

RESEARCH ARTICLE



© OPEN ACCESS Received: 21-03-2022 Accepted: 22-08-2022 Published: 07-11-2022

Citation: Bhadade R, Anuse A, Kute R, Munot M (2022) SEIRD Model for Forecasting Spread of Covid - 19. Indian Journal of Science and Technology 15(41): 2162-2170. https://doi.org/ 10.17485/IJST/v15i41.651

^{*} Corresponding author.

mvmunot@pict.edu

Funding: None

Competing Interests: None

Copyright: © 2022 Bhadade et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN Print: 0974-6846 Electronic: 0974-5645

SEIRD Model for Forecasting Spread of Covid - 19

Raghunath Bhadade¹, Alwin Anuse², Rupali Kute¹, Mousami Munot^{3*}

1 Assistant Professor, School of ECE, Dr. Vishwanath Karad MIT- WPU, Pune, Maharashtra, India

2 Associate Professor, School of ECE, Dr. Vishwanath Karad MIT- WPU, Pune, Maharashtra, India

3 Associate Professor, Department of E&TC Engg, SCTR's Pune Institute of Computer Technology, Maharashtra, India

Abstract

Objectives: To develop a reliable mathematical model in order to predict the evolution of various epidemiological factors and parameters for COVID-19 across the globe. **Methods:** A novel dynamic Susceptible- Exposed-Infected-Recovered-Died (SEIRD) model is proposed in this research. The proposed, two-step approach assumes the infection rate which is dependent on time, to estimate the evolution of various variables of the model. In the first step, parameters like clinical and transmission are estimated, whereas in the second step, simulation of the model is done to predict the outbreak. **Findings:** Making use of this model, the total number of people who are likely to be afflicted by an infectious disease in a closed population over a period of time can be computed theoretically. **Novelty:** The proposed model results into low computational complexity since it is deterministic in nature. Secondly, SEIRD model equations are solved in frequency domain that converts the integraldifferential equations into simple algebraic equations. This further reduces the computational burden.

Keywords: COVID19; Modelling; SEIRD; Clinical; Transmission; Parameters

1 Introduction

It is very critical to model the time evolution of an epidemic especially the COVID-19 pandemic^(1,2). The current models collectively describe the outbreak of the pandemic in a population making use of a restricted number of collective variables. Besides, typically, the number of parameters is very small. SEIRD models⁽³⁾, sub epidemics wave models, Richard models, Generalised Richards models, generalised Growth models, Logistic models, Susceptible-Infectious-Recovered (SIR)⁽⁴⁾ models are the ones they include. The so-called compartmental models class include SIR model assumption for the spread of COVID-19 in different communities⁽⁵⁾, SEIR and other similar models. Several mathematical techniques and methods were conducted to simulate the spread of COVID-19 in various countries.

In this research we have created a SEIRD model for an infectious disease. Making use of this model, the total number of people who are likely to be afflicted by an infectious

disease in a closed population over a period of time can be computed theoretically. The model is divided into three groups. S(t) is the number of susceptible people at any given time, I(t) is the number of infected people at any given time, and R(t) is the number of recovered people at any given time. As indicated by the variable function 't', this model functions dynamically and it is possible for the number of people in each of the three compartments to alter over time. The basic mathematical model that we have selected is the dynamic SEIRD model. The assumption is that the rate of infection decreases with passage of time due to the several measures that the government enforces such as mandatory lockdown and social distancing as these work as mitigation measures. An estimate of the model parameters is made based on the real data by optimizing the model using the actual data as the initial step of the analysis. Next, the parameters that are evaluated are utilized to simulate the forecast. Taking this into account, we feel that there is a need for a software program that will address this issue. We have built this model in R language.

Various researchers have worked on modelling the spread of $corona^{(6-18)}$. The contributions of this research lie in proposed deterministic model resulting into low computational complexity and allowing the better calibration of the model. Low computational complexity and better calibration of the model parameters is not achievable with stochastic model that has been proposed by many other researchers. The performance of the proposed model is compared with the Jenkins modelling procedure⁽¹⁸⁾. The other novelty of this research paper lies in the method used to solve SEIRD equations. SEIRD model equations are solved in frequency domain that converts the integral-differential equations into simple algebraic equations. This further reduces the computational burden.

2 Methodology

Predictive analytics is an area of data analytics that specializes in improved predictions utilizing data gathered from the past and making use of analysis techniques which encompass statistical and learning approaches. The primary goal of predictive analytics is figuring out the existence of patterns in the input data and foreseeing possible events. There are three fundamentals of predictive analytics. They are statistical analysis, predictive modelling and machine learning. Significant abilities of predictive analytics are predictive modelling, statistical analysis, logistic model and linear regression. Choice of the most suitable predictive models and algorithms plays a significant role in arriving at decisions and gaining insights that are efficient and useful. It is crucial to define the problem clearly and collect data with care in order to develop a model that will be able to foresee accurately.

A combination of clinical data parameters and claims is used for most analytical goals. An example of hospitals is a point in case. Hospitals are keen to know how many patients are likely to require admission in intensive care units for a specific period so that the hospitals can set up the required number of beds. Models like these help with such predictions and provide valuable information to hospitals in different countries. The choice of an appropriate model depends on the type of questions that need to be answered and once the information is gathered, how optimization can be undertaken making use of the predictive algorithms.

The proposed model results into a low computational complexity since the model is deterministic in nature. Many researchers have used stochastic model that requires more data which in turn increases the computational burden. Another advantage of the proposed model is that the model parameters can be calibrated. The model can also be analysed and interpreted using the theory of ordinary differential equations



Fig 1. Various Blocks in SEIRD Model

Figure 1 depicts various blocks related to the SEIRD model. Various predetermined values of various parameters as depicted in the block diagram are:

a.	Suscep- tible	(S)	:	Susceptible people are the individuals who have never been contaminated with, and consequently have no resistance against COVID-19
b.	Exposed	(E)	:	Exposed individuals are those who have been contaminated with COVID-19 but are not yet infectious to others.
c.	Infec- tions	(I1)	:	Individuals with mild infections
d.	Infec- tions	(I2)	:	Individuals with severe infections
e.	Infec- tions	(I3)	:	Individuals with critical infections
f.	Recov- ered	(R)	:	Individuals who have recovered and are assumed to be immune from covid-19
g.	Died	(D)	:	Individuals who died due tocovid-19.

There are two datasets used in this experimentation⁽¹⁹⁾. The first dataset is the COVID-19 data repository created by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University⁽²⁰⁾. It consists of COVID- 19 data collected from various countries. The country code-wise number of confirmed, deceased and activated cases is given in the dataset. The second dataset used is from Government of India covid-19 website. This dataset has given covid-19 status with respect to different states of India. It has the information of total no.of cases, active cases, recovered cases and number of deaths.

A model based on mathematics to simulate the spread of infectious diseases was developed by Kermack et al. ⁽²¹⁾. The models that were developed early on classified the population as susceptible, infected (infectious) and recovered (SIR) ⁽⁵⁾ model assumption for the spread of COVID-19 in different communities. With the evolution of the models, more complicated, compartmental models were developed along with stochastic transmission models etc. Each new pandemic or epidemic brings about newer techniques for mathematical modelling. Now, computer software that includes machine learning is also being used. There are three significant genres of infectious disease models. They are models based on:

- Statistics
- Mathematics or mechanistic state space
- Empiricism or machine learning based

The models for infectious diseases used for the purposes of this research are based on mathematics and statistical tools used here have become an intrinsic part of the inputs that are fed into the mitigation measures. These models make it possible for us to test different strategies used in simulations before they are applied to the actual population. Multiple assumptions are made, and many sources of data are tapped into when using these mathematical models for infectious diseases.

The forecasts of the model using virtual data and based on assumptions should be relevant to the real world and the model has to be as close to reality as possible. It is not possible to derive robust forecasts based on assumptions that are either incorrect, or vague or not complete. All effort made in the direction of simplifying, approximating, idealizing and abstraction results in all the models being part descriptions of the procedures that operate in reality. Hence, how robust each model is, has to be evaluated on the fact as to whether the assumptions made by the model relate to reality or not. Many a researcher in India has tried to model this dynamic pandemic mathematically right from the beginning. An analysis done on these models as regards to India showed many differences in terms of forecasts, assumptions, the mode of transmission, the effect different interventions have on the disease or the health care service, etc.

A survey of the current literature on this topic showed that no review of available models has been done. This study has been undertaken as a means to fill this gap by reviewing the existing mathematical models that were utilized to forecast the behaviour of the Coronavirus disease (COVID-19) in the context of the world narrative and to evaluate the correlation between the predicted cases and the actual values of the cases the world over. An in-depth research plan comprising two steps was developed and adopted. To begin with, multiple searches through various databases such as Google scholar, BioRxiv, Medrxiv and Medline were conducted. Any documentation or papers submitted to any of these databases on the subject of COVID were researched in detailed. The research was conducted until as recently as 10:00 am on June 16, 2021 (IST). The next step was searching for articles on COVID manually and also contacting reputed modellers to check for unpublished material based on data on COVID from all over the world.

Data that was mined from GitHub was used to develop the following variables: software used, the kind of mathematical models, the profession of such modelers, whether they had studied the effects of the lockdown, what type of assumptions their models were based on, the crest of the infected numbers and what kind of data they used as well as the source of such data. The primary measures of the summary were the predicted number of COVID cases and the rate of the peak infection. Given that the data used as input for such models was from various sources, variations in the predicted numbers were expected. Such variations were due to the type of mathematical model used, and the assumptions that the model was based on. Coefficients of determination were used to explore the relationship between the predicted numbers and the actual value. Data that was pulled up from 30 different research articles revealed that researchers had started modelling on the public domain data as far back as March 21, 2020. The methods used by these researchers were also different. A study of the different types of models that was conducted shows that about 56% of the studies used the SEIRD model or a variation thereof. Table 1 depicts the various assumptions made by different models regarding R Naught or R0, the period of infection, the time taken for recovery, the serial interval, etc. What can be inferred from this data is that researchers have used the data available on multiple public platforms around the world for deriving their models.



Last updated: 16.06.2021 - 18:56 UTC

Fig 2. Confirmed, Deceased, Estimated, Recovered and Active Number of Cases

In this study we have used the R language and RStudio software for developing the required graphs, measured the peak values and found out the R0 value by changing the parameters in the simulation run-time. In Figure 2 with the help of extracted data, we built a leaflet map to show the number of Confirmed, Deceased, Estimated, Recovered and Active Number of Cases of various countries.

3 Numerical Model and Method

From Figure 1, a time dependent model is created. In this model the population N is classified into sub-population of susceptible (S), Exposed (E), Infectious (I=I1+I2+I3), Recovered (R) and Died (D) for all times t.

N=S+E+I+R+D

$$\frac{dS}{dt} = -(\beta 1 \times I 1 - \beta 2 \times I 2 - \beta 3 \times I 3) \times S$$
(i)

$$\frac{dE}{dt} = (\beta 1 \times I \ 1 + \beta 2 \times I \ 2 + \beta 3 \times I \ 3) \times S - \alpha^* E$$
(ii)

$$\frac{dI}{dt} = -\left(\gamma 1 + \beta \underline{1}\right)^* I \underline{1}$$
(iii)

$$\frac{dI\,2}{dt} = \beta \,1^* I\,1 - \left(\gamma 2 + \beta \underline{2}\right)^* I\,2 \tag{iv}$$

$$\frac{dI3}{dt} = \beta 2^* I 2 - \left(\gamma 3 + \underline{\mu}\right)^* I 3 \tag{v}$$

$$\frac{dR}{dt} = \gamma 1^* I \, 1 + \gamma 2^* I \, 2 + \gamma 3^* I \, 3 \tag{vi}$$

$$\frac{dD}{dt} = \mu^* I \,3 \tag{vii}$$

With initial conditions [S(t0), E(t0), I(t0), R(t0), D(t0)] = [S0, E0, I0, R0, D0] for some initial time t0.

The α parameter is the rate of progression from exposed block to the infectious block. i.e., the rate of incubation. Parameter β is the infection rate, i.e., the probability per unit time that a susceptible individual contracts the disease when coming into contact with a person who has been infected. Parameters γ and μ denote the recovery and death rates respectively⁽⁶⁾.

3.1 Dynamical Model

The current research presents a dynamical model which is the SEIRD type. RStudio was used to implement the SEIRD simulations. Now to make the model more reliable and accurate to actual predictions we have added the following clinical terms into our SEIRD model. All the below reference values are taken from figure 1, table 2, table 3 and table 4.

 $\begin{array}{l} \alpha = 1/(\text{Incubation Period}) \\ \gamma = (1/\text{Duration Mild Infections}) \times \text{Fraction Mild} \\ \beta 1 = (1/\text{DurationMild Infections}) - \gamma 1 \\ \beta 2 = (1/\text{Duration Hospital}) \times (\text{Critical Fraction})/(\text{Critical Fraction} + \text{Severe Fraction})) \\ \gamma 2 = (1/\text{Duration Hospital}) - \beta 2 \\ \mu = (\text{ICU Duration of Time}) \times (\text{CFR}/\text{Critical Fraction}) \\ \gamma 3 = (\text{ICU Duration of Time}) - \mu \\ \text{whereas CFR is the Case Fatality Rate.} \end{array}$

3.2 Parameter Estimation

The input data needed for the SEIRD model is time series data of prior dates from March 25, 2020 to June 16, 2021 and the population size(N=2200000), Initial infected cases were 1000.

A typical dataset for COVID-19 is made up of a daily number of four different types of cases, namely, cases that are confirmed, the number of recovered cases, the number of active cases and the number of cases that have deceased. The number of cases that are active is extracted from the data available by deducting the cases that have recovered and the cases that are deceased from the cases that are confirmed. The variables in the model have five different types of populations, namely, a population that is susceptible, a population that has been exposed, a population with an active infection, a population that has recovered from the infection, and finally, the population that is deceased. The unknown parameters of the model such asS0, $\beta 1$, $\beta 2$, $\beta 3$, α , γ , and μ are determined from the time series data.

In Figure 3, we have made use of logarithmic graphs. As we are aware that on a linear scale graph, the rate of growth keeps increasing and it is possible for the line to become almost vertical and seem to be tending to the Y axis. This could lead to the impression that mitigation measures such as social distancing and other public measures are not benefitting the population. In the meantime, on a logarithmic scale, numbers on the Y-axis do not increase in equal increments. Instead, each interval moves up by a set factor. As an example, 10 could be a factor of any of the following numbers such as, 30 or 450 or 4,500 etc.

4 Results and Discussion

The proposed computational model has been validated against real data for COVID-19 outbreak in Italy. The predictions of the proposed model were compared with actual data and found to be in very good agreement. The model was also applied to Indian scenario and the predictions were with error less than 10 percent for the span between March 2020 and January 2021. The comparative analysis of the proposed model was done with the stochastic model proposed in⁽¹⁸⁾ for forecasting the spread of Covid -19 pandemic in Ethiopia. It was found that both the models performed well in predicting the cases. However, the proposed model is computationally less expensive than the model in⁽¹⁸⁾ as it is deterministic one. Stochastic models require lot of parameter settings and data. Deterministic models are easy to interpret too.



Fig 3. Evolution of cases since outbreak in logarithmic graph

The model considered in this study is fairly simplistic and does not take into account important factors such as possible overloading of the healthcare system, mitigation efforts, behavioural responses to the epidemic, etc. There are more sophisticated and realistic epidemic models that may be able to predict the spread of COVID-19 and the long run number of deaths better than the model studied here. However, those models usually have even more parameters, so one may worry that their Identification would be even more troublesome. The model used in this study requires only three parameters to be estimated; whereas, the statistical model used in ⁽²²⁾ requires estimation of more than ten parameters. The novelty of this study lies in the frequency domain approach that used to solve the SEIRD equations. Use of frequency domain method transforms the integral differential equations into simple algebraic equations. This reduces the computational burden in forecasting the pandemic spread.

Model Simulation Results

The evolution of the variables in the model is calculated by solving the set of differential equations by using the clinical, transmission and intervention parameters are mentioned in Table 1 and Table 2.

			1 0		
Clinical Parameters	Case -1 I/P	Case -2 I/P	Transmission Parameters	Case -1 I/P	Case -2 I/P
Incubation Period	11	5	Mild infections	0.5/day	0.5/day
Mild infections	8	4	Severe infections	0.25/day	0.25/day
Severe infections	15	41	Critical infections	1/day	1/day
Critical infections	6	22			
Death rate critical infections	40	80			
Hospital stays severe infection	5	4			
ICU stay critical	8	3			

Table 1. Clinical and Transmission	Inputs given	to SEIRD Model
------------------------------------	--------------	----------------

Table 2. Intervention Inputs given to SEIRD Model						
Intervention Rate	I/P					
Reduction in transmission from mild infections	30%					
Reduction in transmission from severe infections	1%					
Reduction in transmission from critical infections	5%					

Figure 4 and Figure 7 illustrates predictive value of R0 that tell us the average number of people that a single sick individual infects an entirely susceptible population. Figure 5 shows SEIRD values on algorithmic graph. Figure 6 and Figure 8 shows intervention parameters on a simple and logarithmic scale.



Fig 4. redicted value of R₀



Fig 5. EIRD values are shown in logarithmic graph



Fig 6. imple Graph is Shown with and without intervention parameters.



Fig 8. ogarithmic Graph isshown with and without intervention parameters

5 Limitations

This model works well for a population that is homogeneous and does not migrate. It can be noted in two of the examples that initially the value of the conditions for the arrangement of standard differential is zero. Given such a case, the algorithm for optimization falls flat and the boundaries that have been assessed are equal to the initial values that were guessed. In some of the other cases there has been a second wave in the number of infected cases. In this case the improvement calculation gives unphysical results additionally.

6 Conclusion

The primary focus of this study is the proposal of a SEIRD model. Moreover, we have researched two varying infection rate functions that can be used for analysing the evolution of the COVID-19 outbreak⁽⁷⁾ for any given population N. It is our

firm belief that it is possible to adapt the proposed model to observe different infected cases, recovered cases and cases that are deceased during the various stages of the pandemic. It is for this reason that we have simulated the model by extending the incubation period from 5 days to 10-20 days. We then noted what effect that had on the duration of the simulated COVID-19 model. Moreover, making use of this model, we have been able to predict the value of R0. The proposed model is deterministic one and that is why it is computationally less expensive compared to existing stochastic models. Estimating the model parameters is a challenge in Covid-19 situation. The parameters are likely to vary on a day to day basis. The proposed deterministic model requires almost 50 % less parameters to be estimated compared to stochastic models that reduces the computational burden. Accuracy wise both deterministic and stochastic models perform equally well but deterministic models are computationally less expensive than stochastic models.

References

- 1) Li H, Liu SM, Yu XH, Tang SL, Tanga CK. Coronavirus disease 2019 (COVID-19): current status and future perspectives. Int J Antimicrob Agents . 2020;55(5):105951. Available from: https://doi.org/10.1016/j.ijantimicag.2020.105951.
- 2) World Health Organization. . Available from: https://www.who.int/health-topics/coronavirus#tab=tab_1.
- Chatterjee S, Sarkar A, Karmakar M, Chatterjee S, Paul R. SEIRD model to study the asymptomatic growth during COVID-19 pandemic in India. *Indian Journal of Physics*. 2021;95:2575–2587. Available from: https://doi.org/10.1007/s12648-020-01928-8.
- Mwalili S, Kimanthi M, Ojiambo V, Gathungu D, Mbogo RW. SEIR Model for COVID-19 Dynamics Incorporating the Environment and Social Distancing. BMC Res Notes. 2020;13(352). Available from: https://doi.org/10.1186/s13104-020-05192-1.
- 5) Cooper I, Mondal A, Antonopoulos CG. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons & Fractals.* 2020;139:110057. Available from: https://doi.org/10.1016/j.chaos.2020.110057.
- 6) Maherala'raj, Munirmajdalawieh N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infectious Disease Modelling*. 2021;6:98–111. Available from: https://doi.org/10.1016/j.idm.2020.11.007.
- 7) Alanazi MMSA, Kamruzzaman M, Alruwaili N, Alshammari SA, Alqahtani A, Karime. Measuring and Preventing COVID-19 Using the SIR Model and Machine Learning in Smart Health Care. *Journal of Healthcare Engineering*. 2020;8857346. Available from: https://doi.org/10.1155/2020/8857346.
- 8) Kim Y, Cho N. A Simulation Study on Spread of Disease and Control Measures in Closed Population Using ABM. *Computation*;10(1):2. Available from: https://doi.org/10.3390/computation10010002.
- 9) Squazzoni F, Polhill JG, Edmonds B, Ahrweiler P, Antosz P, Scholz G, et al. Computational Models That Matter During a Global Pandemic Outbreak: A Call to Action. Journal of Artificial Societies and Social Simulation . 2020;23(2):10. Available from: https://doi.org/10.18564/jasss.4298.
- Li XZZ, Gao S, Fu YKK, Martcheva M. Modeling and Research on an Immuno-Epidemiological Coupled System with Coinfection. Bulletin of Mathematical Biology. 2021;83(11):116. Available from: https://doi.org/10.1007/s11538-021-00946-9.
- Tuan NH, Tri VV, Baleanu D. Analysis of the fractional corona virus pandemic via deterministic modeling. *Mathematical Methods in the Applied Sciences*. 2021;44(1):1086–1102. Available from: https://doi.org/10.1002/mma.6814.
- 12) Lin Q, Chiu AP, Zhao S, He D. Modeling the spread of Middle East respiratory syndrome coronavirus in Saudi Arabia. *Statistical Methods in Medical Research*. 2018;27(7):1968–1978. Available from: https://doi.10.1177/0962280217746442.
- 13) Masum M, Masud MA, Adnan MI, Shahriar H, Kim S. Comparative study of a mathematical epidemic model, statistical modeling, and deep learning for COVID-19 forecasting and management. Socio-Economic Planning Sciences. 2022;80:101249. Available from: https://doi.org/10.1016/j.seps.2022.101249.
- 14) Chen Y, Li N, Lourenço J, Wang L, Cazelles B, Dong L, et al. Measuring the effects of COVID-19-related disruption on dengue transmission in southeast Asia and Latin America: a statistical modelling study. *The Lancet Infectious Diseases*. 2022;22(5):657–667. Available from: https://doi.org/10.1016/S1473-3099(22)00025-1.
- 15) Pichugin YA, Malafeyev OA, Zaitseva IV, Shulga AA, Kolesov DN. Forecast of Coronavirus Distribution in a Number of Countries of Europe and Asia: Dynamic and Stochastic Approach. *IOP Conference Series: Earth and Environmental Science*. 2021;666(3):032021. Available from: https://doi.org/10.1088/ 1755-1315/666/3/032021.
- 16) Bu SJ, Cho SB. Time Series Forecasting with Multi-Headed Attention-Based Deep Learning for Residential Energy Consumption. Energies. 2020;13(18):4722. Available from: https://doi.org/10.3390/en13184722.
- Harvey A, Kattuman P. Time Series Models Based on Growth Curves with Applications to Forecasting Coronavirus. *Harvard Data Science Review*. 2020;(S1). Available from: https://doi.org/10.1162/99608f92.828f40de.
- 18) Asmelash Y, Gebretensae, Asmelash D. Trend Analysis and Forecasting the Spread of COVID-19 Pandemic in Ethiopia Using Box- Jenkins Modeling Procedure". International Journal of General Medicine. 2021;14:1485–1498.
- 19) Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*. 2020;20(5):533–534. Available from: https://doi.org/10.1016/S1473-3099(20)30120-1.
- 20) JHU, Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. n.d. Johns Hopkins Coronavirus Resource Center (CRC). Available from: https://coronavirus.jhu.edu/map.html.
- 21) Korolev I. Identification and estimation of the SEIRD epidemic model for COVID-19. *Journal of Econometrics*. 2021;220(1):63–85. Available from: https://doi.org/10.1016/j.jeconom.2020.07.038.
- 22) Liu L, Moon HR, Schorfheide F. Panel forecasts of country-level Covid-19 infections. *Journal of Econometrics*. 2021;220(1):2–22. Available from: https://doi.org/10.1016/j.jeconom.2020.08.010.