Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Data Forecasting and Comparative Analysis Using SIR Model for COVID-19 Outbreak in Bangladesh and India

Md. Imran Zahid[1], Abdullah Ahmed Foisal[1], Md. Mahmud Alam[2*]

[1]Department of Mathematics, University of Barishal, Barishal-8200, Bangladesh
[2]Mathematics Discipline, Khulna University, Khulna, Bangladesh

DOI: 10.36347/sjpms.2023.v10i05.002 | **Received:** 19.12.2022 | **Accepted:** 24.01.2023 | **Published:** 21.06.2023

***Corresponding author:** Md. Mahmud Alam
Mathematics Discipline, Khulna University, Khulna, Bangladesh

Abstract | **Original Research Article**

The aim of this paper is to analyze data and forecasting of the novel coronavirus (Covid-19) using the SIR model in Bangladesh and India. The recovery rate, growth rate, and percentage of daily increment for both countries have been computed. The basic reproduction rate was found to be larger than 1 in both countries. This present study analyzes actual data to predict future susceptibility, recovery, and death till the end of the year. The total deaths, daily deaths, total cases, and daily growth of the two countries for the first wave, and second wave of the Covid outbreak has been compared and shown graphically.

Keywords: SIR Model, Bangladesh, India, Coronavirus (Covid-19), Data Analysis.

INTRODUCTION

The World Health Organization (WHO) proclaimed COVID-19 to be a pandemic on March 11, 2020 (Gupta *et al.*, 2021). Many health experts believe that the new strain of the coronavirus likely originated in bats or pangolins (Cyranoski, 2020). The first transmission to humans was in Wuhan, China. Since then, the virus has spread chiefly from person to person. Coronaviruses are a kind of virus that may infect both animals and humans. The novel coronavirus strain is known as severe acute respiratory syndrome (SARS-CoV-2) (Zhang *et al.*, 2020). Approximately 80% of persons with COVID-19 recover without the need for specialized care. These individuals may suffer flu-like symptoms. Coronaviruses are abundant in several animal species, such as cattle and camels. Although coronavirus transmission from animals to humans is uncommon, this new strain most likely originated in bats, but one study believes pangolins may be the source. However, it is unknown how the virus originally infected people (Kandola, 2020).

In this work, how the pandemic has evolved in Bangladesh and India have been examined. A Susceptible- Infected-Recovered (SIR) model is detailed below. The use of SIR models to study epidemics has been prevalent for decades (Cooper *et al.*, 2020). The Kermack-McKendrick model (Weiss Sir Ronald Ross, 2013) was proposed by Bjornstad *et al.*, 2002. This model is governed by differential equation which is clarifying the high accuracy of mathematical solution. Still, minor parameter variations can cause significant changes in the solution. Using SIR model, the future situation of covid-19 in Bangladesh and India has been predicted. Also, the covid situations in both countries for 1st wave as well as 2nd wave are analyzed.

Data Source

This research uses data from daily and cumulative covid-19 cases in Bangladesh and India. For both countries, the early covid period and delta period are considered. For Bangladesh, this data covers from May 2020 to October 2020 and March 2021 to August 2021. For India, this data covers from March 2020 to August 2020 and March 2021 to August 2021. The data for Bangladesh was obtained from Johns Hopkins University master data set, which is available at (JHU CSSE, 2020). The data for India was obtained from a trusted open-source website, available at (https://data.covid19india.org/). In the following part, a sample exploratory analysis of the incidence data is presented.

Bangladesh

After India, Bangladesh is the most affected country in South Asia. Bangladesh has joined the global coronavirus disease 2019 (COVID-19) pandemic,

which is being caused by severe acute respiratory syndrome coronavirus 2 (SARS CoV-2). In March 2020, the virus was discovered to have spread to Bangladesh. The country's epidemiology institute announced the first three known cases on March 8, 2020. Since then, the epidemic has spread over the whole country on a daily basis, and the number of those infected has grown. Figure 1 shows the monthly cumulative cases for Bangladesh over the entire sample period. On 18-June 2020, confirmed cases exceeded 100,000 (one lakh) per day. On 18-July 2020, confirmed cases exceeded 200,000 (two lakh) people per day. Towards the end of October 2020, the total confirmed cases crossed 11 million people. From May 2020 to October 2020, the total confirmed cases crossed 4 million. Although it increased from June 2020 to July 2020, the infection began to fall in early October 2020. In figure 2, the logarithm of cumulative incidence has a roughly linear trend, implying that cumulative incidence is increasing exponentially which are discovered. By the end of June 2020, the number of persons affected had increased twice over May 2020.

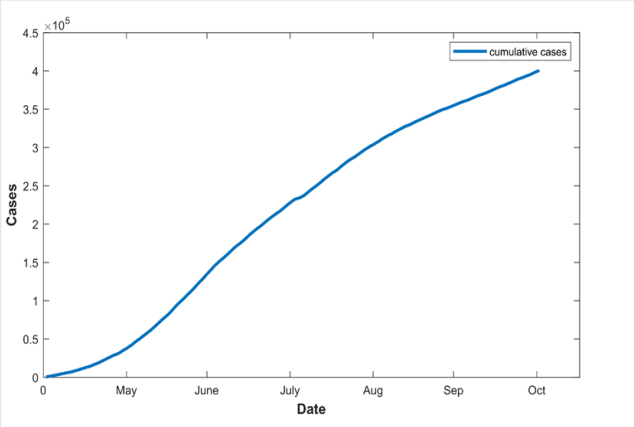

Figure 1. Daily cumulative incidence (linear scale) for the period of March to August, 2020.

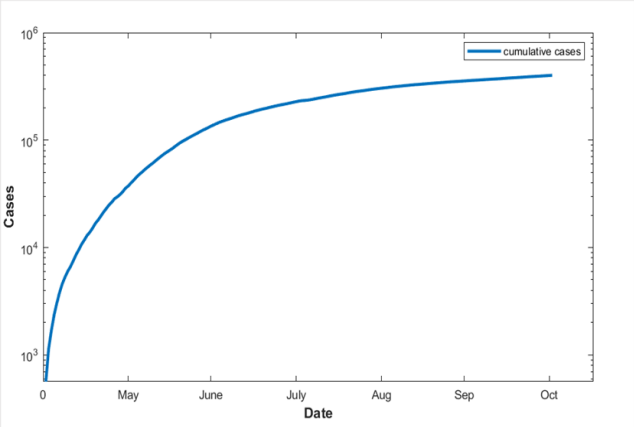

Figure 2. Daily cumulative incidence (log scale) for the period of May to October, 2020.

On May 8, 2021, the first Delta variation was discovered. On May 13, 2021 Bangladesh set a new record of 264 covid deaths in the 24 hours with 12,744 new corona virus infections cases were diagnosed (Rahman *et al.*, 2021). In the first week of June 2021, the Institute of Epidemiology, Disease Control, and Data (IEDCR) released research on genome-sequencing of 50 Covid cases since 16-May 2021, which revealed that 80 percent were infected with the Delta variant. People from Chapainawabganj, Khulna, Gopalganj, and Dhaka were sampled. After sequencing 60 Covid samples taken in Dhaka during the latter week of May and the first week of June, the International Centre for Diarrheal Disease Research, Bangladesh (ICDDR,b) found on June 17, 2021, that 68 percent of individuals who tested positive were infected with the Delta or Indian strain (Hasan *et al.*, 2021).

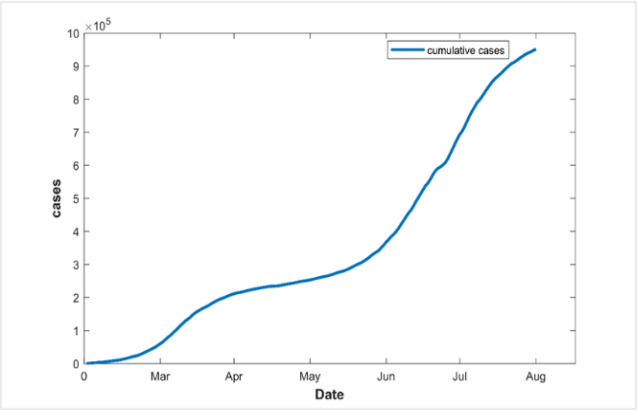

Figure 3. Daily cumulative incidence (linear scale) for the period of March to August, 2021.

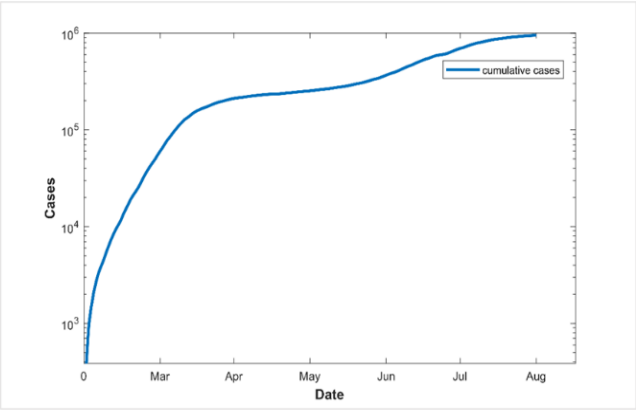

Figure 4. Daily cumulative incidence (log scale) for the period of March to August, 2021.

Figure 3 shows the monthly cumulative cases from March 2021, to August 2021. During this time, Bangladesh was heavily influenced by the Delta variant. During this time, the rate of coronavirus infection rises briefly and then falls, as shown in Fig.3. Later, in mid-May 2021, the covid started to grow rapidly. Figure 4 represents the daily cumulative cases on a log scale for Bangladesh. On April 15, 2021, the total cumulative incidences were 10 times compared to March 15, 2021.

India

On the 30 January, 2020, India reported its first incidence of COVID-19 in Thrissur, Kerala, which quickly grew to three cases by the 3rd of February, 2020 (Ram *et al.*, 2020); all were students returning from Wuhan. Apart from these, no significant rise in transmissions was observed in February 2020. On 4-March 2020, 22 additional cases were announced, with 14 afflicted members of an Italian tourist group among them. Transmissions rose throughout the course of the month after numerous persons with a history of travel to impacted countries and their connections tested positive. On March 12, 2020, a 76-year-old man with a history of travel to Saudi Arabia became India's first COVID-19 fatality. All commercial domestic and international flights were suspended in March 2020. Several cities and states announced that they would restrict public gatherings, dine-in restaurants or order the closure of various non-essential businesses through 31 March 2020, to slow the spread of COVID-19.

Figure 5 shows the monthly cumulative cases for India over the whole sample period. On May 18, 2020, of confirmed cases crossed 1,600,000 (sixteenth lakh) people. The total confirmed cases crossed 3.5 million people at the end of August. From July to August, the number of infected people increased rapidly. In August 2020, the total number of infected people was 1.5 million, but at the end of September 2020, this number crossed 3.5 million. Figure 6 shows that, at the end of August 2020, the number of incidences increased almost 10 times compared to June 2020.

In February 2021, daily cases had fallen to 9,000 per day. However, by early-April 2021, a significant second wave of infections took hold in the country with destructive consequences; on 9-April 2021, India surpassed 1 million active cases, and by 12-April 2021, India overtook Brazil as having the second-most COVID-19 cases worldwide. By late April 2021, India passed 2.5 million active cases and reported an average of 300,000 new cases and 2,000 deaths per day. Some analysts feared this was an undercount. In March 2021, a catastrophic second wave gripped the country.

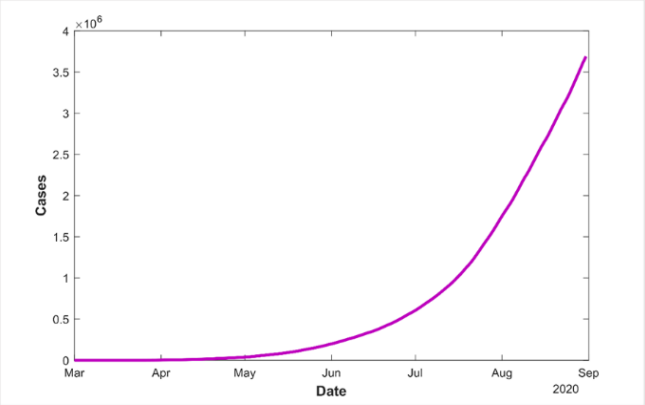

Figure 5. Daily cumulative incidence (linear scale) for the period of March to August, 2020.

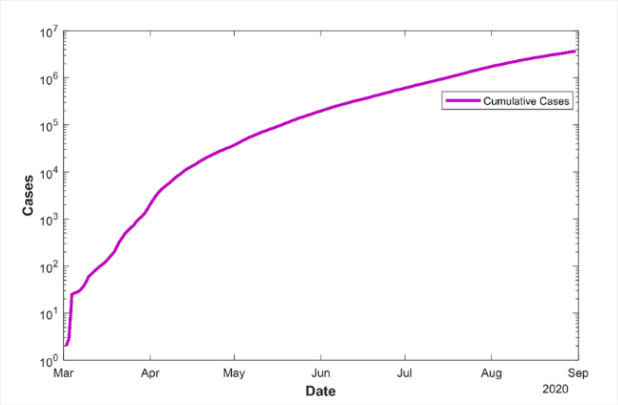

Figure 6. Daily cumulative incidence (log scale) for the period of March to August, 2020.

The situation was aggravated by a shortage of hospital beds, oxygen cylinders and medicines. On April 9, India surpassed Brazil as the world's second worst-affected country. By the end of April 2021, the nation had reported more than 400,000 cases every day. According to estimates, the second wave killed 2 lakh people. India records 4.14 lakh new cases in a day on May 7, 2021. Fig 7 shows the monthly cumulative cases (in linear scale) for India over the whole sample period (From March to September, 2021). On May 5, 2021, the total confirmed cases crossed 1 million. The end of June 2021 the total confirmed cases crossed 2.3 million people. From March 2021 to April 2021, the number of infected people increases rapidly. From figure 8, at mid-May, 2021 the total number of incidences increased almost 10 times compared to March 2021 are observed. After the July 2021, the cumulative cases increase slowly. This means that the second wave is probably end after few months.

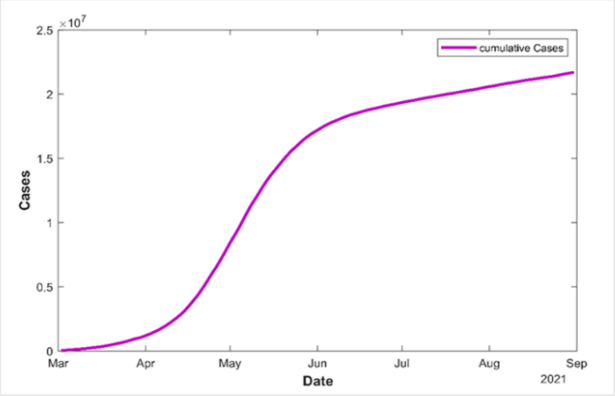

Figure 7. Daily cumulative incidence (liner scale) for the period of March to August, 2021.

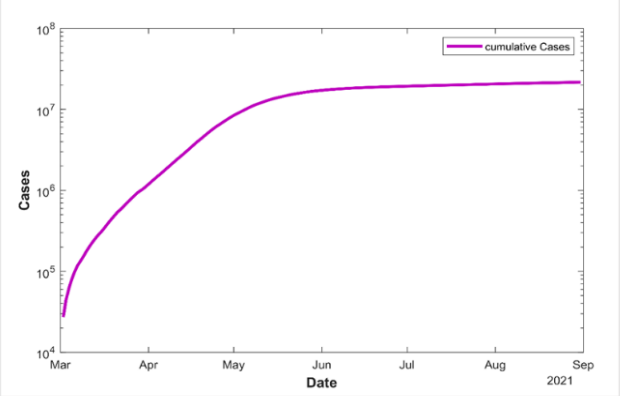

Figure 8. Daily cumulative incidence (log scale) for the period of March to August, 2021.

METHOD AND METHODOLOGY

The SIR model deals with only the number of susceptible, infected and recovered people.

Individuals, who are healthy but may develop the disease, susceptible individuals are referred to as susceptible. This class is commonly represented by S.

Infected people are those who have caught the disease and are now ill with it. The number of infected persons in a class is given by I.

Individuals, who have had the condition but are now healthy. The category of recovered individuals is commonly denoted by R.

Figure 9: SIR model flow chart.

The number of people in each of these groupings increases throughout time. That is, $S(t)$, $I(t)$ and $R(t)$ are functions of time t. The total population size N, satisfying $N = (t) + (t) + (t)$.

The governing equations of this model are

$$\frac{dS}{dt} = -\beta SI \quad (1)$$

$$\frac{dI}{dt} = \beta SI - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \gamma I \quad (3)$$

The model is based on two parameters: β (infection rate per unit time), and γ (recovery rate per unit of time). All three parameters have a positive constant (Chu, 2021).

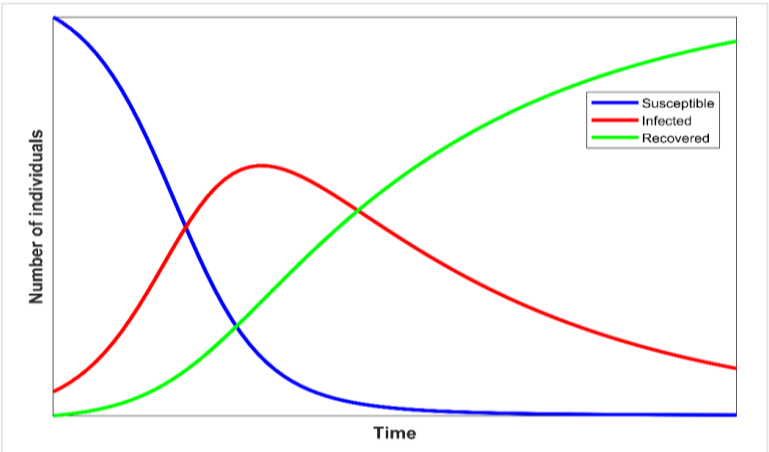

Figure 10: The trend of epidemic growth throughout time

The progress of the variables *S, I* and *R* have been shown in figure 10. Here the ranges are not given because they depend on each particular epidemic (Martínez *et al.*, 2021).

Governing Equations of the Model

The SIR model is based on a set of nonlinear equations. The model is governed by the equations (1)-(4), the formulation of the model described below.

"The parameter beta, infected rate per unit time and it is considered the product of two elements that is, $\beta = p \times y$, where p is the probability that contact with an infected person results in an infection and y is the number of contacts made by an infected person" (Martínez *et al.*, 2021). Therefore, we can say the susceptible people will be $pySI$. Putting the value of py for β, from equation. (1), the following equation has been obtained as:

$$\frac{dS}{dt} = \left\{\begin{matrix}\text{Variation of susceptible people per unit time} \\ \text{due to all infected people}\end{matrix}\right\} = -\beta SI$$

From the definition of γ, equation (3) has been obtained as;

$$\frac{dR}{dt} = \text{Variation in the number of retrieved individuals per unit time} = \gamma I$$

Therefore, from equation (2),

$$\frac{dI}{dt} = \text{Variation in the number of infected people per unit time} = \beta SI - \gamma I$$

From equation (1), the following equation can be written as follows;

$$dS = (-\beta SI)dt$$

This means that we can predict a slight change in the susceptible population for a small change in time. dS can be split as follows;

$$S_{i+1} - S_i = (-\beta S_i I_i)dt$$

Or,

$$S_{i+1} = S_i - (\beta S_i I_i)dt \qquad (4)$$

Proceeding in this way, Eq (2) and Eq (3) become respectively

$$I_{i+1} = I_i + (\beta S_i I_i - \gamma I_i)dt \qquad (5)$$
$$R_{i+1} = R_i + (\gamma I_i)dt \qquad (6)$$

The value of the parameters β (infection rate) and γ (recovery rate) for a given initial susceptible, infected and recover people can be calculated using equation (4) and equation (6).

RESULTS AND DISCUSSION

In Bangladesh, the confirmation of the first cases on March 8, 2020, the covid situation was typical (Haq *et al.*, 2022; Hassan *et al.*, 2020). However, within a month, the virus spreads widely, and Bangladesh becomes part of the first wave of covid-19. We take May 1st to June 30, 2020, daily cases of covid-19 from the Johns Hopkins University data set to establish the SIR model. We assume that only 500,000 of the total population have a chance of getting exposed to the virus and also consider$I_0 = 134{,}698$. Using equation (4) and equation (6) to find the value of β and γ. After that, using this value, we solve the system (1)-(3) with the help of MATLAB Ode Solver (ode45 function). Figure 11 represents the future situation of Covid-19 for Bangladesh till 31-December, 2020. The SIR model shows that the infected cases will decrease after the first week of August 2020. On the other hand, the number of recovered people increased rapidly from mid-July 2020.

The Delta variant, whose presence was first confirmed in Bangladesh on May 8, 2021 (Haq *et al.*, 2022; Hassan *et al.*, 2020). To predict future cases using the SIR model, we choose our data from March 1, 2021, to April 30, 2021. Using this raw data, we calculate our future cases and recover. We also assume that only 700,000 of the total population has a chance of getting exposed to the virus also consider$I_0 = 211{,}124$. After calculating the value of β and γ using equation (4) and equation (6), we solve the system (1)-(3) with the help of MATLAB Ode Solver (ode45 function). After that, we predict the future situation of covid-19 situation till December 31, 2021. Figure 12 represents the future situation of Covid-19 for Bangladesh till 31-December, 2021. According to our prediction, and the infected cases will start decreasing third week of the May, 2021.

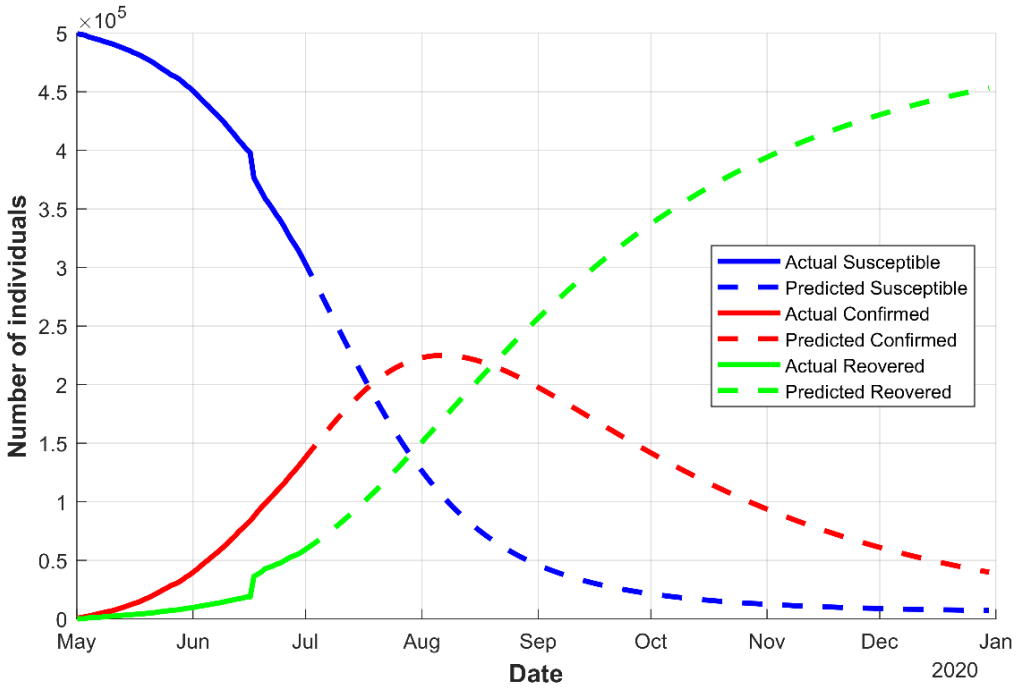

Figure 11: Plots of actual data and predicted estimated using the SIR model (for Bangladesh 2020)

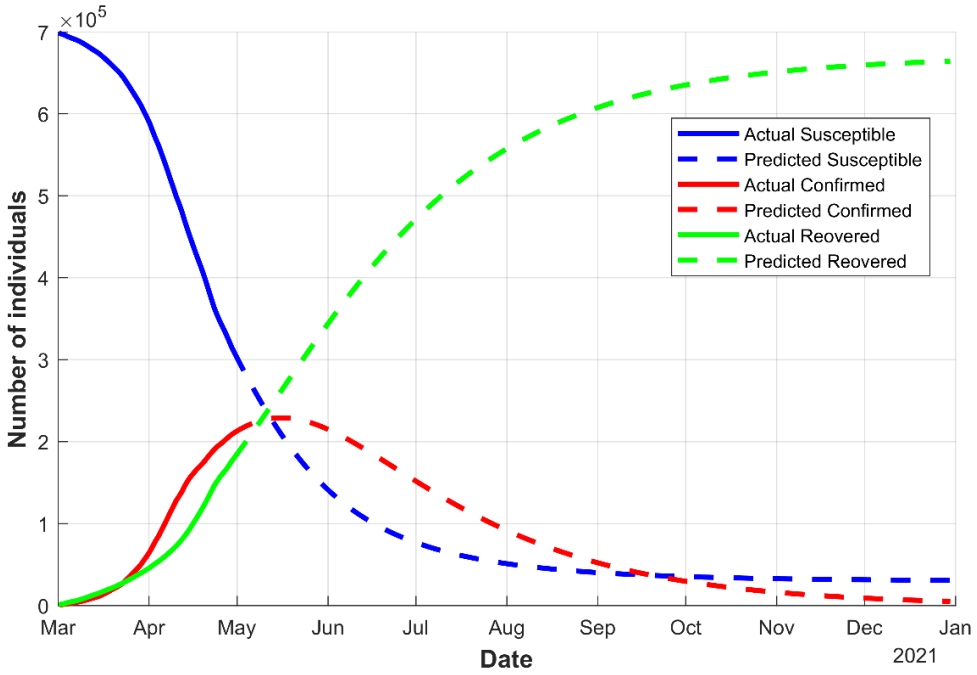

Figure 12: Plots of actual data and predicted estimated using the SIR model (for Bangladesh 2021)

On January 30, 2020, India announced its first COVID-19 case in Thrissur, Kerala. We selected raw data for India from March 1, to April 30, 2020. We assume that only 5,000,000 of the total population have a chance of getting exposed to the virus and also consider$I_0 = 34{,}863$. Using equation (4) and equation (6) calculate the value of β and γ. Then, using this value, we use MATLAB Ode Solver to solve the systems (1)-(3). (ode45 function). Figure 13 represents the future situation of Covid-19 for India till 31-December, 2020. According to the SIR model, infected cases would decline after the third week of August 2020. On the other side, the number of patients who have recovered has been quickly increasing from first week of August 2020.

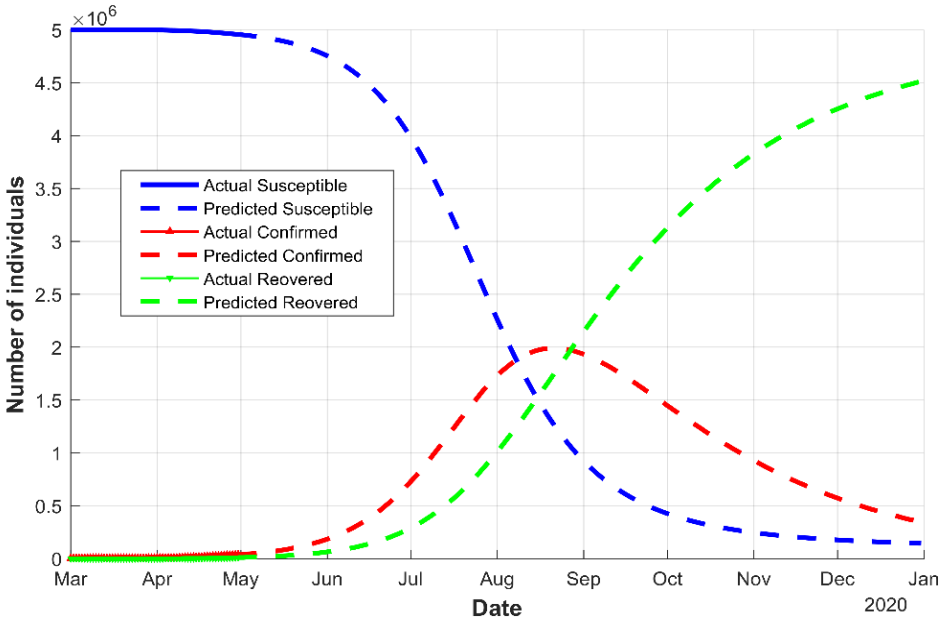

Figure 13: Plots of actual data and predicted estimated using the SIR model (for India 2020)

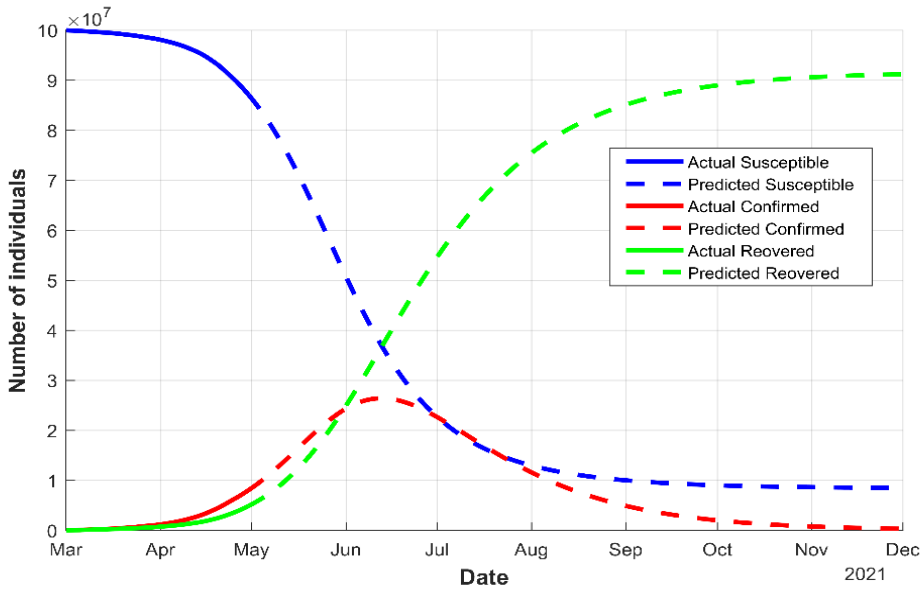

Figure 14: Plots of actual data and predicted estimated using the SIR model (for India 2021)

The Delta variant was first detected in India in October 2020. From our raw data, we choose from March 1 to April 30, 2021, daily covid cases of India, and build the SIR model. We assume 1,000,000,000 of the total population have a chance of getting exposed to the virus and also consider$I_0 = 8{,}045{,}001$. Figure 14 depicts the Covid-19 prospective position for India until December 31, 2020. Infected cases would begin to fall after mid-June, 2021 according to the SIR model.

India's daily cases peaked end of July 2020 with over 55,000 cases per day. On the other hand, Bangladesh confirmed its first cases on March 8, 2020. Infections were low until the end of March, but then began to climb dramatically in April 2020 (Hassan *et al.*, 2020). In the week ending on April 11, 2020, new cases in Bangladesh grew by 1,155 percent, the highest in Asia, ahead of Indonesia, with 186 percent. On May 6, 2020, cases were confirmed in all districts (Hopkins, 2020). Figure 15 represents the comparative data analysis of Bangladesh and India from May to September 2020. During this period, the Daily cases in Bangladesh were low compared to India. The total of cases in India at an exponential rate. But at the same time, Bangladesh faces a linear trend.

Daily cases were shallow compared to India's. The total of cases in India's at an exponential rate. But at the same time, Bangladesh faces a linear trend. India has a higher total death toll than Bangladesh. At first,

India's daily growth rate was high but slowed down. In Bangladesh, the opposite is happening. After July 2020, India's increment ratio falls below 2.

On the other hand, till the last week of August, the increment ratio of Bangladesh is about 3. From March to April 2021, India's and Bangladesh's daily growth rates are almost identical. From the 1st week of May to June, 2021, India's daily increment was high compared to Bangladesh. From the 2nd week of June 2021, India's increment ratio falls below 0.5%, but Bangladesh's daily increment ratio toll over 0.5% to 1%. The total deaths in India toll over 4.5 lakhs till August, where Bangladesh total deaths are below 50 thousand.

Figure 17 represents the Basic reproduction rate for both countries from May to September 2020. From May to July 2020, Bangladesh's basic reproduction rate is so high compared to India. From August 15 to September 2020, Bangladesh's and India's reproduction rates are almost the same.

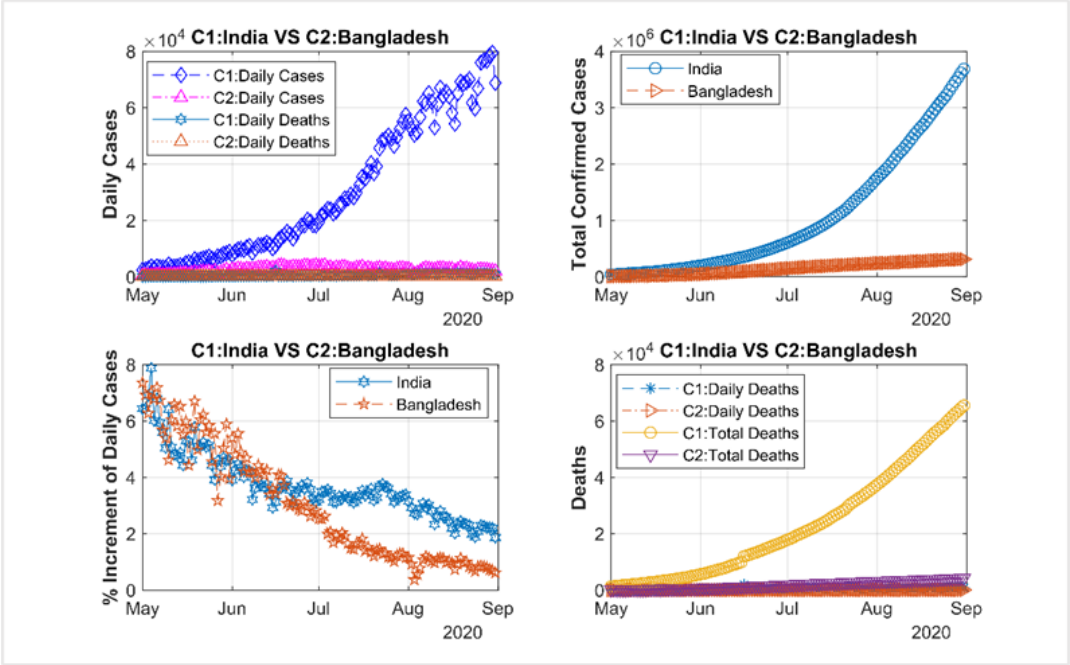

Figure 15: Data compassion between Bangladesh and India (2020)

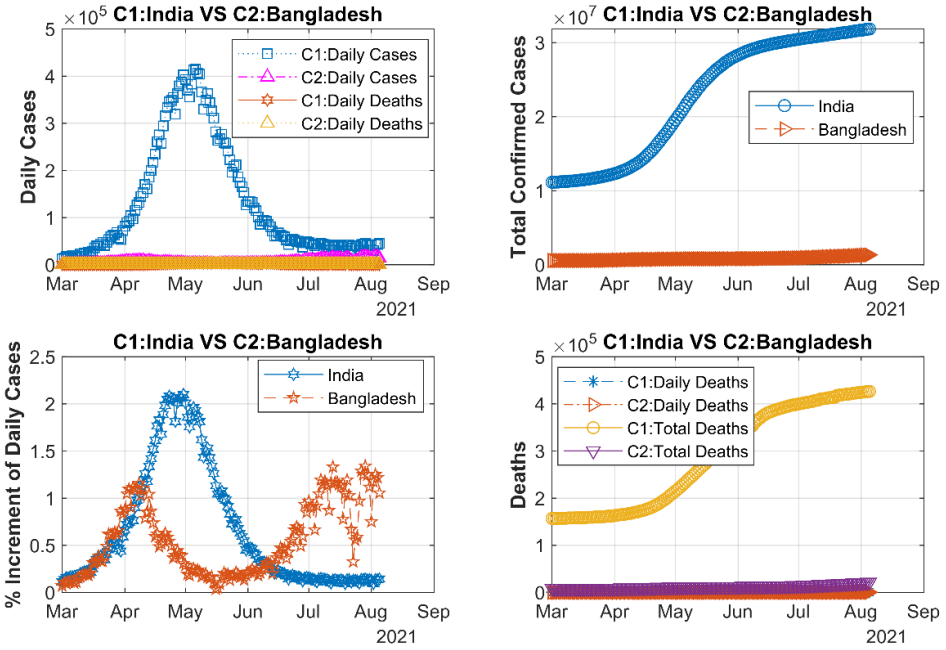

Figure 16: Data compassion between Bangladesh and India (2021)

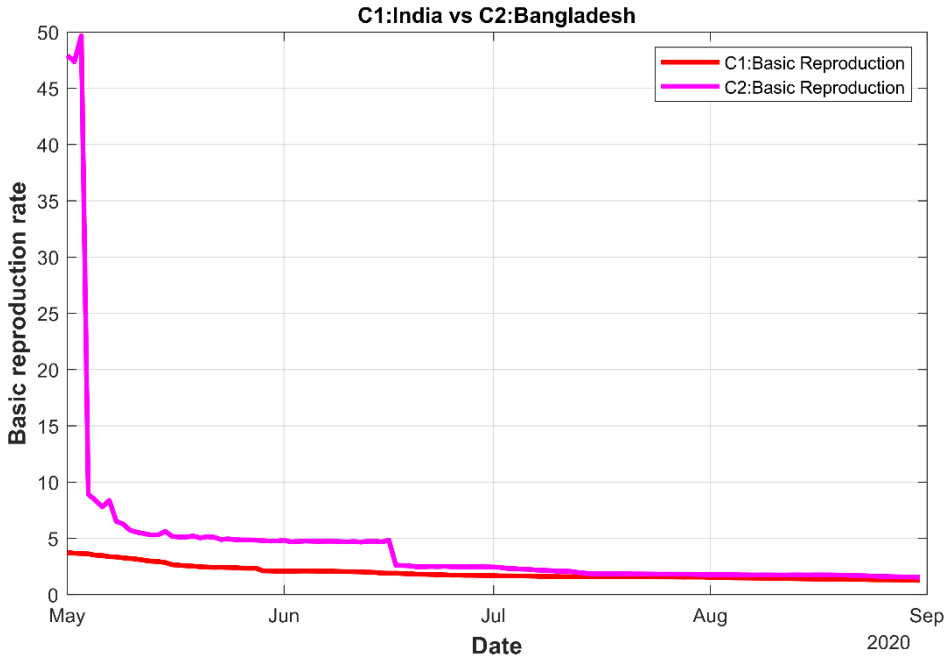

Figure 17: Basic reproduction rate for both country (2020)

Figure 18 presents the basic fertility rates of both countries from March to September 2021. From the 2nd week of March to April 2021, Bangladesh's basic fertility rate is higher than India's. From the end of April to June 2021, the fertility rate of India is higher than that of Bangladesh. Bangladesh had the highest fertility rate on April 15, 2021, 1.8, whereas India had 1.23 in May.

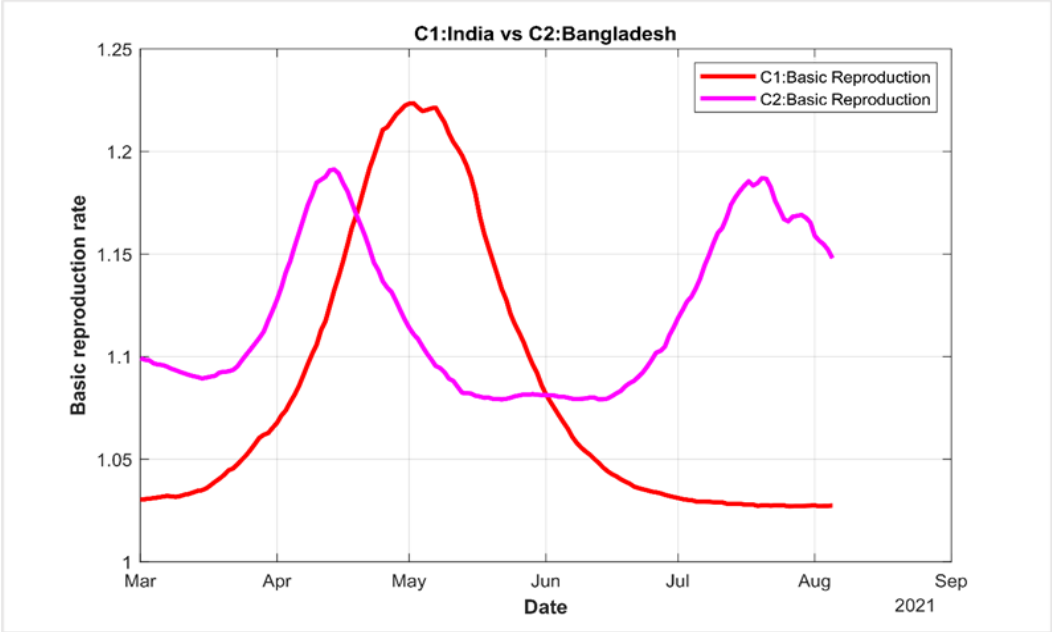

Figure 1: Basic reproduction rate for both countries (2021)

CONCLUSIONS

In Bangladesh and India, two of South Asia's worst-hit nations, a novel coronavirus (COVID-19) outbreak has been documented using a straightforward data analysis. Using data on the daily and cumulative incidence in both nations, the patterns, modeled the incidence, and projected the future incidence, recovery, and fatalities using one typical epidemiology method, the SIR model are examined.

REFERENCES

- Bjornstad, O. N., Finkenstadt, B. F., & Greenfell, B. T. (2002). Dynamics of Measles Epidemics: Estimating Scaling of. *Ecological Monographs*, *72*(2), 169–184.

- Chu, J. (2021). A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. In *PLoS ONE* (Vol. 16, Issue 3 March).
- Cooper, I., Mondal, A., & Antonopoulos, C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons and Fractals*, *139*, 110057.
- CSSE, J. (2020). GitHub - Novel Coronavirus (COVID-19) Cases, provided by JHU CSSE. In *Dataset COVID-19*. https://github.com/CSSEGISandData/COVID-19.
- Cyranoski, D. (2020). Mystery deepens over animal source of coronavirus. *Nature*, *579*(7797), 18–19.
- Gupta, N., Pandey, A., Goyal, A., & Dimri, K. (2021). Effect of COVID-19 on Access to Cancer Care in India. *South Asian Journal of Cancer*, *10*(01), 46–47.
- Haq, I., Hossain, M. I., Saleheen, A. A. S., Nayan, M. I. H., & Mila, M. S. (2022). Prediction of COVID-19 Pandemic in Bangladesh: Dual Application of Susceptible-Infective-Recovered (SIR) and Machine Learning Approach. *Interdisciplinary Perspectives on Infectious Diseases*, *2022*.
- Hasan, S. M. T., Das, S., Faruque, A. S. G., Khan, A. I., Clemens, J. D., & Ahmed, T. (2021). Taking care of a diarrhea epidemic in an urban hospital in bangladesh: Appraisal of putative causes, presentation, management, and deaths averted. *PLoS Neglected Tropical Diseases*, *15*(11), 1–23.
- Hassan, R., Dosar, A. S., Mondol, J. K., Khan, T. H., Noman, A. Al, Sayem, M. S., Hasan, M., & Juyena, N. S. (2020). Prediction of Epidemics Trend of COVID-19 in Bangladesh. *Frontiers in Public Health*, *8*(November 2020), 1–12.
- Hopkins, J. (2020). COVID-19 Map - Johns Hopkins Coronavirus Resource Center. In *Johns Hopkins Coronavirus Resource Center* (p. 1). https://coronavirus.jhu.edu/map.html.
- https://data.covid19india.org/
- Kandola, A. (2020). Coronavirus (COVID-19) origin: Cause and how it spreads. *Medical News Today*, *2*, 1–11.
- Martínez, V., Berihuete, Á., Sánchez-Sánchez, M., & Suárez-Llorens, A. (2021). A bayesian model of covid-19 cases based on the gompertz curve. *Mathematics*, *9*(4), 690–704.
- Rahman, M., Ahmed, A., Hossain, K., Haque, T., Islam, M., & Hussain, M. (2021). Impact of control strategies on COVID-19 pandemic and the SIR model based forecasting in Bangladesh. *Food Safety and Health*, *0*, 35.
- Ram, V. S., Babu, G. R., & Prabhakaran, D. (2020). COVID-19 pandemic in India. In *European Heart Journal* (Vol. 41, Issue 40, pp. 3874–3876).
- Weiss Sir Ronald Ross, H. (2013). The SIR model and the Foundations of Public Health. *MATerials MATemàtics Volum*, *17*(3), 1887–1097. www.mat.uab.cat/matmat
- Zhang, T., Wu, Q., & Zhang, Z. (2020). Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Current Biology*, *30*(7), 1346-1351.e2.