



# Estimation of Mortality Rate of COVID-19 in India using SEIRD Model

Dharmaraja Selvamuthu<sup>1</sup> · Deepak Khichar<sup>1</sup> · Priyanka Kalita<sup>1</sup> · Vidyottama Jain<sup>2</sup>

Accepted: 15 August 2021 / Published online: 20 September 2021  
© Operational Research Society of India 2021

## Abstract

In India, the number of infections is rapidly increased with a mounting death toll during the second wave of Coronavirus disease (COVID-19). To measure the severity of the said disease, the mortality rate plays an important role. In this research work, the mortality rate of COVID-19 is estimated by using the Susceptible-Exposed-Infected-Recovered-Dead (SEIRD) epidemiological model. As the disease contains a significant amount of uncertainty, a fundamental SEIRD model with minimal assumptions is employed. Further, a basic method is proposed to obtain time-dependent estimations of the parameters of the SEIRD model by using historical data. From our proposed model and with the predictive analysis, it is expected that the infection may go rise in the month of May-2021 and the mortality rate could go as high as 1.8%. Such high rates of mortality may be used as a measure to understand the severity of the situation.

**Keywords** COVID-19 · SEIRD model · Mortality rate · Parameter estimation

## 1 Introduction

The number of infections of COVID-19 in India has put it as the second-worst affected country in the world with more than 350,000 reported cases. There is a notable imbalance between the supply and the demand of medicines and oxygen cylinders in the healthcare system. With every passing day, the death toll rises with 3000 deaths reported every day due to COVID-19 pandemic. This infectious disease

---

✉ Dharmaraja Selvamuthu  
dharmar@maths.iitd.ac.in

Vidyottama Jain  
vidyottama.jain@curaj.ac.in

<sup>1</sup> Indian Institute of Technology Delhi, New Delhi, Delhi, India

<sup>2</sup> Department of Mathematics, Central University of Rajasthan, Ajmer 305801, India

is historically different in the following ways. The asymptomatic cases, one of the cause of virus transmission among the healthy population, are practically unnoticed by the healthcare system. Also, the government policies, social movements and actions by healthcare departments of any country are not yet well studied in epidemiology.

In the literature, a variety of stochastic and mathematical models with different complexity levels were presented and discussed ([1–4], etc.). Few of them validated the COVID-19 pandemic results and predicted the trajectory of COVID-19 in order to match the available data of the region under consideration ([1–3], etc.). Note that all these models follow one common approach, i.e., introducing distinguishable compartments to all the possible states of a person living in an affected region. Stanford Model [5] assumed different compartments such as exposed, symptomatic, pre-symptomatic, hospitalized, and recovered for a COVID-19 patient's journey from beginning to end. This assumption led to a large number of unpredictable interactions when the parameters of the mathematical models are tuned accordingly. With every single moment, the rapid rise in COVID-19 cases caused the exact data recording task challenging and unreliable. Consequently, mathematical modeling with such data may be questionable and the confidence interval of the obtained results may be uncertain. Thus, the trade-off between the description and accuracy of the mathematical model should be given the the greatest preference.

The best way to measure mortality is by calculating rate of mortality rather than counting the number of deaths as the number of deaths is heavily influenced by the number of people who are at the risk of dying. Therefore, the mortality rate is considered as the fundamental criterion to assess the state of any disease. In this research work, the mortality rate is estimated with the historical data. It is observed that the results shown in this work are significantly relevant to the actual data and contain less absolute percentage of error.

The rest of this paper is organized as follows: Sect. 2 represents a review over the relevant mathematical models developed for COVID-19 outbreak. The Susceptible-Exposed-Infected-Recovered-Dead (SEIRD) model is proposed and the mechanism of mortality rate calculation is described in Sect. 3. Further, case study of COVID-19 pandemic by using India data is discussed in Sect. 4. Finally, concluding remarks are summarized in Sect. 5.

## 2 Literature review

During COVID-19 pandemic, the prediction and the mortality rate may be estimated by using various epidemiological models, namely, Susceptible-Infected-Recovered (SIR), Susceptible-Exposed-Infectious-Removed (SEIR), Susceptible-Exposed-Infectious-Removed-Death (SEIRD), etc. In the last three decades, a lot of research works were carried out to measure the mortality rate. A few relevant ones are as follows, Piccolomin and Zama [6] proposed a forced SEIRD model for analysis and forecast of the COVID-19 spread in the Italian region. By comparing the predicted results with the actual data, it was concluded that the model is quickly adapted to monitor various infected areas at different epidemic stages.

To predict about the COVID-19 outbreak, its dynamic behaviour should be considered as the major factor. Therefore, Rapolu et al. [7] studied a time-dependent SEIRD model to anticipate the COVID-19 transmission. A SEIRD model based on partial differential equations was studied by Viguerie et al. [8]. They presented a strong qualitative agreement between the simulated data of the spatio-temporal COVID-19 spread and epidemiological data for the Italian region of Lombardy. The identification and estimation of the SEIRD model for COVID-19 are discussed in [9]. The identification of the model from the observed number of deaths and confirmed cases is not up to the mark. Thus, by using the Monte Carlo method, Korolev [9] finds some fairly accurate estimations for reproduction numbers. Ala'raj et al. [10] developed a dynamic hybrid model based on SEIRD and studied the properties related to the COVID-19 pandemic. Their model analyzed the real-time data and provided long-term and short-term forecasts with confidence intervals.

He et al. [2] proposed a SEIR model based on the conditions of hospitals, quarantine, and external inputs for COVID-19 spread. Lopez and Rodo [11] presented the modified SEIR compartmental model by incorporating the effects of varying proportions of lock-downs responsible for minimizing the COVID-19 infection. Further, Carcione et al. [12] extended a SEIR epidemiological model to analyze the tailored measures of epidemic control. The tailored measures of epidemic control contained group-specific protection and the use of tracing apps.

Roda et al. [13] investigated that a SIR model performed better than a SEIR model in representing the confirmed-case data using the Akaike Information Criterion (AIC) for COVID-19. Mondal and Antonopoulo [14] adapted SIR modelling for the prediction of the novel COVID-19 disease and studied its functionality. To track the transmission and recovering rate at time  $t$ , Chen et al. [1] performed a time-dependent SIR model for COVID-19.

Okuonghaea and Omame [15] discussed the impact of various non-pharmaceutical control measures (government and personal) on the population dynamics of COVID-19 for Lagos and Nigeria. Through the numerical simulations, they showed the effect of control measures such as common social distancing, use of face mask, and case detection on the dynamics of COVID-19 spread.

Several other researchers discussed different approaches to predict the COVID-19 outbreak and to capture the trend of various COVID-19 cases. Singhal et al. [4] studied two different models; the first model is the parametric model for various parameters relating to the spread of the virus, while the second one is a non-parametric model based on the Fourier decomposition method (FDM), fitted on the available data. They provided the results for India, Italy, and the United States of America (USA). The cases of COVID-19 were compressed into a short period and were strongly shaped by population aggregation and heterogeneity by Rader et al. [16]. By pairing their estimates with globally comprehensive data on human mobility, they predicted that crowded cities worldwide could experience more prolonged epidemics. Different models on Auto-Regressive Integrated Moving Average (ARIMA) were developed to predict the epidemiological behavior of COVID-19 in Italy, Spain, and France [3].

All over the world, during January 2020–June 2021, mortality is mostly affected by COVID-19, analyzed by John Hopkins University & Medicine [17]. Liang et al. [18] studied the factors related to cross-country variation in COVID-19 mortality. The

COVID-19 mortality rate is positively related to the proportion of the population aged 65 or older and transport infrastructure quality. However, it is negatively related to the COVID-19 test number per 100 people, government effectiveness, and the number of hospital beds. An optimal control model for the spreading of the COVID-19 pandemic was discussed by Dhaiban and Jabbar [19]. The effect of population density on COVID-19 spread and related mortality in India was investigated by Bhadra et al. [20]. By using correlation and regression analysis of infection and mortality rates due to COVID-19, they found moderate relation between COVID-19 spread and population density.

### 3 Model description

In this section, a variant of the SEIR epidemiological model is used, namely, SEIRD, which computes the number of people infected (or dead) with a contagious illness for a total population at a given time.

The total population, say  $N$ , is partitioned into five compartments:

- $\mathcal{D}_n$  : The number of deceased individuals till the  $n$ th day.
- $\mathcal{R}_n$  : The number of recovered individuals till the  $n$ th day.
- $\mathcal{I}_n$  : The number of infectious individuals on the  $n$ th day. These individuals may infect others. This compartment is called the infectious individuals compartment.
- $\mathcal{E}_n$  : The number of exposed individuals on the  $n$ th day. These individuals are already infected with the virus but are not yet infectious. This compartment is called the exposed individuals compartment. Over the time, these individuals may become cause of spreading the disease while transitioning into the infectious individuals compartment.
- $\mathcal{S}_n$  : The number of susceptible individuals on the  $n$ th day. When a susceptible person comes in close contact with an infectious one, he may contract the virus and will be transitioned into the exposed individuals compartment.

Therefore, for a given  $n$ th day, we have  $\mathcal{S}_n + \mathcal{E}_n + \mathcal{I}_n + \mathcal{R}_n + \mathcal{D}_n = N$ . While transitioning from one compartment to another, individuals behavior will be governed by the following system of equations:

$$\begin{aligned}
 \mathcal{S}_{n+1} &= \mathcal{S}_n - \sum_{i=0}^{\mathcal{I}_n} \Lambda_n^i \\
 \mathcal{E}_{n+1} &= \mathcal{E}_n + \sum_{i=0}^{\mathcal{I}_n} \Lambda_n^i - \sum_{i=0}^{\mathcal{E}_n} \Phi_n^i \\
 \mathcal{I}_{n+1} &= \mathcal{I}_n + \sum_{i=0}^{\mathcal{E}_n} \Phi_n^i - \sum_{i=1}^{\mathcal{I}_n} \Gamma_n^i - \sum_{i=1}^{\mathcal{I}_n} \Delta_n^i \\
 \mathcal{R}_{n+1} &= \mathcal{R}_n + \sum_{i=1}^{\mathcal{I}_n} \Gamma_n^i \\
 \mathcal{D}_{n+1} &= \mathcal{D}_n + \sum_{i=1}^{\mathcal{I}_n} \Delta_n^i
 \end{aligned} \tag{1}$$

where

- $\Lambda_n^i$  is the number of individuals that are currently in  $\mathcal{S}_n$  and will be infected by the  $i$ th individual of  $\mathcal{I}_n$  on the  $n$ th day. These will be transitioning into  $\mathcal{E}_{n+1}$ .
- $\Phi_n^i$  is 1 or 0 depending on whether an  $i$ th individual of  $\mathcal{E}_n$  will go to  $\mathcal{I}_{n+1}$  (and hence become infectious) or not.
- $\Gamma_n^i$  is 1 or 0 depending on whether an  $i$ th individual of  $\mathcal{I}_n$  will transition to  $\mathcal{R}_{n+1}$  (after recovering from the disease).
- $\Delta_n^i$  is 1 if the virus results in the death of the  $i$ th of  $\mathcal{I}_n$ , thereby transitioning to  $\mathcal{D}_{n+1}$ ; otherwise 0.

Here, for  $n, i \in \mathbb{N}$ ,  $\Phi_n^i$ ,  $\Gamma_n^i$ , and  $\Delta_n^i$  are assumed to be independently and identically distributed random variables. Let  $\epsilon$ ,  $\gamma$  and  $\delta$  be the respective expectations of  $\Phi_n^i$ ,  $\Gamma_n^i$ , and  $\Delta_n^i$ . Here,  $\epsilon$  represent the rate of exposed becoming infectious,  $\gamma$  represent the probability of an infectious individual recovering in a day and  $\delta$  represent the probability of an infectious individual dying in a day. Another variable  $\beta$  represent the number of infectious people interact with others per day. Though the case of  $\Lambda_n^i$  is more subtle than it seems,  $\forall n, i \in \mathbb{N}$ , these are assumed to be independently and identically distributed if the population is infinite. Due to the finite population, infectious individuals are bound to make contact with those who have already recovered and hence are immune for the time. Since  $\beta$  can be considered as the number of infectious people interact with others per day and the probability of the occurrence of  $\Lambda_n^i$  is given by  $\frac{S_n}{N}$ , it is observed that  $\Lambda_n^i$  is a binomial distributed random variable with parameters  $\beta$  and  $\frac{S_n}{N}$ . Thus, the expression for expectation of  $\Lambda_n^i$  becomes:

$$E[\Lambda_n^i] = \beta \frac{S_n}{N}. \tag{2}$$

Let  $E[\mathcal{S}_n] = S_n$ ,  $E[\mathcal{E}_n] = E_n$ ,  $E[\mathcal{I}_n] = I_n$ ,  $E[\mathcal{R}_n] = R_n$  and  $E[\mathcal{D}_n] = D_n$ . Using the process governing equations, the expectation for the random sum of  $\Lambda_n^i$ , by Wald’s equation, is given by [21]

$$E \left[ \sum_{i=0}^{\mathcal{I}_n} \Lambda_n^i \right] = E[\mathcal{I}_n]E[\Lambda_n^i] = \beta I_n \frac{S_n}{N}. \quad (3)$$

Similarly, the expectation for the random sum of other random variables are [21]

$$E \left[ \sum_{i=0}^{\mathcal{E}_n} \Phi_n^i \right] = E[\mathcal{E}_n]E[\Phi_n^i] = \epsilon E_n \quad (4)$$

$$E \left[ \sum_{i=1}^{\mathcal{I}_n} \Gamma_n^i \right] = E[\mathcal{I}_n]E[\Gamma_n^i] = \gamma I_n \quad (5)$$

$$E \left[ \sum_{i=1}^{\mathcal{I}_n} \Delta_n^i \right] = E[\mathcal{I}_n]E[\Delta_n^i] = \delta I_n. \quad (6)$$

Thus, the expected number of individuals in different compartments are as follows [21]:

$$S_{n+1} = S_n - \beta I_n \frac{S_n}{N} \quad (7)$$

$$E_{n+1} = E_n + \beta I_n \frac{S_n}{N} - \epsilon E_n \quad (8)$$

$$I_{n+1} = I_n + \epsilon E_n - \gamma I_n - \delta I_n \quad (9)$$

$$R_{n+1} = R_n + \gamma I_n \quad (10)$$

$$D_{n+1} = D_n + \delta I_n. \quad (11)$$

In the proposed model, the values of the parameters are taken as  $\epsilon = 0.2$  as mentioned in [22] and  $\gamma = 0.1$  as mentioned in [22]. The parameters  $\delta$  and  $\beta$  will be estimated by using the historical data in Sect. 4.

### 3.1 Mortality rate

The mortality rate on the  $n$ th day, say  $MR_n$ , which depends on the ratio of the number of people who died due to COVID-19 and the number of people who got infected on that day, is calculated as

$$MR_n = \frac{D_n - D_{n-1}}{I_n}, \quad \forall n \in \mathbb{N}. \quad (12)$$

This may provide inaccurate results. The reason is that patients, who died on a particular day, were infected by the virus roughly two weeks before. To obtain the better results, it should be considered into account. Backer et al. [23] mentioned that the time between the contraction of the virus and the onset of symptoms is approximately 14 days. On the other side, Huang et al. [24] and Wang et al. [25] reported that the median time from the onset of the symptoms to being admitted into the intensive care unit (ICU) is near about ten days. Additionally, the World Health Organization (WHO) mentioned that the time between the onset of symptoms and death ranges from two weeks to 8 weeks [26]. Hence, there is a requirement to introduce mortality rate based on the ratio of the number of people who died and the number of people who got infected few days back.

Therefore, true mortality rate, say  $TMR_n$ , which is obtained by dividing the number of deaths on  $n$ th day by the number of confirmed COVID-19 infections  $k$ th days before of that  $n$ th day, is defined as

$$TMR_n = \frac{D_n - D_{n-1}}{I_{n-k}} \quad (13)$$

where  $k \in \mathbb{N}$ . This calculation counts  $k$  as 14 as mentioned in [23].

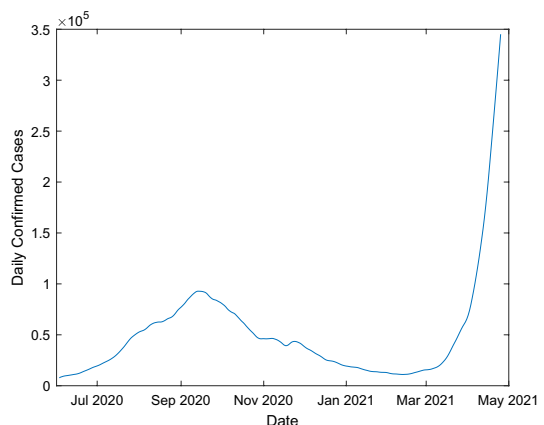
## 4 Results and discussion

The model described in Sect. 3 will be employed to obtain the results for the prediction of the mortality rate with the India data [27].

### 4.1 India data

Figure 1 depicts the variation in daily confirmed cases in India during the past few months. While India was celebrating its success in combating the first wave of COVID-19, the number of COVID-19 infections again began to increase in

**Fig. 1** Daily confirmed cases in India



March 2021 and afterwards. The number of infections per day was more in comparison with the earlier reported cases in India. Consequently, India unexpectedly entered deep into the second wave of coronavirus. In the total tally of infections, India reached at the second position after USA.

The main cause of this sudden surge was considered as a mutant of the earlier version, commonly known as the double mutant. This particular strain was found in many countries, like, UK, South Africa, etc., other than India. Though it has been a subject of global interest, but it is still not confirmed scientifically. The transmission of the virus from an infected person to a susceptible one is very much similar as it happened in the first wave (Oct 2020). Since the spread of this virus is dependent on location, person, etc., it is difficult to characterize the behavior of this virus at the equal level all over India. One other prime factor behind the massive increase in the number of cases is the increased testing drive.

## 4.2 Parameter estimation

It is very important that parameter values should be estimated appropriately and efficiently. Values of the parameters  $\gamma$  and  $\epsilon$  are considered as provided by Kuniya et al. [22]. This section is devoted for estimating the values of  $\delta$  and  $\beta$  by using the historical data for India [28]. These parameters are prone to fluctuate due to their correlation with the political state of the region. Though the lockdown is supposed to be the best way to reduce the virus transmission, uplifting the previously imposed restrictions may tend to increase the people's interaction in person. Further, it is also observable that these parameters are different for disparate regions. Therefore, the parameters should be estimated based on the data to relate the model with the actual scenario. Kuniya et al. [22] proposed an estimation method in which the parameter is a simple random variable and the estimate is a single value. Since such methods may be inaccurate at times, a basic method to find a time-dependent estimate for the parameters  $\delta$  and  $\beta$  using historical data is intended to present in this paper to improve the accuracy of the model.

The parameter  $\delta$  is estimated by taking the average of  $\hat{\delta}_n$ , where by using Eq. (11),  $\hat{\delta}_n$  is calculated as follows

$$\hat{\delta}_n = \frac{\Delta D_n}{I_{n-1}}.$$

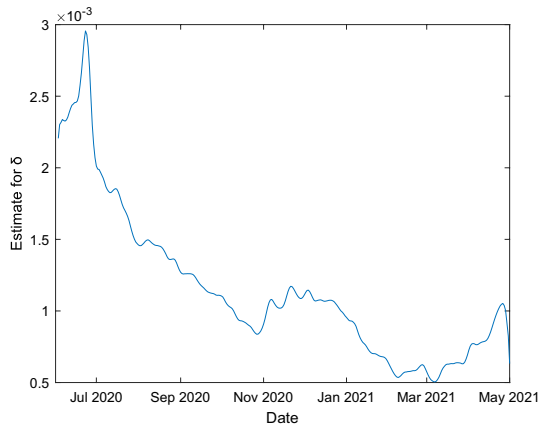
Here,  $\Delta D_n$  represents the difference between the number of deceased individuals on  $n - 1$ th day and  $n$ th day, i.e.,  $\Delta D_n = D_n - D_{n-1}$  and,  $I_{n-1}$  represents the data of the number of infected individuals on  $n - 1$ th day. By using the historical data,  $\delta$  is estimated as 0.0012.

Further, the following steps are for the estimation of  $\beta$ , say  $\hat{\beta}_n$ .

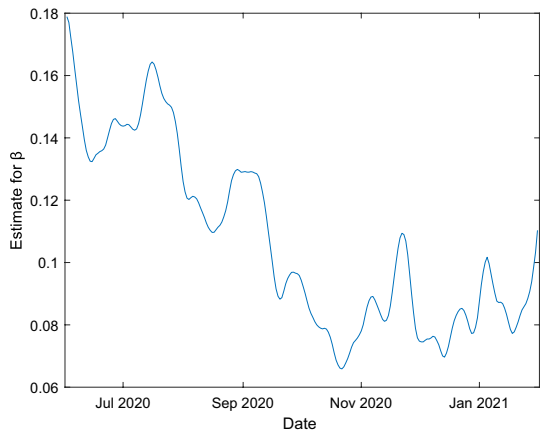
Step 1: By using Eq. (9) and using the data of the number of infected individuals till the  $n$ th day, the number of exposed individuals may be calculated by



**Fig. 2** Estimate for  $\delta$



**Fig. 3** Estimate for  $\beta$



$$E_{n-1} = \frac{1}{\epsilon}(\Delta I_n + (\gamma + \delta)I_{n-1})$$

where  $\Delta I_n$  represents the difference between the number of infected individuals on the  $n - 1$ th day and  $n$ th day, i.e.,  $\Delta I_n = I_n - I_{n-1}$ ,  $\gamma = 0.1$  [22] and  $\epsilon = 0.2$  [22].

Step 2: By using Eq. (8), the estimate for  $\beta$  is made from the exposed individuals as

$$\hat{\beta}_n = \frac{N}{I_{n-2}S_{n-2}}(\Delta E_{n-1} + \epsilon E_{n-2}) \tag{14}$$

where  $\Delta E_{n-1}$  represents the difference between the number of exposed individuals on the  $n - 2$ th day and  $n - 1$ th day, i.e.,  $\Delta E_{n-1} = E_{n-1} - E_{n-2}$ .

Results of the proposed estimation process of  $\delta$  and  $\beta$  are shown in Figs. 2 and 3, respectively.

These estimations are for the first wave in India. Notice that there is an overall downward trend with seasonal fluctuations. Such behavior may be attributed to the overall incessant efforts of the authorities and public to keep the virus in check. The second wave, as discussed earlier, is assumed to follow similar transition patterns as the first wave. The next section discusses the predictions made for the second wave.

### 4.3 Model predictions

The expected number of infectious and deceased individuals are predicted using the model (Eq. (1)) and data [27] provided in the previous sections. Infectious means the total number of active infections at a point of time, which is equal to the cumulative sum of all daily reported infections minus the cumulative sum of daily removals (removals include both the recovered and deceased individuals). For deceased individuals, the average daily number of deaths is predicted.

To run the model and simulate the number of infectious individuals, a starting point is needed. This work makes predictions starting from April 2021 till August 2021. Naturally, 1st April 2021 is chosen as the initial point and the model is run using Eq. (9). Figure 4 exhibits the predictions for several active infections. It is observed from Fig. 4 that in the one-month prediction, i.e., in May 2021 the absolute percentage error for infectious individuals (calculated as  $\frac{|I_n(\text{actual}) - I_n(\text{predicted})|}{I_n(\text{actual})} \times 100\%$ ) is less than 20%. After the one-month prediction, the absolute percentage error increases, and from June 15, 2021 onwards the absolute percentage error decreases.

The second wave of coronavirus in India looks much harsher than the first one. The peak infections are much ahead and the daily impending death toll is creating havoc within the country.

From the predictions, it is inferred that the peak of the second wave is expected to occur somewhere in May 2021 with expected active infections of well over 4 million. The second wave is higher than the first one but it is expected that the number of infections returning to a manageable scale by August 2021.

The predictions for the number of daily deceased individuals are shown in Fig. 5. It is noticed from Fig. 5 that in the one-month prediction, i.e., in May 2021 the absolute percentage error is less than 30%. The absolute percentage error for the number of daily deceased individuals is calculated as

$$\frac{|(D_n - D_{n-1})(\text{actual}) - (D_n - D_{n-1})(\text{predicted})|}{(D_n - D_{n-1})(\text{actual})} \times 100\%.$$

After the one-month prediction, the absolute percentage error slightly increases and decreases but the fluctuation is less than 37% and from June 23, 2021, onwards the absolute percentage error decreases.

The daily count for the deaths due to COVID-19 is expected to peak near the end of May 2021 with the count going as high as 5000. In the next section, mortality rate is estimated and predicted using the results obtained here.

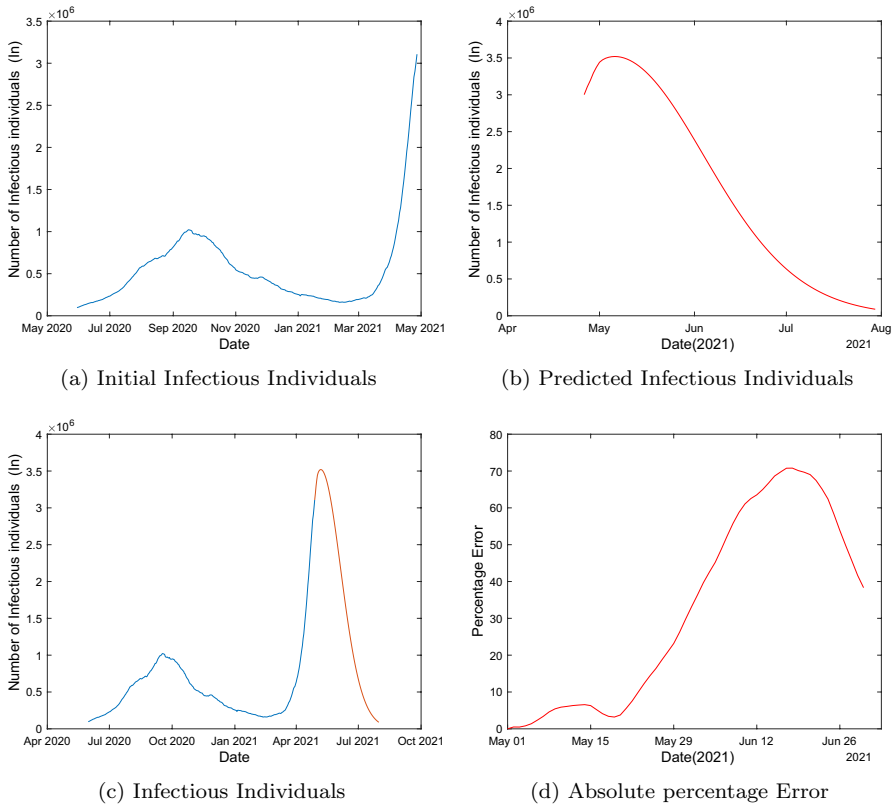


Fig. 4 Infectious individuals

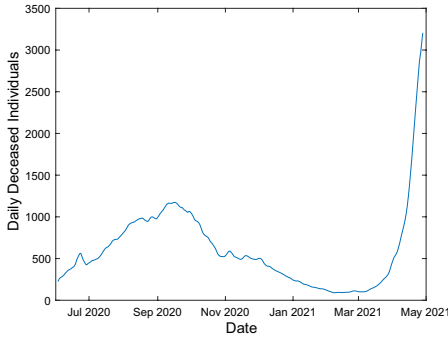
### 4.4 Estimation of true mortality rate

As mentioned earlier, the way usual mortality rate calculated using Eq. (12) is fundamentally inaccurate because of the lag between the time of infection and the time of death. Figure 6 shows the estimates for mortality rate.

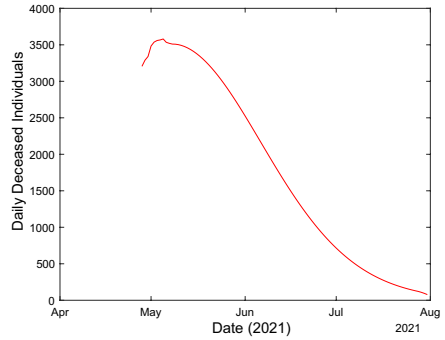
Further, the TMR formula is used to obtain the accurate mortality rate and the results are shown in Fig. 7.

TMR is found to somewhat higher than the usual one. TMR shows high death rates and an upward trend during the period March–April, 2021, which seems realistic. It is claimed that TMR provides a more accurate picture of the disease.

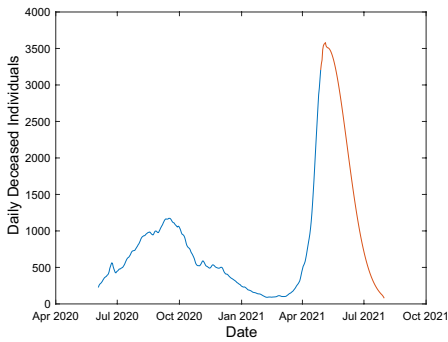
Further, Eq. (13) is applied to the predicted number of infections and deaths from the model, depicted in Fig. 8. It predicts that the mortality rate is expected to stay high (more than 1%) for most of the time post-March 2021, while peaking with a value of almost 1.8% near the beginning of May 2021. The absolute percentage error for TMR is calculated as  $\frac{|TMR(actual) - TMR(predicted)|}{TMR(actual)} \times 100\%$ . Also, in the initial predicted month for TMR, i.e., in May 2021, the absolute percentage error is less than 30%. After the one-month prediction, the absolute percentage



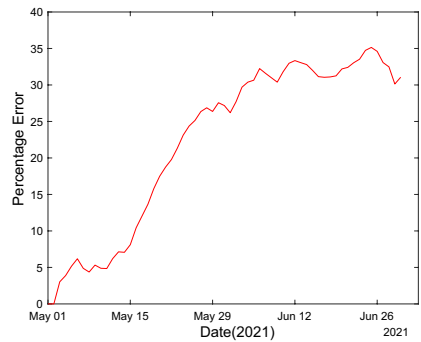
(a) Actual Daily Deceased Individuals



(b) Predicted Daily Deceased Individuals



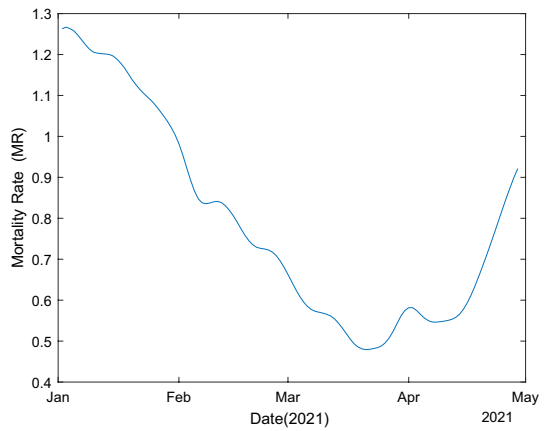
(c) Daily Deceased Individuals



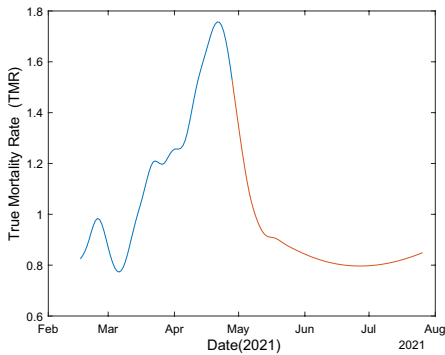
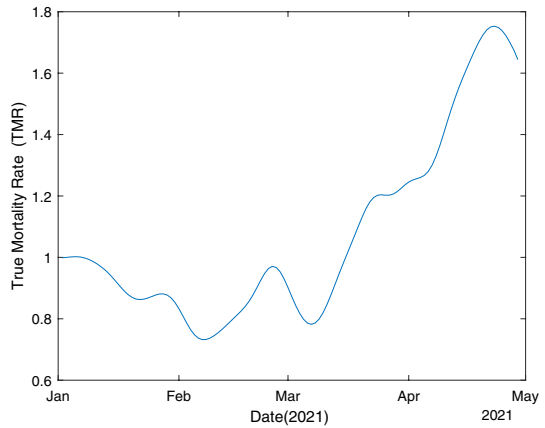
(d) Absolute percentage Error

**Fig. 5** Daily deceased individuals

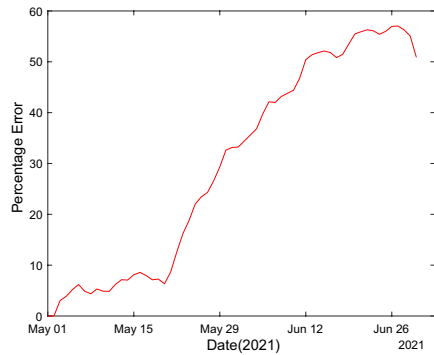
**Fig. 6** Mortality rate



**Fig. 7** True mortality rate



(a) Predicted True Mortality Rate



(b) Absolute percentage Error

**Fig. 8** Predicted true mortality rate

error increases, and from June 26, 2021, onwards the absolute percentage error decreases.

### 5 Conclusions and future work

In this paper, the SEIRD model is presented as a tool to predict the dynamic behavior of COVID-19 cases in India using available data of the reported cases of historical data. The SEIRD or SEIR model consists of susceptible, exposed, infectious, recovered and dead cases while SIR model is restricted to susceptible, infectious and recovered cases. Since SIR model does not consider the latent stage (the exposed individuals), it will be an inappropriate model to employ in the proposed scenario. Therefore, SEIR & SEIRD models can provide better approximation as compare to SIR model [12]. The mortality rate of COVID-19 is estimated by taking

into account the dynamic behavior of the spread of the disease. It was inferred that the infections are expected to peak in May-2021 and the mortality rate could go as high as 1.8%. Such high rates may be used as a measure to understand the severity of the pandemic. There is an advantage that this predictive methodology can be applied in any region, if the historical data is available. This analysis will be useful in making important decisions such as health care policies in India. In future work, the estimated results can be validated by the simulation model. In future, our proposed model can also be extended by obtaining well established reproduction rate (R) along with mortality rate. Also, the scenarios effects (base, best, worst, etc.) on a monthly basis helping policies on restrictions like lock down can also be incorporated in future.

**Acknowledgment** We thank the editor and the anonymous reviewers for their insightful comments, which helped in improving the paper. We thank Mr. Nimesh Sangwan for preparing the initial draft of this paper. This research work is supported by the Department of Science and Technology, India.

**Authors Contribution** All authors confirm the responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

**Data availability** The datasets generated and/or analysed during the current study are available in the following weblink: Ministry of Health and Family Welfare, Government of India. <http://nhm.gov.in/>

## Declarations

**Conflict of interest** The authors declare that they have no competing interests

**Ethical approval** Manuscripts reporting studies involving human participants, human data or human tissue: Not applicable

**Consent for publication** manuscript does not contain data from any individual person, hence, it is “Not applicable”

## References

1. Chen, Y.C., Lu, P.E., Chang, C.S., Liu, T.H.: A time-dependent SIR model for COVID-19 With undetectable infected persons. *IEEE Trans. Netw. Sci. Eng.* **7**(4), 3279–3294 (2020)
2. He, S., Peng, Y., Sun, K.: SEIR modeling of the COVID-19 and its dynamics. *Nonlinear Dyn.* **101**, 1667–1680 (2020)
3. Ceylan, Z.: Estimation of COVID-19 prevalence in Italy, Spain, and France. *Sci. Total Environ.* **729**, 1–7 (2020)
4. Singhal, A., Singh, P., Lall, B., Joshi, S.D.: Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos, Solut. Fractals* **138**, 1–8 (2020)
5. Potential Long-Term Intervention Strategies for COVID-19. <http://covid-measures.stanford.edu/>
6. Piccolomini, E.L., Zama, F.: Monitoring Italian COVID-19 spread by a forced SEIRD model. *Plos one* **15**(8), 1–17 (2020)
7. Rapolu, T., Nutakki, B., Rani, T.S., Bhavani, S.D.: A Time-Dependent SEIRD Model for Forecasting the COVID-19 Transmission Dynamics. *medRxiv* 2020.05.29.20113571
8. Viguerie, A., Lorenzo, G., Auricchio, F., Baroli, D., et al.: Simulating the spread of COVID-19 via a spatially-resolved susceptible-exposed-infected-recovered-deceased (SEIRD) model with heterogeneous diffusion. *Appl. Math. Lett.* **111**, 1–9 (2021)

9. Korolev, I.: Identification and estimation of the SEIRD epidemic model for COVID-1. *J. Econ.* **220**, 63–85 (2021)
10. Ala'raj, M., Majdalawieh, M., Nizamuddin, N.: Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infect. Dis. Modell.* **6**, 98–111 (2021)
11. López, L., Rodó, X.: A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: simulating control scenarios and multi-scale epidemics. *Results Phys.* **21**, 1–15 (2021)
12. Carcione, J.M., Santos, J.E., Bagaini, C., Ba, J.: A simulation of a COVID-19 epidemic based on a deterministic SEIR model. *Front. Public Health* **8**, 1–13 (2020)
13. Roda, W.C., Varughese, M.B., Han, D., Li, M.Y.: Why is it difficult to accurately predict the COVID-19 epidemic? *Infect. Dis. Modell.* **5**, 271–281 (2020)
14. Mondal, A., Antonopoulo, C.G.: A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solit. Fractals* **139**, 1–14 (2020)
15. Okuonghaea, D., Omame, A.: Analysis of a mathematical model for COVID-19 population dynamics in Lagos, Nigeria. *Chaos, Solut. Fractals* **139**, 1–18 (2020)
16. Rader, B., Scarpino, S.V., Nande, A., et al.: Crowding and the shape of COVID-19 epidemics. *Nat. Med.* **26**, 1829–1834 (2020)
17. Mortality Analysis, John Hopkins University & Medicine. <https://coronavirus.jhu.edu/region/us/alabama>
18. Liang, L.L., HungTseng, C., Ho, H.J., Wu, C.Y.: COVID-19 mortality is negatively associated with test number and government effectiveness. *Scientif. Rep.* **10**, 1–7 (2020)
19. Dhaiban, A.K., Jabbar, B.K.: An optimal control model of COVID-19 pandemic: a comparative study of five countries. To appear in OPSEARCH (2021). <https://doi.org/10.1007/s12597-020-00491-4>
20. Bhadra, A., Mukherjee, A., Sarkar, K.: Impact of population density on Covid-19 infected and mortality rate in India. *Modell. Earth Syst. Environ.* **7**, 623–629 (2021)
21. Castaneda, L.B., Arunachalam, V., Dharmaraja, S.: Introduction to probability and stochastic processes with applications. Wiley, New Jersey (2012)
22. Kuniya, T.: Prediction of the epidemic peak of coronavirus disease in Japan. *J. Clin. Med.* **9**(3), 1–7 (2020)
23. Backer, J.A., Klinkenberg, D., Wallinga, J.: Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from Wuhan, China, 20–28 January 2020. *Euro Surveill.* **25**(5), 1–6 (2020)
24. Huang, C., Wang, Y., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **395**, 497–506 (2020)
25. Wang, D., Hu, B., et al.: Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *Jama* **323**(11), 1061–1069 (2020)
26. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>
27. Ministry of Health and Family Welfare, Government of India. <http://nhm.gov.in/>
28. Dharmaraja, S., Das, D.: Introduction to statistical methods, design of experiments and statistical quality control. Springer, New York (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.