed to Sajid Ali; sajidali.qau@hotmail.com
and Fred Roland Muwanika; frmuwanika@yahoo.co.uk

Received 22 January 2022; Revised 29 March 2022; Accepted 13 June 2022; Published 28 June 2022

Academic Editor: Philip Broadbridge

Copyright © 2022 Muhammad Taimoor 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.

Pakistan is currently facing the fourth wave of the deadly coronavirus, which was first reported in Wuhan, China, in December 2019. This work utilizes the epidemiological models to analyze Pakistan's COVID-19 data. The basic susceptible, infected, and recovered (SIR) model is studied assuming Bayesian and time-series SIR (tSIR) approaches. Many studies have been conducted from different perspectives, but to the best of our knowledge, no study is available using the SIR models for Pakistan. The coronavirus incubation period has been set to 14 days across the globe; however, this study noticed that the assumption of 14 days is not suitable for Pakistan's data. Furthermore, on the basis of R_0 , we infer that COVID-19 is not a pandemic in Pakistan, as it was in other nations, such as the United States, India, Brazil, and Italy, among others. We attribute this to the best strategy adopted by the Government of Pakistan to minimize the burden of COVID-19 cases in Pakistani hospitals. It is also noticed that the posterior-based SIR (pSIR) model with uniform prior to R_0 and Poisson distribution (of log-likelihood) provides better results as compared to other distributions. From time-series SIR (tSIR), we observed that the value of the reporting rate (ρ) is less than 1 which means that cases are underreported.

1. Introduction

There are numerous types of coronaviruses, some of which cause severe threats to human life. Coronavirus gets its name from its shape—"corona" means "crown." The virus's top layers are coated with spike proteins, which act as a crown. The fatal COVID-19 disease started in 2019 which is a respiratory sickness caused by coronavirus. It first appeared in Wuhan, China, near the end of December 2019. The scientists initially treated it as pneumonia, but it spread at an alarming rate all over the world and became the first pandemic of the twenty-first century with a high reproductive rate. The International Committee on Taxonomy of Viruses named it severe acute respiratory syndrome coronavirus-2 [1]. However, on March 11, 2021, the World Health Organization (WHO) named it coronavirus disease-19 (COVID-19). According to WHO, there have been

532,201,219 confirmed cases, with 6,305,358 deaths worldwide, till June 6, 2022.

According to data from recent examinations, the severity of the disease differs epidemiologically by race, gender, and age [2]. It is almost two years, and no one knows how long it will stay around. Scientists are working hard to control it, but no particular solution has been found yet. The majority of patients infected with COVID-19 have mild-to-moderate symptoms and recover without any additional therapy. The researchers now believe that COVID-19 is transmitted mostly by droplets, and virus particles are dispersed into the atmosphere when an infectious agent coughs, laughs, talks, sings, sneezes, or breaths out. The most typical symptoms of COVID-19 include exhaustion, high temperature, sinus infection, discomfort, conjunctivitis, sore throat, diarrhea, headache, taste and odor loss, discoloration of the fingers or toes, lack of oxygen, chest pain, and speech and

movement impairments. The WHO advised the following precautions against COVID-19:

- (i) Hands should be cleaned often with soap or any alcohol-based liquid for at least 20 seconds
- (ii) Keep a safe distance (minimum 2 meters) away from somebody coughing or sneezing
- (iii) When physical separation is not possible, use a mask
- (iv) Avoid touching your eyes, nose, or mouth
- (v) Use tissue or bent elbow while coughing or sneezing
- (vi) If you are sick, stay at home and take care of yourself
- (vii) Seek medical help if you have a fever and cough or feel difficulty in breathing
- (viii) Masks can prevent others if someone has infection and vice versa

The first incidence of COVID-19 in Pakistan was recorded and reported by the Ministry of Health, Government of Pakistan, on February 26, 2020, in Karachi, Sindh. On the same day, Pakistan's Federal Ministry of Health confirmed another case in Islamabad. Within fifteen days, there were twenty confirmed cases out of 471 suspected cases. Sindh province had the greatest number, followed by Gilgit-Baltistan. All of the confirmed cases had traveled from Iran, Syria, or London, and the situation in Pakistan is still alarming. On February 12, 2020, Pakistan's Ministry of National Health Services, Regulation, and Coordination unveiled a plan titled "National Action Plan for Preparedness and Response to Corona Virus Disease Pakistan," with the goal of controlling virus spread and strengthening country and community emergency preparedness to ensure a timely, convenient, and prompt solution for future occurrences. The government introduced various measures over time to control the spread of coronavirus in the country. It includes containment measures, border control, quarantine houses, countrywide smart lockdowns, cordoning off areas, testing and contact tracing, establishing a field epidemiology laboratory program with the help of WHO, implementation of SOPs, initiation of an awareness campaign, production of ventilators, and other economic measures [3–5].

If we consider the current COVID-19 pandemic and try to develop a model that predicts how this mortal disease will behave in the upcoming months in Pakistan, it is plausible to think how many of those susceptible can be infected by a single infectious person. Suppose a single infection-carrying person can infect three others who are in contact with him. Then, the model will be like $I_t = I_0^t$, where I_t is the total number of infected people at time t and I_0^t is the initial number of infected people. This equation gives us the sudden increase in the disease and predicts that the epidemic will never end. Thus, if we use this single ordinary differential equation (ODE) to model the epidemic, it would not capture the real

situation and will show totally different and coarse graphics as compared to the real situation. That is why the SIR model is preferred. The SIR model is time-dependent and depends upon the initial values.

Weiss [6] used the SIR model in public health, whereas Finkenstädt and Grenfell [7] used time-series SIR (tSIR) modeling for children's disease measles. Grenfell et al. [8] studied the dynamics of measles using the tSIR model with the aim of providing a suitable empirical and theoretical test to address the importance of noise versus nonlinearity and temporal predictability of population wealth. Katris [9] used the tSIR, autoregressive integrated moving average (ARIMA), feedforward artificial neural network, and adaptive regression splines for forecasting the outbreak of COVID-19 in Greece. They compared different countries using R_0 and the mean absolute percentage error (MAPE). Pasquali et al. [10] introduced a beta distribution for the observation error in a SIR model to model the underdetection and removed individuals in the compartmental model for COVID-19.

Chen et al. [11] conducted a study on COVID-19 by assuming two scenarios, i.e., the time-wise prediction, time-dependent parameters (β and γ) and the undetectable number of infected individuals, to get a more accurate prediction using the time-dependent SIR (tSIR) model. Deo et al. [12] studied the transmission rate of COVID-19 in India using the tSIR model. They estimated that the total infections crossed 9 million with 1 million critical cases. They also estimated R_0 for smart and full lockdowns in different phases. Postnikov [13] conducted a study on COVID-19 using the simple SIR model for parameter estimation and future predictions. Metcalf et al. [14] conducted a study related to rubella in Mexico to capture seasonality, stochasticity, and region-wise variation. Lavielle et al. [15] extended the SIR model to model the COVID-19 data taken from Johns Hopkins University on the basis of the daily confirmed, active, death, recovered, and cumulative number of cases for each compartment. They applied the extended model to several countries like Switzerland, Italy, and the US.

Bjornstad et al. [16] estimated the transmission rates of measles for England and Wales using the tSIR model. Deo and Grover [17] noticed that the unreported pathogens are more threatening than reported and quarantined ones. Their proposed model was susceptible-infected (quarantined/free)-removed-deceased (SI(Q/F)RD). The estimated values of R_0 of undetected pathogens for California and Florida were 1.464 and 1.612, respectively. On the other hand, R_0 for reported pathogens were 0.497 and 0.359 for respective states at the time of the study. Fang et al. [18] conducted a study on the contagion moral force of COVID-19 to imitate the dispersion of COVID-19 using the susceptible-exposed-infected-recovered (SEIR) model. Waris et al. [3] assessed COVID-19 in Pakistan using different factors like hospital capacity, isolation, quarantine, and vaccination facilities.

Bruignano et al. [19] studied the multiregional extension of the SIR model for the COVID-19 outbreak in Italy to capture the effect of misdiagnosed infected and recovered pathogens. They used the susceptible, diagnosed-infected, undiagnosed-infected, diagnosed-recovered, and undiagnosed-recovered (SI_2R_2) model. Ferrari et al. [20] conducted a study about the seasonal incidence change

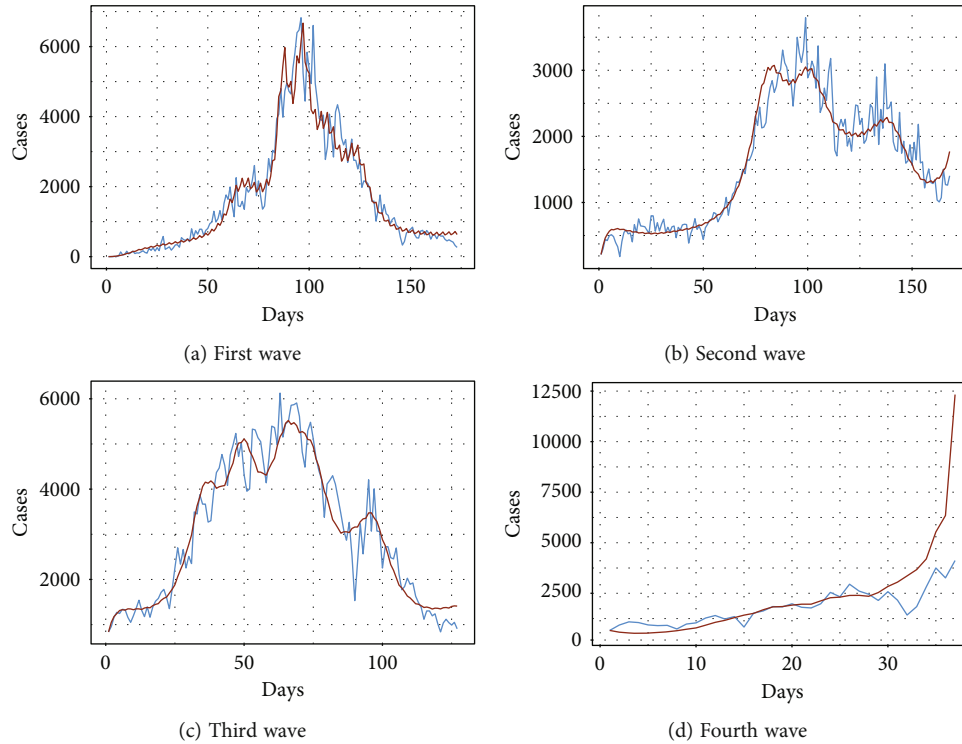


FIGURE 1: Wave-wise fitted tSIR model.

of measles in sub-Saharan Africa. They estimated seasonal/yearly variation by applying the tSIR model to 17-year-long time-series (monthly) data of Niamey (1986-2002). Althaus [21] estimated the basic reproductive number of the Ebola virus in West Africa. Li et al. [22] estimated the parameters of the stochastic SIR model using the maximum likelihood and Bayesian approaches with reference to media coverage of COVID-19.

There are different statistical and mathematical (deterministic) models that exist for epidemiological modeling. The existing models can also be used to predict the basic reproductive number of any pandemic like COVID-19. The SIR model, which is also known as the compartmental model, is used to predict the basic reproductive ratio. The SIR is the simplest form of the compartmental models, and all other models are the derivatives of this basic model. The SIR model has three different compartments, susceptible, infected, and recovered. The basic assumption of the traditional SIR model is that the infected (I) and susceptible (S) populations are mixed uniformly and that the overall population (N) remains constant throughout time [23]. Many studies, including time-series models for forecasting and mathematical models, are available for modeling the COVID-19 epidemic and for its prediction. Yousaf et al. [24] used the ARIMA model for forecasting COVID-19 in Pakistan which was the first study to make short-term forecasting about COVID-19 cumulative confirmed cases in Pakistan. The number of confirmed cases was increasing at a higher rate as compared to the number of recoveries. Contrary to their forecasted cases, the reality was totally different. Thus, we prefer to use compartmental modeling for the prediction of the basic reproductive number and fore-

casting purpose which is somehow performing better than statistical or time-series models used in epidemiological studies. Thus, the primary objective of the study is to predict the reproductive number and forecast COVID-19 cases in Pakistan using the SIR model. Further objectives of the study are to estimate the parameters (contact rate (β), recovery rate ($1/\gamma$), force of infection (λ), and reproductive number (R_0)) of the model using Pakistan's COVID-19 data. In addition, the study investigates whether cases are underreported or overreported and decides whether it is a pandemic or not in Pakistan.

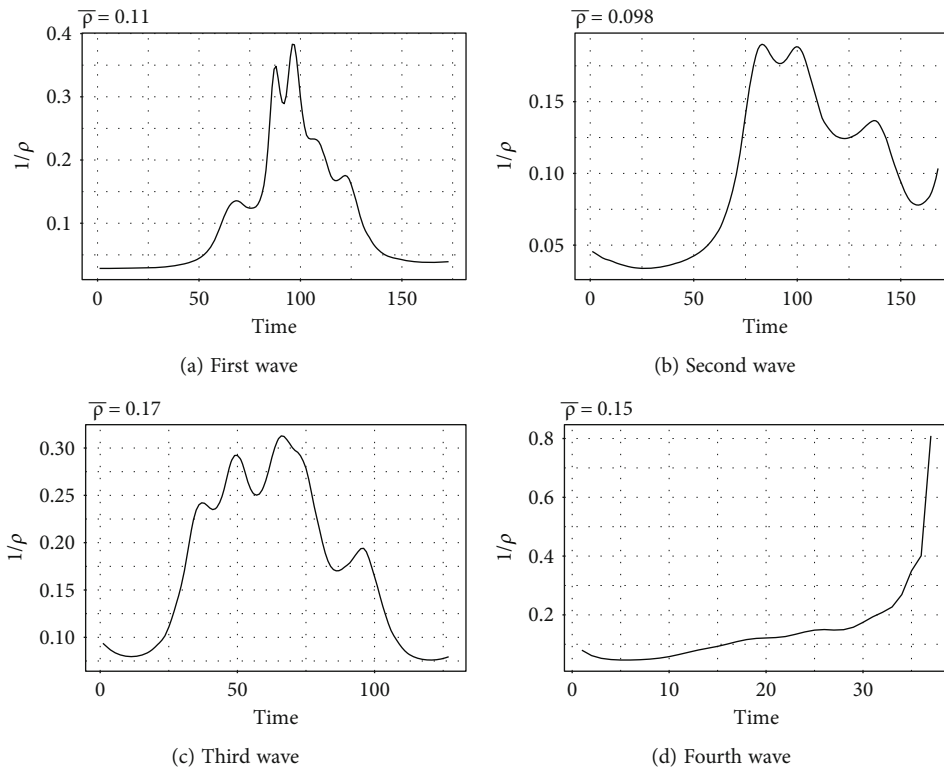
The rest of the study is categorized as follows. Section 2 presents the simple SIR, posterior-based SIR, and tSIR models. Analyses using the tSIR, SIR, and pSIR are discussed in Sections 3–5. Section 6 presents some concluding remarks and recommendations.

2. The SIR Model

The SIR model [1], introduced by Kermack and McKendrick in 1927, is a simple epidemiological model. This primary and simplest deterministic model is used for modeling the trend of epidemics (infectious diseases) and for obtaining its future predictions [25]. The population used in this model comprises three compartments which are as follows: (i) susceptible (S): the population not yet infected but can be infected by contact with an infected individual; (ii) infected (I): the number of persons who have been infected by the disease and can now transmit it; and (iii) recovered (R): the number of individuals who got infected but now have been shifted from the infected to the recovered compartment or are dead. Hence, $S + I + R = N$. The special

TABLE 1: Estimates of the tSIR model using data (500 days) and wave-wise data.

Parameters	Data	First wave	Second wave	Third wave	Fourth wave
$\bar{\beta}$	1.80E-06	2.53E-06	2.94E-05	1.93E-05	3.39E-06
$\bar{\rho}$	0.1290	0.1070	0.0980	0.1740	0.1460
\bar{S}	2.29E+06	2.24E+06	2.27E+06	2.29E+06	2.27E+06
S_0	7.79E-04	8.61E-03	9.97E-03	1.00E-02	1.00E-02
I_0	6.53E-08	1.58E-07	2.10E-05	3.96E-05	3.68E-05
α	0.8050	0.8200	0.5700	0.6100	0.7900

FIGURE 2: Wave-wise behavior of parameter ρ estimated by the tSIR model.

feature of SIR is that it is used to find the basic reproductive number of a pandemic, denoted by R_0 . R_0 basically tells us how many susceptible people can be infected (on average) by a single pathogen. If it is less than 1, we conclude that the pandemic will end very soon, and if it is greater than 1, then it is a pandemic and will take time to end. If R_0 is greater than 3, the whole population will be infected [26].

The mathematical form of the SIR model is

$$\begin{aligned}
 \frac{dS}{dt} &= -\beta S(t)I(t), \\
 \frac{dI}{dt} &= \beta S(t)I(t) - \gamma I(t), \\
 \frac{dR}{dt} &= \gamma I(t),
 \end{aligned} \tag{1}$$

where $dS/dt + dI/dt + dR/dt = 0$. The basic reproductive number R_0 is computed by $R_0 = (\beta/\gamma)N$.

2.1. Bayesian Approach to the SIR Model. In the Bayesian approach, the SIR model is slightly different from the simple SIR model as the Bayesian approach treats R_0 as a random variable, and one can assume negative binomial, Poisson, normal, or log-normal distribution as the prior distribution. A slight modification of the SIR model is given below.

$$\begin{aligned}
 P_{\text{inf}} \frac{dS}{dt} &= -SI \frac{R_0}{N}, \\
 P_{\text{inf}} \frac{dI}{dt} &= \left(\frac{R_0}{N} S - 1 \right) I, \\
 P_{\text{inf}} \frac{dR}{dt} &= I.
 \end{aligned} \tag{2}$$

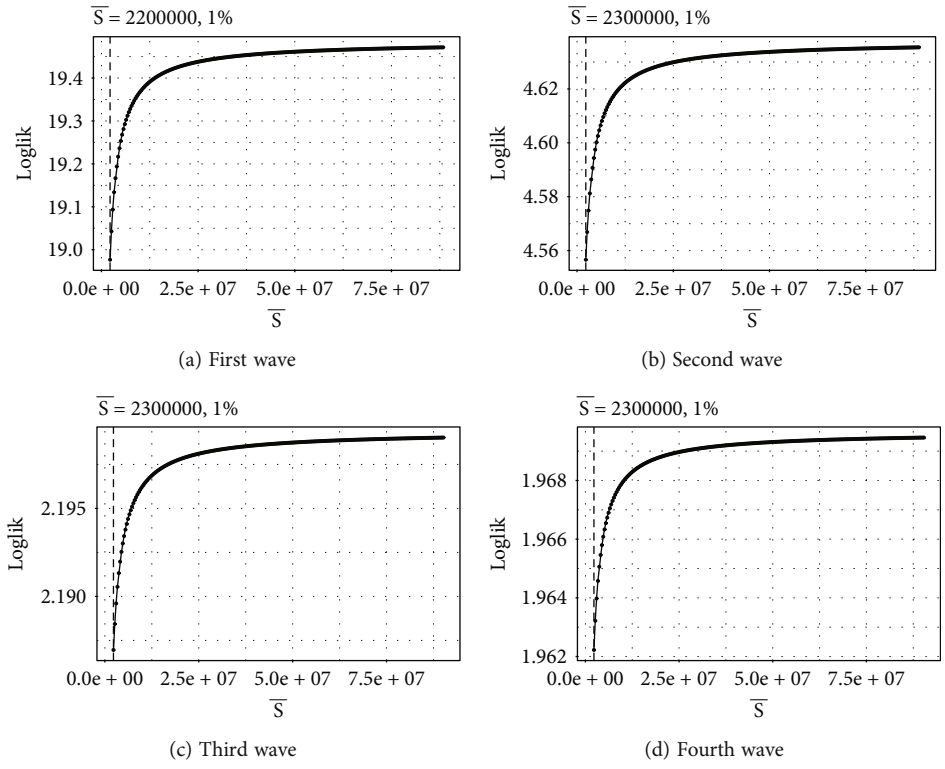


FIGURE 3: Wave-wise upcoming graphical behavior of the total number of susceptible people.

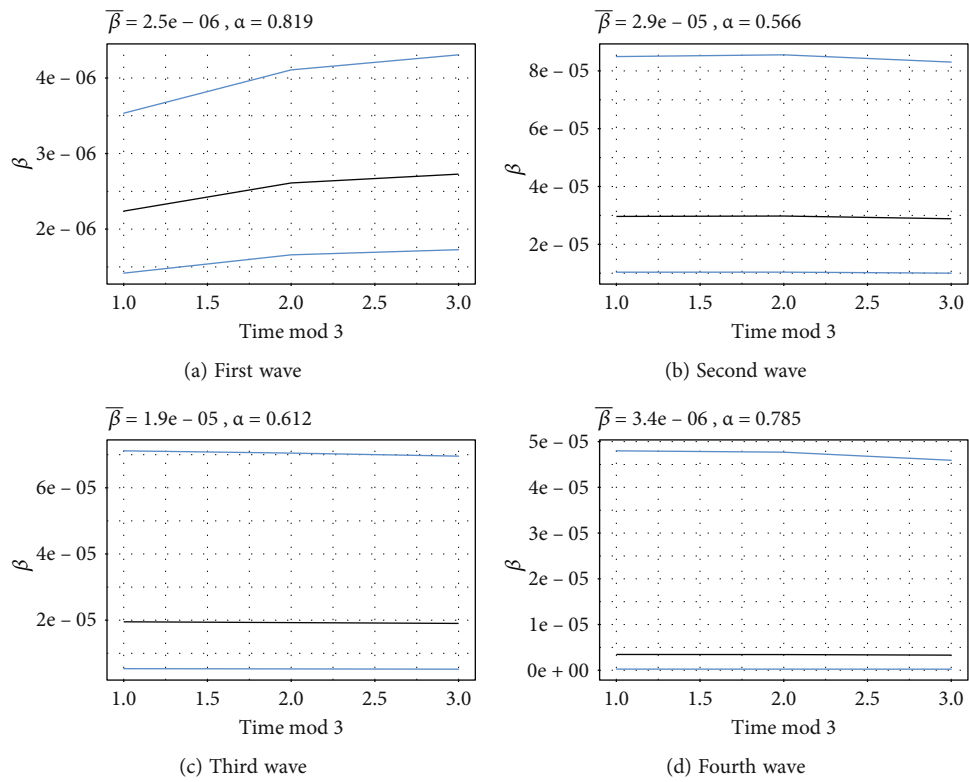


FIGURE 4: Wave-wise behavior of the contact rate (β).

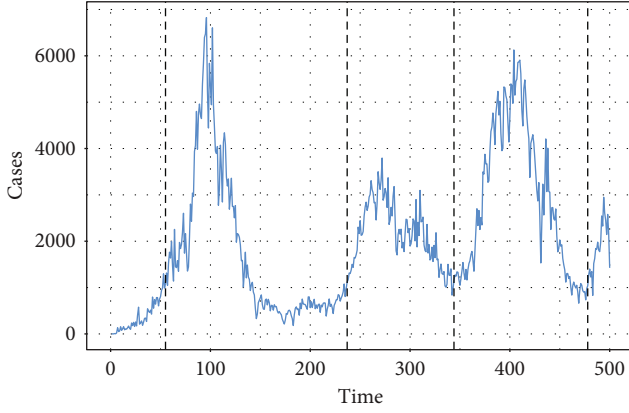


FIGURE 5: Outbreak plot of the data (500 days).

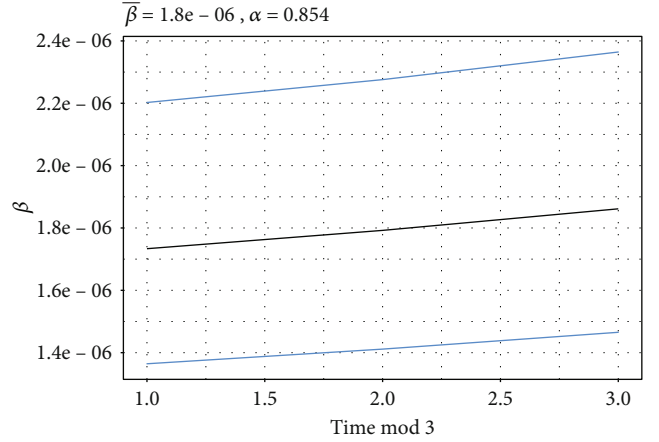
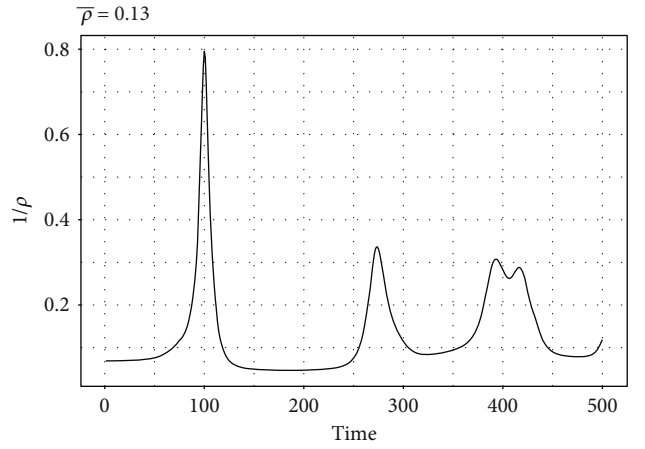
It is worth mentioning that the interpretation of the model is the same as that of the simple SIR model. This model has two parameters, R_0 and infectious period (P_{inf}). In this study, we use uniform distribution as a prior distribution of R_0 :

$$f(R_0) = \begin{cases} \frac{1}{b-a}, & \text{for } R_0 \in [a, b], \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

2.2. Time-Series SIR Model (tSIR). A time-series susceptible-infected-recovered (tSIR) model proposed by [7] is employed to study the dynamical behavior of the COVID-19 data. To this end, the data comprise two discrete variables, i.e., “infected cases” and “susceptible.” The SIR model can be expressed by the following three ODEs:

$$\begin{aligned} \frac{dS}{dt} &= \xi P(t) - \frac{\lambda S(t)}{N} - \mu(t)S(t), \\ \frac{dI}{dt} &= \frac{\lambda S(t)}{N} - \gamma I(t) - \mu(t)I(t), \\ \frac{dR}{dt} &= \gamma I(t) - \mu(t)R(t), \end{aligned} \quad (4)$$

where $\xi \in (0, 1)$, $\xi P(t)$ is the fraction of respiratory patients with SARS-2 infection, and ξ might be approximated as a constant near 1 (e.g., 0.9), but it will depend on the time in general. λ is the force of infection which can be expressed as $I\beta/N$, μ is the total number of deaths at time t , and γ is the recovery parameter. In the cyclical occurrences of the epidemic, the transmission rate β would vary with time. The traditional models give us estimates of each parameter against seasonal data. As we know, the tSIR model is tractable in the situations where seasonality occurs, but it also has some flaws. First, it takes only one main variable, “reported cases.” Second, there might be many underreported and overreported cases. Therefore, in this study, we use the “people having respiratory problems” to fulfill the requirements of the tSIR model. It is noticed from the literature that on a daily basis, almost 13,000 persons suffer from respiratory issues (https://www.aku.edu/news/Pages/News_Details.aspx?nid=NEWS-000849). Thus, to capture under- or overreporting, the tSIR model is modified as

FIGURE 6: Behavior of the contact rate (β) for the first five hundred days of COVID-19.FIGURE 7: Behavior of the reporting parameter (ρ).

$$S_{t+1} = P_{t+1} - S_t - I_{t+1}, \quad (5)$$

$$\mathbb{E}[I_{t+1}] = \beta_{t+1} S_t I_t^\alpha, \quad (6)$$

where P_{t+1} and I_{t+1} are the one-day-ahead forecasted number of people having respiratory problems and reported cases, respectively. Similarly, β_{t+1} is the one-day-ahead forecasted contact rate given by the tSIR model. Under the tSIR framework, the R function “runtSIR” first fits a simple regression model between the cumulative reported cases and the cumulative number of people having respiratory problems. As there should be a linear relationship between reported cases and people with respiratory problems, from the slope of the fitted regression line, we can conclude that the cases are either underreported or overreported. We name this slope as ρ_r , and it tells us whether the cases are under- or overreported. If the value of ρ is near 1, one may assume that cases are almost entirely reported, and if it is less than 1, one can conclude that cases are underreported. However, if its value is more than one, one can assume that instances are being reported excessively. The “runtSIR” function also gives us S_t

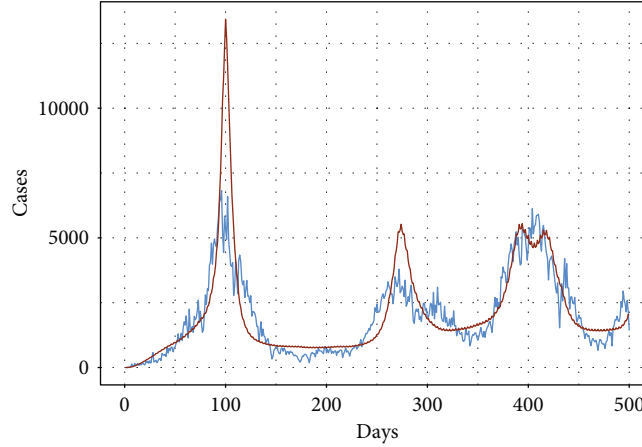


FIGURE 8: Graph of the fitted tSIR model to the data (500 days).

TABLE 2: MAPE and MAE using the first 500 days as training data and 10 days as testing data using different distributions.

Distributions	Link	MAPE(t)	MAE(t)	$\bar{\rho}(t)$	MAPE(f)	MAE(f)
Quasi-Poisson	Log	38.09	619.71	0.13	18.5862	650.3652
Poisson	Log	38.16	621.04	0.13	18.7826	651.3254
Gaussian	Log	36.58	617.39	0.13	17.9658	643.2589
Gaussian	Identity	36.95	620.27	0.13	17.4062	640.2154

and β_t , susceptible dynamics with respect to time t and contact rate of the single pathogen over time t , respectively. After taking the expectation of equations (5) and (6), one can get the following log-linear equation:

$$\log I_{t+1} = \log \beta_{t+1} + \log (Z_t + \bar{S}) + \alpha \log I_t, \quad (7)$$

where Z_t is the residual of the fitted regression model and α is the estimated homogeneity parameter by using the generalized linear model (GLM). \bar{S} is the mean number of susceptible people of the overall series, and α is a parameter which describes the intensity of the epidemic. Furthermore, the “runTSIR” fits the above log-linear relationship and resimulates the SIR model (forward and backward) by using the estimated parameters.

3. Analysis Using the tSIR Model

Figures 1(a)–1(d) present the behavior of the wave-wise fitted tSIR model for the first four waves of COVID-19 in Pakistan using the optimum values of parameters (listed in Table 1). In particular, the blue curve shows the actual number of daily reported cases while the gray curve presents the behavior of the cases using the tSIR model. On the basis of prior lag daily cases, it is noticed that the first wave is better fitted by the tSIR. However, in the second wave, there is a small amount of uncertainty between the fitted number and the observed number of instances. There is a variation in the third wave, but it appears to be a better match overall. However, when we look at the fourth wave, a significant

increase in the fitted curve is noticed, which indicates that the disease may pose a serious threat in the upcoming days.

Figure 2 shows the wave-wise behavior of $\bar{\rho}$ of COVID-19 data for Pakistan.

Figure 3 reports the daily total number of susceptible people estimated by the tSIR model for four waves of COVID-19, where \bar{S} is the mean number of susceptible people. To depict this figure, initially, 1% of the population is taken as the susceptible population.

Figure 4 presents the behavior of the contact rate (β) along with its respective estimated (by the tSIR model) intervals for four waves of COVID-19, respectively. Furthermore, it represents α , which is the disease intensity parameter.

3.1. Five-Hundred-Day Analysis. Figure 5 depicts the daily number of cases for the first five hundred days of COVID-19. One can observe that there are four waves, and it can be observed that the highest number of cases is observed for the first wave. It is worth mentioning that to depict the figure, we set 1000 as the outbreak threshold.

Figure 6 depicts the evolution of β for the first 500 days of COVID-19 data in Pakistan, as well as its estimated intervals. It also reflects the severity of the disease (α). The contact rate appears to be growing with time, which may lead to an increase in the daily number of cases.

The behavior of the estimated reporting rate (ρ) using the first 500 days of COVID-19 data for Pakistan is shown in Figure 7. One can observe that in the case of the first wave, there are a less number of cases which are actually positive and not reported as compared to all other waves of COVID-19.

TABLE 3: MAPE and MAE of the first wave using different distributions.

Distributions	Link	MAPE	MAE	$\bar{\rho}$
Quasi-Poisson	Log	65.92	300.89	0.11
Poisson	Log	65.31	300.58	0.11
Gaussian	Log	79.39	305.72	0.11
Gaussian	Identity	30.32	288.94	0.11

TABLE 4: MAPE and MAE of the second wave using different distributions.

Distributions	Link	MAPE	MAE	$\bar{\rho}$
Quasi-Poisson	Log	14.73	190.36	0.098
Poisson	Log	14.74	190.26	0.098
Gaussian	Log	14.81	190.67	0.098
Gaussian	Identity	14.25	188.79	0.098

TABLE 5: MAPE and MAE of the third wave using different distributions.

Distributions	Link	MAPE	MAE	$\bar{\rho}$
Quasi-Poisson	Log	13.04	335.08	0.17
Poisson	Log	13.05	335.51	0.17
Gaussian	Log	13.05	335.44	0.17
Gaussian	Identity	12.85	334.22	0.17

TABLE 6: MAPE and MAE of the fourth wave using different distributions.

Distributions	Link	MAPE	MAE	$\bar{\rho}$
Quasi-Poisson	Log	33.13	730.06	0.15
Poisson	Log	33.02	727.96	0.15
Gaussian	Log	33.43	733.96	0.15
Gaussian	Identity	33.01	704.81	0.15

Figure 8 shows the fitted tSIR model to the data (500 days). The blue curve represents the observed number of daily instances, whereas the gray curve represents the estimated curve using the tSIR model. Furthermore, the predicted number of cases in the middle of the first wave of COVID-19 is extremely high when compared to the observed cases. Similarly, the second-wave model gives an indication of more instances than the observed cases. However, the predicted values after 350 days seem somewhat better than the preceding waves.

Table 2 contains the summary of the mean absolute percentage error ($MAPE(t)$), mean absolute error ($MAE(t)$), and $\bar{\rho}$ for training data ($\bar{\rho}(t)$) by the tSIR using different distributions for the first 500 days of training data. The $MAPE(f)$ and $MAE(f)$ for the 10-day projected data are also shown in Table 2. As it can be seen that for the Gaussian distribution with the link “log” in both the training and forecasted (tested) data, the tSIR model performs better. Similarly, Tables 3–6 exhibit the MAPE, MAE, and $\bar{\rho}$ for four waves, respectively. Finally, we conclude that the Gaussian distribution with the link “identity” is appropriate.

Table 1 lists the estimated parameters for 500 days for four waves by the tSIR model.

4. Analysis Using the SIR Model

Figure 9 presents the four-wave analysis with the SIR model using the 14-day infectious period and considering the total population of Pakistan as the susceptible population. Contrary to the tSIR model, each wave has a distinct number of starting infected people with rapid exponential growth, indicating that COVID-19 is an epidemic in Pakistan and will infect a large population. This also indicates that the traditional SIR model should not be used for the data. Figure 10 presents the fitted SIR model for four waves with decreased susceptible population violating the assumption of an infectious period for 14 days, as stated by WHO. The fitted curves are produced using the optimal values of parameters β and γ , which result in the smallest error sum of squares (ESS). If we assume a 14-day infectious period, the daily number of cases increases dramatically and does not decrease as shown in Figure 9. Furthermore, the contact rate is too low in both

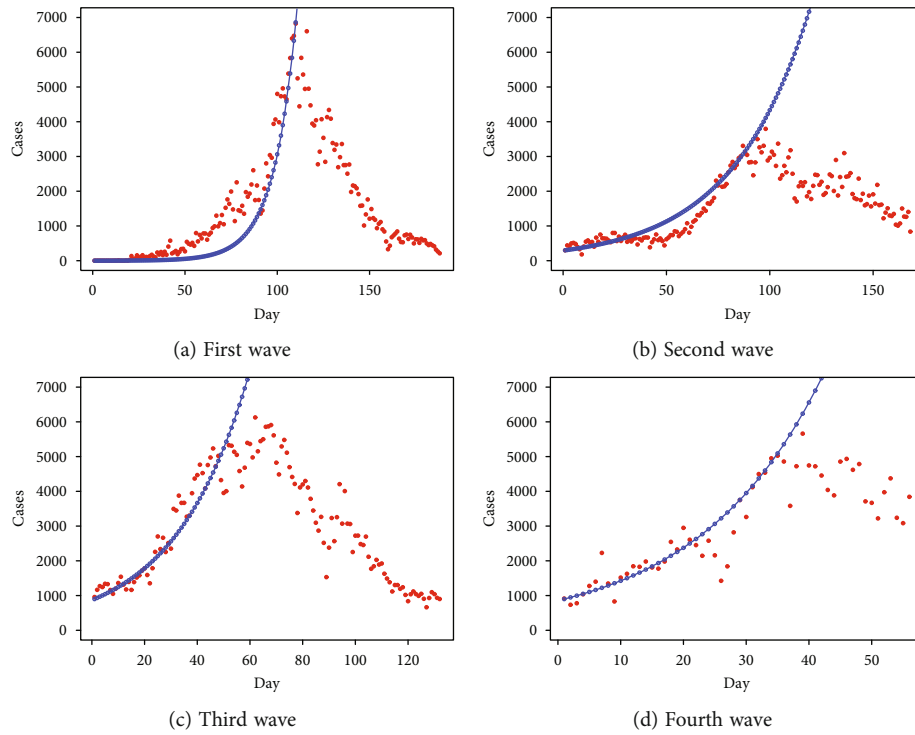


FIGURE 9: Wave-wise fitted SIR model with the 14-day infectious period.

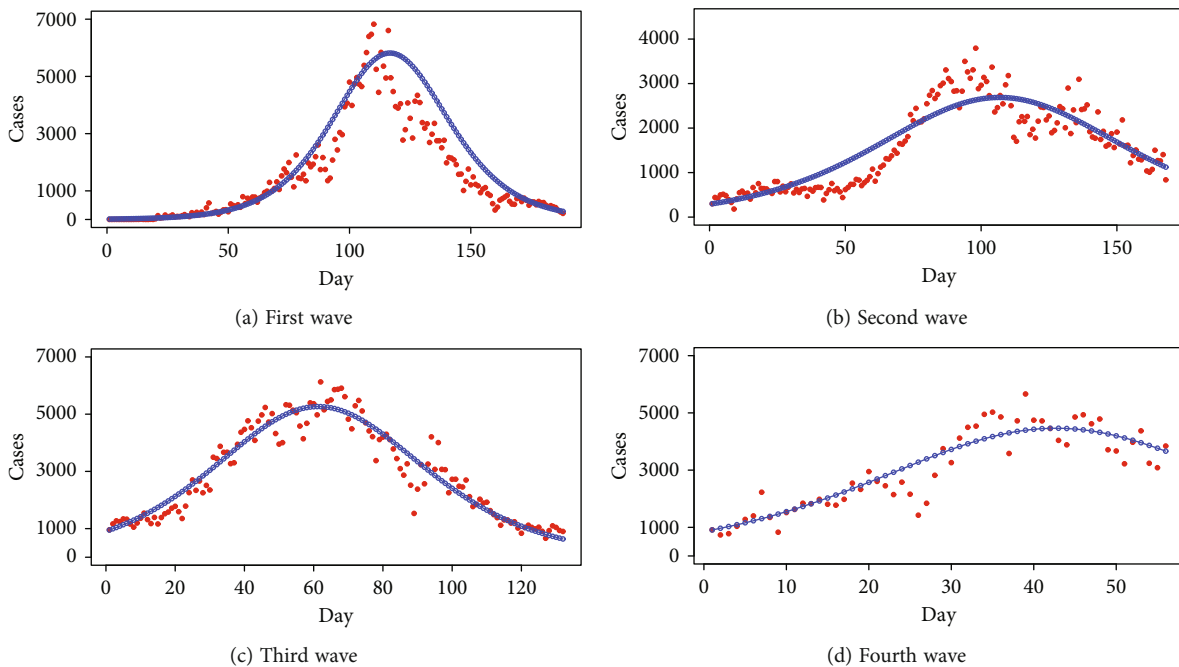


FIGURE 10: Wave-wise fitted SIR model.

situations (with optimal settings and 14-day infectious duration). Thus, we may infer that though disease exists in Pakistan, it is not a pandemic like it is in other nations. Figure 11 shows the fitted SIR model to wave-wise data with errors.

Figure 12 depicts the wave-wise optimum values of the contact rate (β) with the minimum residual sum of squares

(RSS). It can be seen that the first wave has a higher rate of infectious pathogen interaction than the others. However, this is due to the fact that the government did not take any strict steps against COVID-19 in the beginning. Due to smart and complete lockdowns, pathogen contact rates are minimal in other wave cases. Similarly, Figure 13 depicts

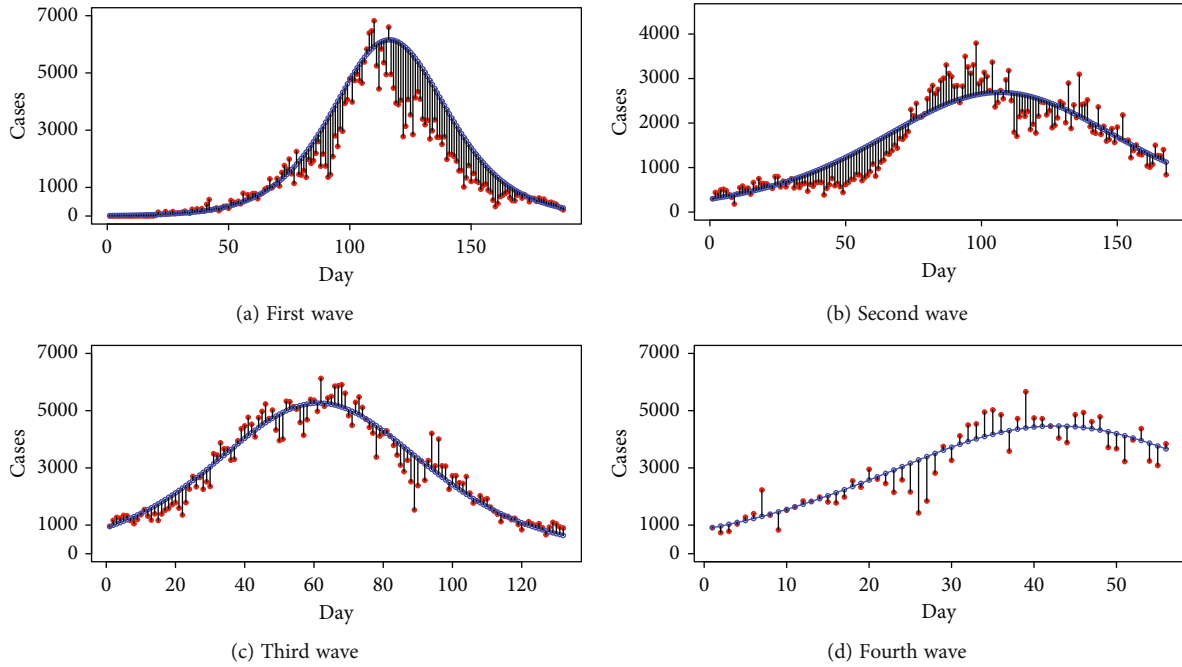


FIGURE 11: Wave-wise fitted SIR model with errors.

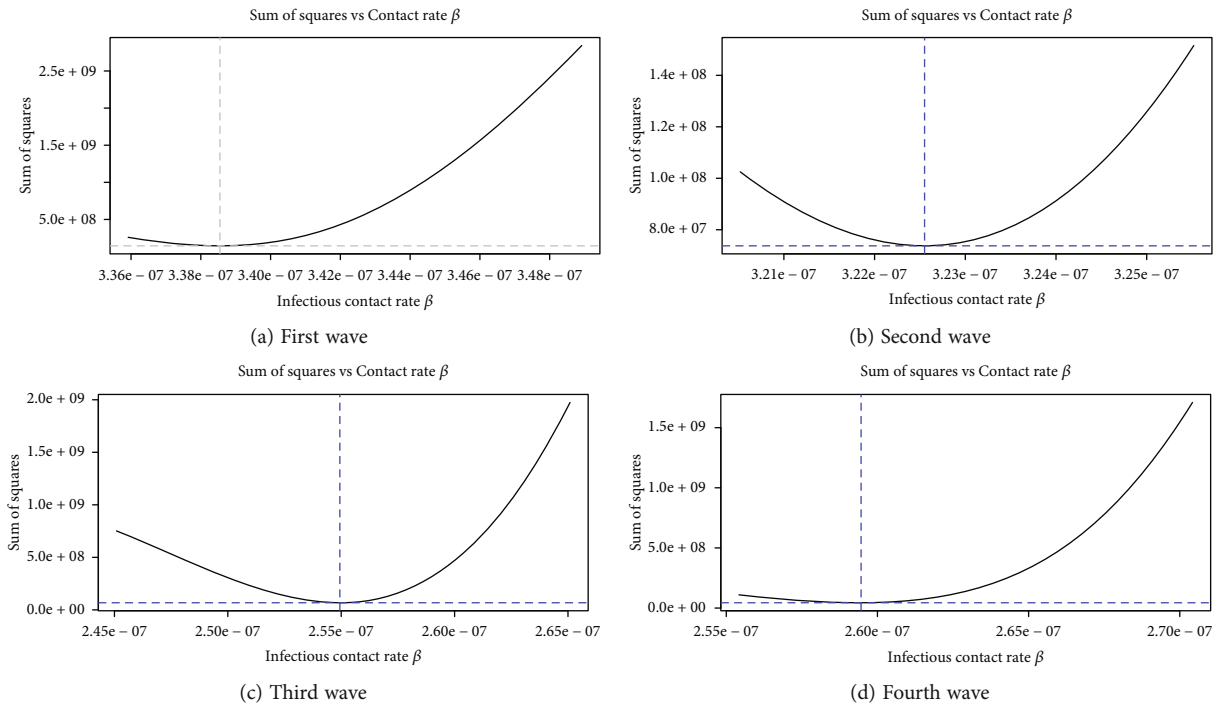


FIGURE 12: Optimum values of the contact rate (β) for four waves.

the optimum values of the recovery rate (γ) for four waves, and it is noticed that the recovery rate (γ) is greater in the first wave and steadily decreases in subsequent waves. This happened because individuals used to adhere to the SOPs and scientists also advised anti-

bodies to combat the virus. It is also noticed from the figures reported in the supplementary text that the intensity of illness rises as the pathogen contact rate increased. Also, the ESS increased due to the rise of β and γ .

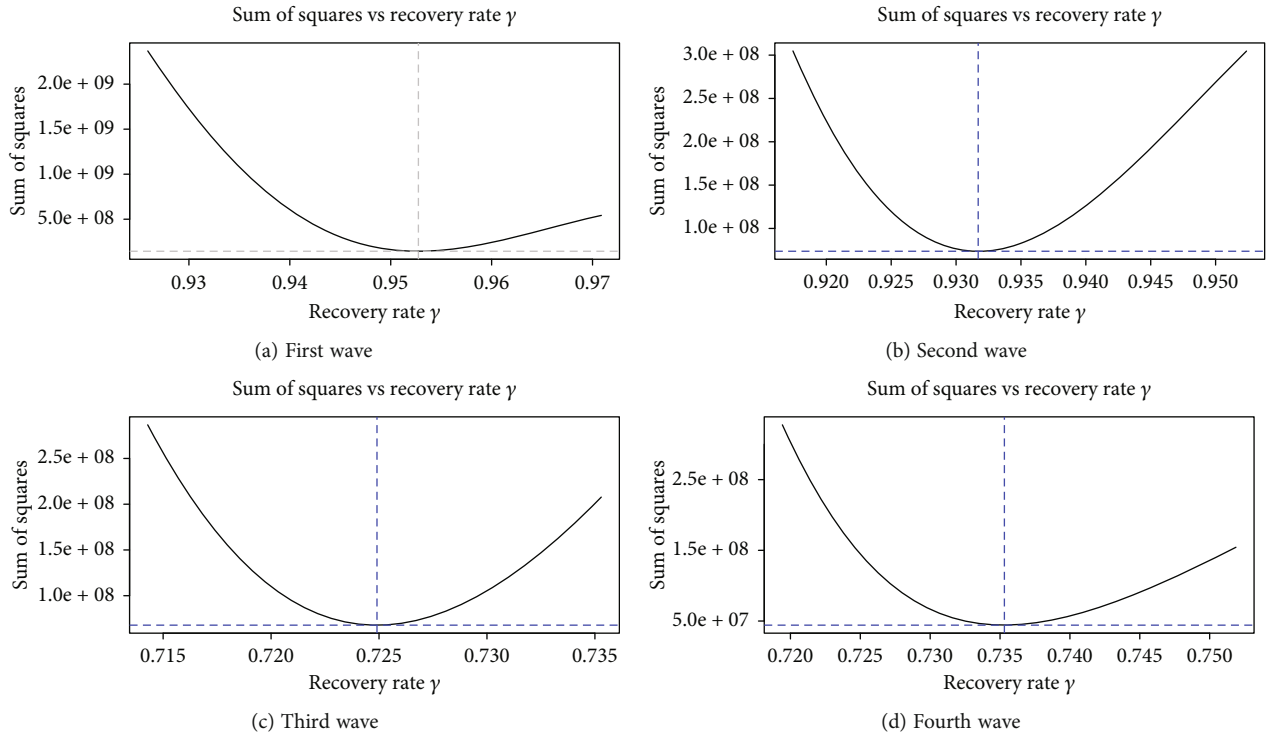


FIGURE 13: Optimum values of the recovery rate (γ) for four waves.

TABLE 7: Wave-wise estimated parameters of the SIR model.

Parameters	First wave	Second wave	Third wave	Fourth wave
β	3.448483E-7	2.65091E-7	2.8442E-7	4.26428E-7
γ	0.9709	0.7637	0.8079	1.2175
RSS	116,470,946	27,059,876	29,087,890	19,050,557
R_0	1.0656	1.0413	1.0561	1.0507

Table 7 lists the parameters estimated using the SIR model, and one can observe that R_0 in all waves of COVID-19 in Pakistan is about 1.05, which is a strong indication that COVID-19 is not endemic in Pakistan.

5. Analysis Using the Bayesian Approach to the SIR Model

Figure 14(a) depicts the estimated number of susceptible, infected, and recovered people by using the pSIR model for the first-wave data assuming uniform distribution as the prior distribution of R_0 and Poisson distribution for the log-likelihood function. It can be seen that this model behaves very similarly to the traditional SIR model. Here, we used the optimal parameter (estimated from the SIR model) values with the lowest RSS. Furthermore, the overall number of susceptible people is gradually reducing, near 2,600,000 after 188 days. Similarly, the recovery rate is increased with time, approaching 400,000 after 188 days. Furthermore, a large number of everyday instances follow a basic SIR model. Figure 14(b) shows the estimated number

of susceptible, infected, and recovered people for second-wave data using the pSIR model. After 150 days of second-wave data, the overall number of susceptible people is steadily reducing, approaching 2,800,000. Similarly, after 150 days, recoveries exceed 250,000. Furthermore, the 113th day of the second-wave data had the highest number of daily positive cases. Similarly, Figures 14(c) and 14(d) present the estimated behavior of susceptible, infected, and recovered people for the third and fourth waves of COVID-19, respectively.

Figure 15 depicts the wave-wise anticipated number of infected people, and by comparing this figure with Figure 10, one can see that there is no discernible difference between the SIR and pSIR models, except that SIR is a simple mathematical model, whereas the pSIR is a Bayesian SIR model.

Table 8 compares uniform, normal, and Poisson distributions as the prior distribution of R_0 while taking different distributions of log-likelihood on the basis of RSS for each wave of COVID-19 in Pakistan. One can see that in using uniform distribution as the prior distribution of R_0 and

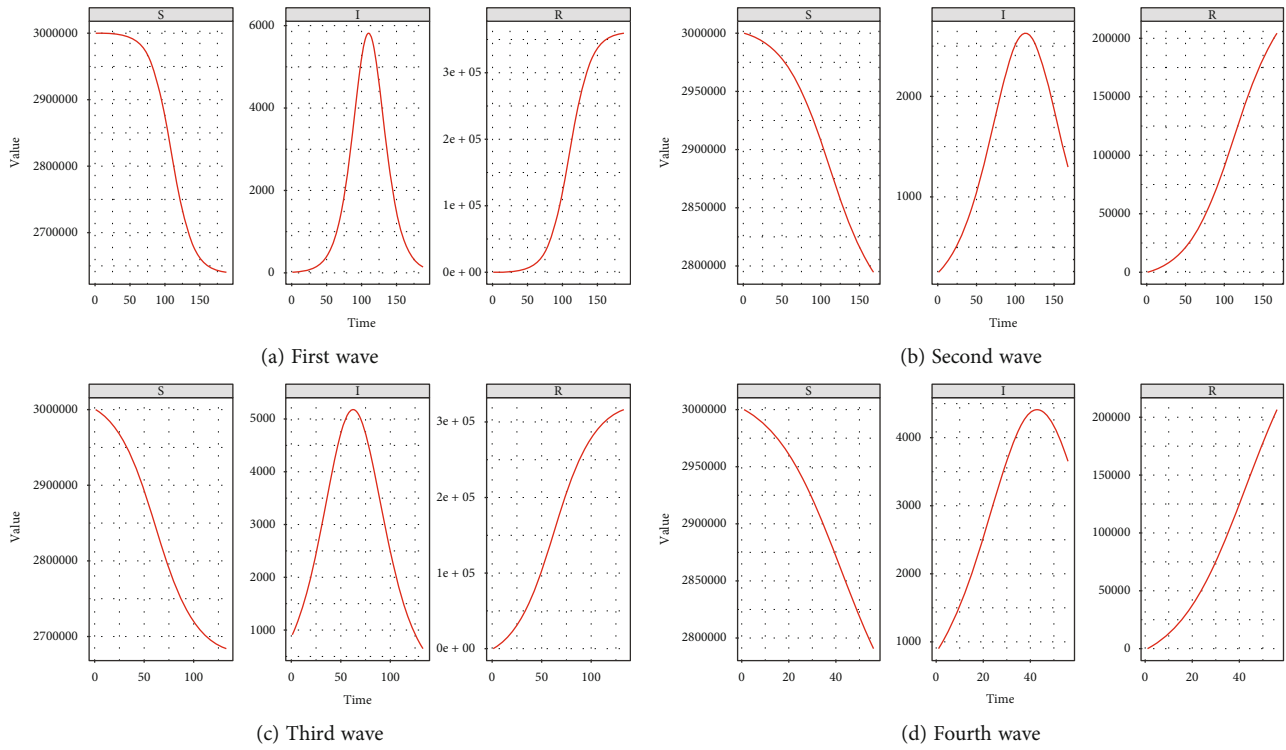


FIGURE 14: Wave-wise estimated number of susceptible, infected, and recovered people by the pSIR model.

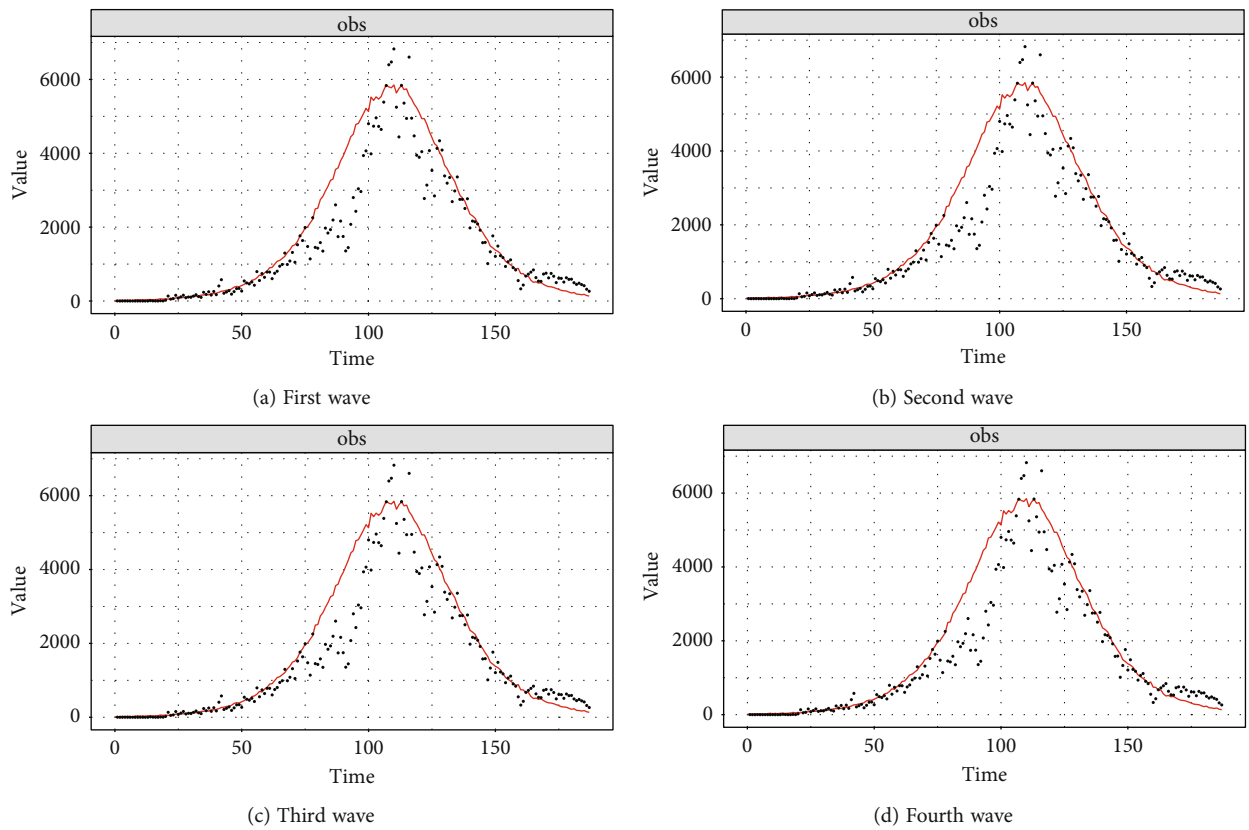


FIGURE 15: Wave-wise fitted pSIR model to the data.

TABLE 8: Comparison of different prior and likelihood distributions of R_0 .

Prior distribution	Log-likelihood distribution	Waves	R_0	P_{inf}	RSS
Uniform	Poisson	1	1.0656	0.9709	116,725,361
		2	1.0413	0.7637	29,235,687
		3	1.0561	0.8080	30,256,324
		4	1.0507	1.2175	21,548,964
Normal	Normal	1	1.0656	0.9709	119,623,567
		2	1.0413	0.7637	30,887,555
		3	1.0561	0.8080	31,998,541
		4	1.0507	1.2175	21,999,219
Poisson	Poisson	1	1.0656	0.9709	121,938,652
		2	1.0413	0.7637	31,658,742
		3	1.0561	0.8080	32,526,487
		4	1.0507	1.2175	23,145,985

Poisson distribution for the log-likelihood, the RSS is the least for all waves as compared to other combinations of distributions.

6. Conclusion and Recommendations

In the past, many pandemics have occurred throughout the world, and epidemiological modeling is always an attractive research field in predicting disease dynamics and making optimal decisions. Currently, the deadly COVID-19 pandemic has affected our lives. In this study, we focused on the COVID-19 data from Pakistan and modeled it using epidemiological models. To this end, we used the tSIR model for disease prediction using people with respiratory issues. We observed that the forward simulation of the model for the data was very close to the observed data. It is also discovered that COVID-19-positive cases are underreported using the tSIR model because everyone is frightened of this deadly virus and does not report the COVID-19 test result, and even test reports are biased/forged. As a result, the first estimated number of people who are vulnerable is just 1% of the total population.

Next, we used the two SIR models; one is the traditional model, and the other one is the pSIR model (using uniform distribution as prior distribution of R_0). Taken together, the COVID-19 pandemic characteristics are inconsistent with the SIR modeling paradigm, and the dynamics of this pandemic are influenced by a number of variables. The main reason for this outcome is the lack of reliable data. If we follow the assumptions of the model, the SIR models generate relatively coarse fitted curves, as we saw in the analysis. If we utilize the optimal parameter values, R_0 approaches 1, which means that if we use the fixed infectious period (as determined by WHO), R_0 will be small. This implies that COVID-19 is not a pandemic in Pakistan. One outcome which is often understated is that the simple SIR model, with some minor modifications, does a remarkably good job at predicting the size, extent, and shape of a single wave. That is why the many more sophisticated extensions of SIR all seem to largely agree, which has been a great advantage for the health response teams. By the middle of the first wave,

our hospitals were overrun with infections, and thus, without a question mark, the Pakistani government has responded quickly to this deadly virus by implementing clever and full lockdowns, public awareness campaigns, quarantines, and screening centers. To combat future pandemics, we recommend that the government establish additional hospitals at each district level. In addition, there should be a secure and proper data reporting mechanism.

There are several problems in this study that have gone unnoticed. For example, we did not use vaccination data and avoided the cases that were exposed. Only the persons with respiratory issues were utilized as a variable connected to the daily number of cases in this study. Thus, the tSIR model may be studied with “daily deaths” as a primary variable. Furthermore, the extended SIR models may be used to examine wave-by-wave data.

Data Availability

The dataset used in this article is available at <https://ncoc.gov.pk/>.

Conflicts of Interest

The authors have no relevant financial or nonfinancial interests to disclose.

Supplementary Materials

The supplementary text contains the additional tables and figures. (*Supplementary Materials*)

References

- [1] Y. Wan, J. Shang, R. Graham, R. S. Baric, and F. Li, “Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus,” *Journal of Virology*, vol. 94, no. 7, 2020.
- [2] T. Singhal, “A review of coronavirus disease-2019 (COVID-19),” *The Indian Journal of Pediatrics*, vol. 87, no. 4, pp. 281–286, 2020.

- [3] A. Waris, A. U. Khan, M. Ali, A. Ali, and A. Baset, "COVID-19 outbreak: current scenario of Pakistan," *New Microbes and New Infections*, vol. 35, p. 100681, 2020.
- [4] H. Akhtar, M. Afridi, S. Akhtar et al., "Pakistan's response to COVID-19: overcoming national and international hypes to fight the pandemic," *JMIR Public Health and Surveillance*, vol. 7, no. 5, p. e28517, 2021.
- [5] M. Imran, B. Ahmad, F. Maqbool et al., "Preventive measures and active response against COVID-19 taken by all four provinces, two independent territories, and the federal state of Islamic Republic of Pakistan, Islamabad," *Islamabad. Soc Work Public Health*, vol. 37, no. 4, pp. 381–396, 2022.
- [6] H. H. Weiss, "The SIR model and the foundations of public health," *Materials Matematics*, vol. 3, pp. 1–17, 2013.
- [7] B. F. Finkenstädt and B. T. Grenfell, "Time series modelling of childhood diseases: a dynamical systems approach," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 49, no. 2, pp. 187–205, 2000.
- [8] B. T. Grenfell, O. N. Bjornstad, and B. F. Finkenstädt, "Dynamics of measles epidemics: scaling noise, determinism, and predictability with the TSIR model," *Ecological Monographs*, vol. 72, no. 2, pp. 185–202, 2002.
- [9] C. Katris, "A time series-based statistical approach for outbreak spread forecasting: application of COVID-19 in Greece," *Expert Systems with Applications*, vol. 166, article 114077, 2021.
- [10] S. Pasquali, A. Pievatolo, A. Bodini, and F. Ruggeri, "A stochastic SIR model for the analysis of the COVID-19 Italian epidemic," 2021, <http://arxiv.org/abs/210207566>.
- [11] Y. C. Chen, P. E. Lu, C. S. Chang, and T. H. Liu, "A time-dependent SIR model for COVID-19 with undetectable infected persons," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3279–3294, 2020.
- [12] V. Deo, A. R. Chetiya, B. Deka, and G. Grover, "Forecasting transmission dynamics of COVID-19 in India under containment measures—a time-dependent state-space SIR approach," *Statistics and Applications*, vol. 18, no. 1, pp. 157–180, 2020.
- [13] E. B. Postnikov, "Estimation of COVID-19 dynamics "on a back-of-envelope": does the simplest SIR model provide quantitative parameters and predictions?," *Chaos, Solitons & Fractals*, vol. 135, p. 109841, 2020.
- [14] C. J. E. Metcalf, O. Bjornstad, M. Ferrari et al., "The epidemiology of rubella in Mexico: seasonality, stochasticity and regional variation," *Epidemiology & Infection*, vol. 139, no. 7, pp. 1029–1038, 2011.
- [15] M. Lavielle, M. Faron, J.-D. Zeitoun, and J. H. Lefèvre, "Extension of a SIR model for modelling the propagation of Covid-19 in several countries," *medRxiv*, 2020.
- [16] O. N. Bjornstad, B. F. Finkenstädt, and B. T. Grenfell, "Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model," *Ecological Monographs*, vol. 72, no. 2, pp. 169–184, 2002.
- [17] V. Deo and G. Grover, "A new extension of state-space SIR model to account for underreporting - an application to the COVID-19 transmission in California and Florida," *Results in Physics*, vol. 24, p. 104182, 2021.
- [18] Y. Fang, Y. Nie, and M. Penny, "Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: a data-driven analysis," *Journal of Medical Virology*, vol. 92, no. 6, pp. 645–659, 2020.
- [19] L. Brugnano, F. Iavernaro, and P. Zanzottera, "A multiregional extension of the SIR model, with application to the COVID-19 spread in Italy," *Mathematical Methods in the Applied Sciences*, vol. 44, no. 6, pp. 4414–4427, 2021.
- [20] M. J. Ferrari, R. F. Grais, N. Bharti et al., "The dynamics of measles in sub-Saharan Africa," *Nature*, vol. 451, no. 7179, pp. 679–684, 2008.
- [21] C. L. Althaus, "Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa," *PLOS Currents*, vol. 6, p. 2, 2014.
- [22] C. Li, Y. Pei, M. Zhu, and Y. Deng, "Parameter estimation on a stochastic SIR model with media coverage," *Discrete Dynamics in Nature and Society*, vol. 2018, 7 pages, 2018.
- [23] I. Cooper, A. Mondal, and C. G. Antonopoulos, "A SIR model assumption for the spread of COVID-19 in different communities," *Chaos, Solitons & Fractals*, vol. 139, p. 110057, 2020.
- [24] M. Yousaf, S. Zahir, M. Riaz, S. M. Hussain, and K. Shah, "Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan," *Chaos, Solitons & Fractals*, vol. 138, p. 109926, 2020.
- [25] R. M. Anderson, "Discussion: the Kermack-McKendrick epidemic threshold theorem," *Bulletin of Mathematical Biology*, vol. 53, no. 1-2, pp. 1–32, 1991.
- [26] P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen, "Complexity of the basic reproduction number (R_0)," *Emerging Infectious Diseases*, vol. 25, no. 1, pp. 1–4, 2019.