Using SIR Model and Recurrence Formula to Predict the Spread of COVID-19 in Sambalpur: A Mathematical Study

S. Kapoor and Roushnee Naik*

Department of Science and Mathematics, Regional Institute of Education (NCERT), Bhubaneswar, Odisha, India
E-mail: roushneenaik@gmail.com
**Corresponding Author*

Received 01 September 2021; Accepted 29 October 2021;
Publication 23 December 2021

Abstract

Corona Virus has spread across the globe and is creating havoc. Lockdown is being imposed worldwide depending on the number of cases. Everyone are advised to wear masks, follow social distancing, and use hand sanitizers to keep them safe. But all these precautions are not enough to curb the spread of the disease. People are still getting infected even after taking proper precautions and obeying the lockdown rule. We need to know in advance the approximate number of infected people so that we can devise better precautionary measures to curb the spread of the virus. So we use a simple SIR Model and solve it using basic differentiation and integration techniques and then use recurrence formula approach to predict the spread of COVID-19 in the city of Sambalpur of Odisha state. We compare the outcome of the model with the real time data and we arrive at the conclusion that the model is efficient in predicting the spread using the recurrence formula till the date 05/06/2021.

Keywords: COVID-19, SIR Model, recurrence formula, spread.

Journal of Graphic Era University, Vol. 9_2, 231–256.
doi: 10.13052/jgeu0975-1416.928

1 Introduction

1.1 History of COVID-19

According to Mayo Clinic [1] Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is the virus that causes COVID-19 (Coronavirus Disease 2019), the respiratory disease that caused the COVID-19 pandemic. The World Health Organization declared the outbreak a public health emergency of international concern on January 30, 2020 and declared a pandemic on March 11, 2020. As stated by the National Institutes of Health, it is the successor to SARS-CoV-1, which caused the SARS outbreak in 2002–2004.

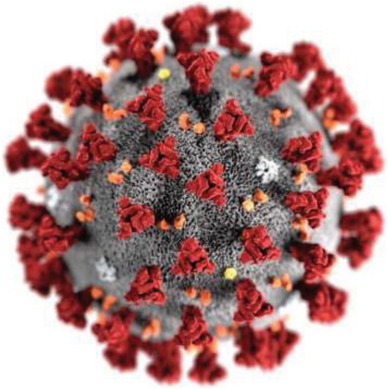

Figure 1 SARS-CoV-2.

The first known case was discovered in Wuhan, China, in December 2019. Since then, the disease has spread around the world, causing a pandemic. Symptoms vary, but usually include fever, cough, headache, fatigue, shortness of breath, and loss of smell and taste. Symptoms can begin 1 to 14 days after exposure to the virus. At least one-third of infected people will not have obvious symptoms.

According to WHO [2] the virus spreads when people come into contact with contaminated respiratory particles released by an infected person through breathing, talking, coughing, sneezing or singing. Particles vary in size, from tiny airborne particles that can be suspended in the air for hours to larger droplets that quickly fall to the ground. Transmission occurs when particles are inhaled or directly come in contact with the mucous membranes of the eyes, nose, or mouth (for example, by coughing) or when they are touched with contaminated hands. People can be contagious for up to 20 days, and they can spread the virus even if they do not show any symptoms.

Preventive measures include maintaining physical or social distance, isolation, ventilation of indoor spaces, covering the nose and mouth when coughing and sneezing, washing hands, and keeping unwashed hands away from the face. It is recommended to use masks or face shields in public places to minimize the risk of transmission. Several vaccines have been developed, and many countries have begun mass vaccination campaigns.

As with infection, there is a delay from when a person is first infected until the first symptoms appear. The average delay for COVID-19 is four to five days. Most people will develop symptoms 2 to 7 days after exposure, and almost all people will develop at least one symptom within 12 days.

It was estimated in September 2020 that, an infected person will infect two or three other people. This is more contagious than the flu, but not as good as measles. It usually spreads in clusters, where the infection can be traced back to indexed cases or geographical locations. The "super transmission incident" has a main function, that is, several people are infected by one person. Infected people can spread the virus to others for up to two days before they show symptoms, even if they never show up. People remain infectious for 7 to 12 days in moderate cases and up to two weeks in severe cases.

SARSCoV2 has spread in and out of China. Now it is spreading in the United States, India and the world, which means that people are reluctantly contracting and spreading the coronavirus. This increasing global spread has now become a pandemic. The first case of COVID-19 in Odisha was found in its capital Bhubaneswar on March 16, 2020, and has since spread throughout the state. The COVID-19 pandemic has spread to more than 180 countries. It has caused more than 300,000 deaths worldwide. It doesn't even have a cure. Although a vaccine has been developed, even if a person is vaccinated, he can still become infected with the disease and spread it among the susceptible ones.

Therefore, there must be a way to curb this disease. Many countries have declared national lockdown. It is recommended that people wear masks, wash their hands from time to time, and maintain social distancing. Although these methods help prevent infection, they are not fool proof. Better ways to stop this disease needs to be devised.

1.2 History of Sambalpur

Figure 2 Locating Sambalpur City in Odisha.

Sambalpur, a city of North-western Odisha State, eastern India. The District is surrounded by Deogarh District in the East, Bargarh District in the west, Jharsuguda District in the North and Sonepur and Angul District in the South. The District of Sambalpur has a history full of events including Indian freedom struggle representing the different section of the society. Sambalpur is the fifth largest city in the Indian state of Odisha covering a geographical area of 6702 sq. km. It is located in a lowland valley along the banks of the river Mahanadi. Sambalpur is one of the ancient place of India according to [3].

Figure 3 Map of Sambalpur.

Sambalpur is vulnerable to heat wave, flood, drought, fire accidents, factories hazards etc. for this reason some good preventive measures have been taken by the authorities. COVID-19 spread across the city of Sambalpur exponentially. On May 9, Sambalpur had figured among the top 10 districts in the country with the highest percentage of people testing positive for COVID-19. Lockdowns and Shutdowns were being imposed on various part of the city depending upon the number of daily cases.

1.3 Mathematical Models

To find ways to reduce the risk of infection, we must understand how quickly the disease spreads. We need a way to predict the spread. We still need to know when the pandemic ends.

Mathematicians have developed various methods to predict the spread of a disease. These methods are the mathematical models.

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modelling as stated in Wikipedia [4].

A model may help to explain a system and to study the effects of different components, and to make predictions about behaviour.

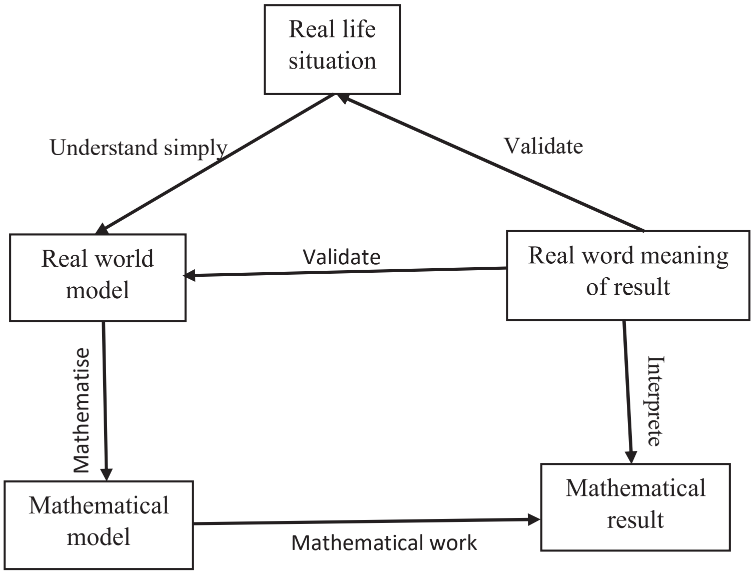

Figure 4 Flowchart depicting a mathematical model.

Mathematical models are extremely important in improving our understanding of population dynamics of infectious diseases. Models are a crucial tool to help with controlling and preventing the spread of disease based on scientific evidence.

Mathematical models can be of several types:

- Exponential model
- Logarithmic model
- SIR model
- Other variants of SIR model.

We will use a simple **SIR** model and use it to study the spread of COVID-19.

As COVID-19 is a flu like disease we assume that **SIR** model may be able to predict its spread.

2 Literature Review

Alexandria [5] said that mathematical models can be used as a mechanism or tool to help explain, describe, and predict the behaviour of physical phenomena. Epidemiological research is an area where models are particularly useful. There are many epidemiological mathematical models, from the simplest to the most complex. As scientists gain a better understanding of how disease spreads, they continue to work to create models that fit whatever problems they encounter.

The exponential model was popularized by Thomas Robert Malthus, says Alexei [6].

As said by Alexandria [5] the exponential model has an unrestricted pattern. It assumes that the growth rate is proportional to the population. This model is useful only to predict the early stages of disease transmission and is too simple for long term use. It can be used to predict the impact on the resources to be used. It does not predict the maximum infection that can occur.

The logistic model was introduced by Pierre Verhulst in 1838 who formulated it as a model of population growth by adjusting the model of exponential growth by Thomas Robert Malthus. Alexandria [5] states that the nature of the logarithmic model is very restrictive. The population tends towards a fixed k value; commonly referred to as carrying capacity. If the population is small with respect to k, then the pattern approximates that of the exponential model.

The logarithmic model has the sigmoid shape characteristics of many models predicting the spread of diseases but its parameters do not take into account the factors such as individuals who have recovered from the disease and have gained lifelong immunity or the randomness in infection. Also, to use the logarithmic model, it is necessary to know in advance what percentage of the population will be infected.

The SIR model, also known as the Kermack McKendrick model, is one of the earliest achievements of mathematical epidemiology created by William Kermack, Anderson McKendrick to model the spread of epidemics in larger and closed populations. It assumes that the population, N, is classified into three distinct groups of individuals, represented by the letters S, I, and R. The mathematical model of COVID-19 transmission constructed in this study is driven by research of Charlotte [7]. The model proposed in Charlotte [7] is built from the SIR model.

Anwar et al. [8] developed a Mathematical Model for Coronavirus Disease 2019 (COVID-19) containing isolation class. It was assumed that human to human contact is the potential cause for the COVID-19 outbreak. Therefore isolation of the infected human can reduce the risk of future COVID-19 spread. So, the total population was divided into five compartments:

1. Susceptible
2. Exposed
3. Infected
4. Isolated
5. Recovered

The susceptible population came in contact with the isolated and the infected population. The infected individuals, the individuals showing no symptoms apparently but have the disease in weak form inside their bodies, must be sent to isolated class in different rates.

The findings of Anwar et al. [8] shows that human to human contact is the potential cause of the outbreak of COVID-19. The coronavirus spreads through contact and the new infections induces the epidemic. So, isolation of the infected humans can overall reduce the risk of future COVID-19 spread.

The model developed by Giulia et al. [9] considers eight stages of infection:

1. Susceptible (S) – uninfected;
2. Infected (I) – asymptomatic or pauci-symptomatic infected;
3. Diagnosed (D) – asymptomatic infected, detected;
4. Ailing (A) – symptomatic infected, undetected;
5. Recognised (R) – (symptomatic infected, detected);

6. Threatened (T) – infected with life threatening symptoms, detected;
7. Healed (H) - recovered;
8. Extinct (E) – dead.

Collectively termed as SIDARTHE.

This model distinguishes between infected individuals depending on the severity of their symptoms. It compares simulation results with real data on the COVID-19 epidemic in Italy. It also models the possible scenarios of implementation of countermeasures.

It is a new mean field epidemiological model for the COVID-19 pandemic in Italy that extends the classical SIR Model.

It omits the probability rate of being susceptible again after recovering from the infection.

For the COVID-19 epidemic in Italy, it estimates the model parameters based on data from 20 February 2020 to 5th April 2020 and shows how the restrictions have affected the epidemic.

The SIDHARTHE dynamic system consists of eight ordinary differential equations, describing the evolution of the population over time. It was solved for the variables and the result was found that testing is important as it will help in detecting the asymptomatic individuals and preventing them from further spread. It predicts that had it not been for the preventive measures, more than 70,000 people could die in Italy in the first year. The outcomes suggested that the strict lockdown measures should continue for preventing spread of the disease.

The findings confirm that the adopted social distancing measures are necessary and effective, and should be promptly enforced at the earliest stage. Lockdown measures can only be relieved safely in the presence of widespread testing and contact tracing. Combining lockdown and population wide testing is key to rapidly ending the COVID-19 pandemic.

Pakwan et al. [10] proposed a model to analyse and to understand the transmission dynamics of the COVID-19 pandemic in Bangkok, Thailand. It is divided into seven compartments namely

1. Susceptible (S)
2. Exposed (E)
3. Symptomatically infected (I_s)
4. Asymptomatically infected (I_a)
5. Quarantined (Q)
6. Recovered (R)
7. Death (D)

The next- generation matrix approach was used to compute the basic reproduction number R_0 of the proposed model. The model's analysis and numerical results prove that the consistent use of face masks would go a long way in reducing the COVID-19 pandemic.

Saad et al. [11] proposed a model for predicting the spread of COVID-19 using the SIR Model and machine learning for the smart healthcare and the well-being of the citizens of KSA. It forecasts the situation for the upcoming 700 days. Mathematical analysis and simulation results were presented here to forecast the progress of outbreak and its possible end for the 3 situations;

i. no actions
ii. lockdown
iii. new medicine

The lockdown case delays the peak point by decreasing the number of infection. New medicines have a significant impact on infected curve by decreasing the number of infected people over time. The simulations predict that the highest cases might occur between 15 and 30 November 2020. It also suggests that the virus might be fully under control only after June 2021. Its reproductive rate shows that measures such as government lockdowns and isolation of individuals are not enough to stop the pandemic. So the authorities as soon as possible, should apply a strict long-term containment strategy to reduce the epidemic size successfully.

Abdy et al. [12] present a SIR Model for COVID-19 with fuzzy parameters. They have constructed the model using factors of vaccination, treatment, obedience in implementing health protocols, and the corona virus loads. Parameters of the infecting rate, recovery rate, and the death rate due to COVID-19 are constructed as a fuzzy number, and their membership functions are used in the model as fuzzy parameters. The model analysis uses the generation matrix method to obtain the basic reproduction number and the stability of the model's equilibrium points. Simulation results shows that differences in coronavirus-loads will also cause differences in transmission of COVID-19. Likewise, the factors of vaccination and obedience in implementing health protocols have the same effect in slowing or stopping the transmission of COVID-19 in Indonesia.

Omar et al. [13], suggested a new mathematical model describing the evolution of the COVID-19 in countries subject to the state of a health emergency. Based on the classical compartment models, two new compartments are added to the classical SIR model, representing the people who have benefitted from the awareness programmes and they could not respect the national quarantine for their personal reasons, called the partially controlled people. Second are people who know about the pandemic and respect the lockdown by staying at home, called the totally controlled people. The impact of these population on the severity of the COVID-19 infection was investigated. Using real COVID-19 propagation data they have estimated the totally and partially controlled populations. The findings are that the number of totally controlled people rose slowly from the beginning of the national lockdown. The population at risk in Morocco is more than 10 million in the first month of the closure and they need an efficient control strategy to save more lives.

3 Problem Formulation

3.1 Objectives

In Literature we saw that many advanced version of SIR Model was able to predict to some extent, the spread of COVID-19 across different regions. So, let us see if the simple SIR Model is effective in predicting the spread of COVID-19.

Before we proceed to develop the simple **SIR** Model, we state the following objectives for the model

- Prediction of number of daily cases.
- To check the effectiveness of the existing preventive measures.
- Predicting the end of the Pandemic.
- Effectiveness of the simple **SIR** Model in predicting the spread of COVID-19 in Sambalpur.

We will use the simple SIR model as developed by Charlotte et al. [7] to find an equation which we will convert into a recursive formula and use the formula recursively on the population of Sambalpur to predict the spread of COVID-19 in Sambalpur.

3.2 SIR Model

Assumptions of SIR model:

- Total size of population is assumed to be constant i.e. no immigration or emigration occurs.
- The population is not initially immune to the disease.
- Once a person is immune, he won't again get the disease.
- Individuals within the population follow the law of mass action.

- The deaths in the population occurs only due to COVID-19. No natural deaths or deaths due to any other disease occurs.
- The number of susceptible individuals in a particular day depends on the number of susceptible individuals of the previous day.

3.2.1 Law of mass action

Law of mass action states that the rate of chemical reaction is directly proportional to the product of activities or the concentration of the reactants.

In this scenario, the people act as the reactants. Any infected person in random can come in contact with any healthy person and transfer the disease. It assumes that the transmission of the disease is directly proportional to the population size.

3.2.2 Formulation

The total population of Sambalpur is divided into 3 parts.

1. **Susceptible** ($\boldsymbol{S}$) – The part of population that hasn't got the disease but can get the disease.
2. **Infected** ($\boldsymbol{I}$) – The part of the population that is infected.
3. **Removed** ($\boldsymbol{R}$) – The part of the population that has been removed i.e. recovered from the disease or has died from the disease.

It is assumed that at any point of time:

$$\boldsymbol{S} + \boldsymbol{I} + \boldsymbol{R} = \textit{constant} \ = \boldsymbol{N} = 1041099$$

Differentiating both sides, we get

$$\rightarrow \frac{d\boldsymbol{S}}{dt} + \frac{d\boldsymbol{I}}{dt} + \frac{d\boldsymbol{R}}{dt} = 0 = \frac{d\boldsymbol{N}}{dt}$$

where t = time.

Setting the Initial Conditions

$$\boldsymbol{S}(\mathbf{0}) = 1013314 = \boldsymbol{S}_{\mathbf{0}}$$

$$\boldsymbol{I}(\mathbf{0}) = 1946 = \boldsymbol{I}_{\mathbf{0}}$$

$$\boldsymbol{R}(\mathbf{0}) = 25750 = \boldsymbol{R}_{\mathbf{0}}$$

New infections occur as a result of contact between the infected ones and the susceptible ones. So when a new infection occurs, the infected individuals move from susceptible to infected category. In this simple model there is no way individuals can enter or leave the population. So we get the first differential equation:

$$\frac{d\boldsymbol{S}}{d\boldsymbol{t}} = -r\boldsymbol{S}(\boldsymbol{t})\boldsymbol{I}(\boldsymbol{t})$$

where,
$r =$ disease transmission rate.

Also the Infected individuals moved from infected to removed group. This happens at the rate $\sigma\boldsymbol{I}(\boldsymbol{t})$.

So we get the following 2 differential equations as well:

$$\frac{d\boldsymbol{I}}{d\boldsymbol{t}} = r\boldsymbol{S}(\boldsymbol{t})\boldsymbol{I}(\boldsymbol{t}) - \sigma\boldsymbol{I}(\boldsymbol{t})$$

$$\frac{d\boldsymbol{R}}{d\boldsymbol{t}} = \sigma\boldsymbol{I}(\boldsymbol{t})$$

where, $\sigma =$ rate of removal (recovery + death)

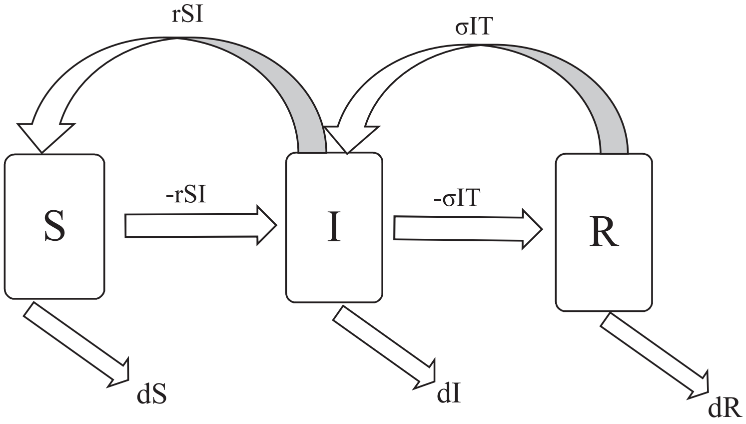

Figure 5 Diagram depicting the transfer of individuals in between the 3 compartments in the SIR Model.

So we get the following system of linear equations:

$$\frac{d\boldsymbol{S}}{d\boldsymbol{t}} = -r\boldsymbol{S}(\boldsymbol{t})\boldsymbol{I}(\boldsymbol{t}) \tag{1}$$

$$\frac{d\boldsymbol{I}}{d\boldsymbol{t}} = r\boldsymbol{S}(\boldsymbol{t})\boldsymbol{I}(\boldsymbol{t}) - \sigma\boldsymbol{I}(\boldsymbol{t}) \tag{2}$$

$$\frac{d\boldsymbol{R}}{d\boldsymbol{t}} = \sigma\boldsymbol{I}(\boldsymbol{t}) \tag{3}$$

We assume that $\sigma > 0$ and $r > 0$.

Also, it is evident that $\boldsymbol{S_0} > 0$ and $\boldsymbol{I_0} > 0$, else disease wouldn't have spread.

Dividing (3) from (1);

$$\frac{d\boldsymbol{S}}{d\boldsymbol{R}} = -\frac{r\boldsymbol{SI}}{\sigma I}$$

$$\rightarrow \frac{d\boldsymbol{S}}{d\boldsymbol{R}} = -\frac{r\boldsymbol{S}}{\sigma}$$

Let us assume $\rho = \frac{\sigma}{r}$, we get:

$$\rightarrow \frac{d\boldsymbol{S}}{d\boldsymbol{R}} = -\frac{\boldsymbol{S}}{\rho}$$

$$\rightarrow \frac{d\boldsymbol{S}}{\boldsymbol{S}} = -\frac{d\boldsymbol{R}}{\rho}$$

Integrating both sides, we get:

$$\rightarrow \int \frac{d\boldsymbol{S}}{\boldsymbol{S}} = \int -\frac{d\boldsymbol{R}}{\rho}$$

$$\rightarrow \int \frac{d\boldsymbol{S}}{\boldsymbol{S}} = -\int \frac{d\boldsymbol{R}}{\rho}$$

$$\rightarrow ln\boldsymbol{S} = -\frac{\boldsymbol{R}}{\rho} + C \tag{4}$$

At $t = 0$,

$$\rightarrow ln\boldsymbol{S_0} = -\frac{\boldsymbol{R_0}}{\rho} + C$$

$$\rightarrow C = ln\boldsymbol{S_0} + \frac{\boldsymbol{R_0}}{\rho} \tag{5}$$

Using (5) in (4):

$$\rightarrow ln\boldsymbol{S} = -\frac{\boldsymbol{R}}{\rho} + ln\boldsymbol{S_0} + \frac{\boldsymbol{R_0}}{\rho}$$

$$\rightarrow ln\boldsymbol{S} - ln\boldsymbol{S_0} = \frac{(\boldsymbol{R_0} - \boldsymbol{R})}{\rho}$$

$$\rightarrow \ln\left(\frac{\boldsymbol{S}}{\boldsymbol{S_0}}\right) = \frac{(\boldsymbol{R_0} - \boldsymbol{R})}{\rho}$$

$$\rightarrow \frac{\boldsymbol{S}}{\boldsymbol{S_0}} = e^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{R}}{\rho}\right)}$$

where S and R are the number of susceptible and removed individuals respectively.

We can rewrite the equation as follows:

$$\rightarrow \frac{\boldsymbol{S_\infty}}{\boldsymbol{S_0}} = e^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{R_\infty}}{\rho}\right)} \tag{6}$$

where,

$\mathbf{S}_\infty =$ no. of susceptible individuals at t $= \infty$
$\mathbf{I}_\infty =$ no. of infected individuals at t $= \infty$
$\mathbf{R}_\infty =$ no. of removed individuals at t $= \infty$

at t $= \infty$, we assume that the pandemic ends, so $\mathbf{I}_\infty = 0$

$$\rightarrow \mathbf{R}_\infty = \mathbf{N} - \mathbf{S}_\infty \tag{7}$$

Using (7) in (6):

$$\rightarrow \frac{\boldsymbol{S_\infty}}{\boldsymbol{S_0}} = e^{\left[\frac{\boldsymbol{R_0}-(\boldsymbol{N}-\boldsymbol{S_\infty})}{\rho}\right]}$$

$$\rightarrow \frac{\boldsymbol{S_\infty}}{\boldsymbol{S_0}} = \mathrm{e}^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{N}+\boldsymbol{S_\infty}}{\rho}\right)}$$

$$\rightarrow \boldsymbol{S_\infty} = \boldsymbol{S_0}\left[e^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{N}+\boldsymbol{S_\infty}}{\rho}\right)}\right]$$

We can modify the above equation and write it as a recursive relation:

$$\rightarrow \boldsymbol{S_{n+1}} = \boldsymbol{S_0}\left[e^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{N}+\boldsymbol{Sn}}{\rho}\right)}\right] \tag{8}$$

where,

$\boldsymbol{S_n} =$ number of susceptible individuals of day n
$\boldsymbol{S_{n+1}} =$ number of susceptible individuals of day $n+1$

Equation (8) is justified as we have assumed that the number of susceptible individuals in a day depends on the number of susceptible individuals of the previous day.

As per census, Sambalpur district has a population of 1,041,099.

$\rightarrow$ So we assume $\boldsymbol{N} = 1041099$

According to the district dashboard of Sambalpur [14] we have the following data:

Table 1 Actual data of infected, recovered and deaths

Date	Total Infected (**I**)	Total Recovered	Total Deaths
01-06-2021	1946	25750	89
02-06-2021	1813	26067	91
03-06-2021	1711	26368	91
04-06-2021	1553	26686	92
05-06-2021	1372	27003	93
06-06-2021	1176	27306	93
07-06-2021	1024	27559	95
08-06-2021	990	27685	95
09-06-2021	930	27875	95
10-06-2021	837	28068	97
11-06-2021	742	28228	98
12-06-2021	646	28365	98
13-06-2021	600	28472	98
14-06-2021	551	28574	99
15-06-2021	488	28674	100
16-06-2021	435	28765	101
17-06-2021	341	28896	101
18-06-2021	320	28961	102
19-06-2021	311	29002	102
20-06-2021	281	29063	102
21-06-2021	257	29115	104
22-06-2021	247	29150	107
23-06-2021	247	29185	111
24-06-2021	245	29220	113
25-06-2021	275	29220	113

We have assumed that

$$\text{Removed population } (\boldsymbol{R}) = \text{total deaths} + \text{total recovered} \tag{9}$$

Also we have taken $\boldsymbol{N} = \boldsymbol{S} + \boldsymbol{I} + \boldsymbol{R}$

$$\rightarrow \boldsymbol{S} = \boldsymbol{N} - \boldsymbol{I} - \boldsymbol{R} \tag{10}$$

Using (9) and (10) in Table 1 we get the values of Removed and Susceptible individuals that is listed in the following table:

Table 2 Actual data of susceptible

Date	Total Susceptible (S)	Total Infected (I)	Total Removed (R)
01-06-2021	1013314	1946	25839
02-06-2021	1013128	1813	26158
03-06-2021	1012929	1711	26459
04-06-2021	1012768	1553	26778
05-06-2021	1012631	1372	27096
06-06-2021	1012524	1176	27399
07-06-2021	1012421	1024	27654
08-06-2021	1012329	990	27780
09-06-2021	1012199	930	27970
10-06-2021	1012097	837	28165
11-06-2021	1012031	742	28326
12-06-2021	1011990	646	28463
13-06-2021	1011929	600	28570
14-06-2021	1011875	551	28673
15-06-2021	1011837	488	28774
16-06-2021	1011798	435	28866
17-06-2021	1011761	341	28997
18-06-2021	1011716	320	29063
19-06-2021	1011684	311	29104
20-06-2021	1011653	281	29165
21-06-2021	1011623	257	29219
22-06-2021	1011595	247	29257
23-06-2021	1011556	247	29296
24-06-2021	1011521	245	29333
25-06-2021	1011491	275	29333

4 Analysis and Finding

From the previous chapter we have seen,

$$\rightarrow \boldsymbol{S_{n+1}} = \boldsymbol{S_0}\left(e^{\left(\frac{\boldsymbol{R_0}-\boldsymbol{N}+\boldsymbol{S_n}}{\rho}\right)}\right)$$

Rearranging the term we get,

$$\rho = \frac{\boldsymbol{R_0} - \boldsymbol{N} + \boldsymbol{S_n}}{\ln\left(\frac{\boldsymbol{S_{n+1}}}{\boldsymbol{S_0}}\right)} \tag{11}$$

where,

$\boldsymbol{S_n}$ = number of susceptible individuals of day n
$\boldsymbol{S_{n+1}}$ = number of susceptible individuals of day $n+1$
$\rightarrow$ And $\boldsymbol{N} = 1041099$

Putting the values from Table 2 and the values of $\boldsymbol{R}_0$ and $\boldsymbol{S}_0$, we get the following values of ρ which is listed in the following table:

Table 3 Calculated data of ρ

Date	Total Susceptible	ρ
01-06-2021	1013314	
02-06-2021	1013128	1737728
03-06-2021	1012929	1631520
04-06-2021	1012768	1742208
05-06-2021	1012631	1864285
06-06-2021	1012524	2000194
07-06-2021	1012421	2058628
08-06-2021	1012329	1995824
09-06-2021	1012199	1935591
10-06-2021	1012097	1935540
11-06-2021	1012031	1962990
12-06-2021	1011990	2006947
13-06-2021	1011929	1996728
14-06-2021	1011875	1994227
15-06-2021	1011837	2012125
16-06-2021	1011798	2021772
17-06-2021	1011761	2058978
18-06-2021	1011716	2042771
19-06-2021	1011684	2028103
20-06-2021	1011653	2027405
21-06-2021	1011623	2023739
22-06-2021	1011595	2013129
23-06-2021	1011556	1990891
24-06-2021	1011521	1972887
25-06-2021	1011491	1940391

It is evident why we could not get the value of ρ for the date 01/06/2021.

So, we got values of ρ for 24 days. All the values of the ρ are almost equal. So we take the average of all the ρ.

$$\rho_{average} = 1958108$$

$$\rho = \frac{\sigma}{r}$$

r as calculated by a study is approximately 0.2

$$\rightarrow \sigma = \rho * \mathrm{r} = 391,621.6$$

Using this value of ρ_{average} in (11) we can predict the further spread i.e. we can find the values of $\mathbf{S}_n$ for other days, by using the Equation (11).

Using the recursive relation repetitively we get predict the values of S_n which is listed in the following table:

Table 4 Calculated data of S

Date	ρ	$\mathrm{S_n}$
01-06-2021	1958108	1013314
02-06-2021	1958108	1012307
03-06-2021	1958108	1011787
04-06-2021	1958108	1011518
05-06-2021	1958108	1011380
06-06-2021	1958108	1011308
07-06-2021	1958108	1011271
08-06-2021	1958108	1011252
09-06-2021	1958108	1011242
10-06-2021	1958108	1011237
11-06-2021	1958108	1011234
12-06-2021	1958108	1011233
13-06-2021	1958108	1011232
14-06-2021	1958108	1011232
15-06-2021	1958108	1011232
16-06-2021	1958108	1011231
17-06-2021	1958108	1011231
18-06-2021	1958108	1011231
19-06-2021	1958108	1011231
20-06-2021	1958108	1011231
21-06-2021	1958108	1011231
22-06-2021	1958108	1011231
23-06-2021	1958108	1011231
24-06-2021	1958108	1011231
25-06-2021	1958108	1011231
26-06-2021	1958108	1011231
27-06-2021	1958108	1011231
28-06-2021	1958108	1011231
29-06-2021	1958108	1011231
30-06-2021	1958108	1011231

We know that $\mathbf{S} + \mathbf{I} + \mathbf{R} = \mathbf{N}$

$$\rightarrow \mathbf{I} + \mathbf{R} = \mathbf{N} - \mathbf{S} \tag{12}$$

Using (13) in Table 4. We get the values of $\boldsymbol{I_n} + \boldsymbol{R_n}$

Table 5 Calculated data of $\mathrm{I_n} + \mathrm{R_n}$

Date	$\mathrm{S_n}$	$\mathrm{I_n}$ + $\mathrm{R_n}$
01-06-2021	1013314	27785
02-06-2021	1012307	28792
03-06-2021	1011787	29312
04-06-2021	1011518	29581
05-06-2021	1011380	29719
06-06-2021	1011308	29791
07-06-2021	1011271	29828
08-06-2021	1011252	29847
09-06-2021	1011242	29857
10-06-2021	1011237	29862
11-06-2021	1011234	29865
12-06-2021	1011233	29866
13-06-2021	1011232	29867
14-06-2021	1011232	29867
15-06-2021	1011232	29867
16-06-2021	1011231	29868
17-06-2021	1011231	29868
18-06-2021	1011231	29868
19-06-2021	1011231	29868
20-06-2021	1011231	29868
21-06-2021	1011231	29868
22-06-2021	1011231	29868
23-06-2021	1011231	29868
24-06-2021	1011231	29868
25-06-2021	1011231	29868
26-06-2021	1011231	29868
27-06-2021	1011231	29868
28-06-2021	1011231	29868
29-06-2021	1011231	29868
30-06-2021	1011231	29868

Comparing actual data with the data we got from the model we get the following table:

Table 6 Comparison of actual data with calculated data

Date	$\mathrm{S_n}$ (From Data)	$\mathrm{I_n + R_n}$ (From Data)	$\mathrm{S_n}$ (From SIR Model)	$\mathrm{I_n + R_n}$ (From SIR Model)
01-06-2021	1013314	27785	1013314	27785
02-06-2021	1013128	27971	1012307	28792
03-06-2021	1012929	28170	1011787	29312
04-06-2021	1012768	28331	1011518	29581
05-06-2021	1012631	28468	1011380	29719
06-06-2021	1012524	28575	1011308	29791
07-06-2021	1012421	28678	1011271	29828
08-06-2021	1012329	28770	1011252	29847
09-06-2021	1012199	28900	1011242	29857
10-06-2021	1012097	29002	1011237	29862
11-06-2021	1012031	29068	1011234	29865
12-06-2021	1011990	29109	1011233	29866
13-06-2021	1011929	29170	1011232	29867
14-06-2021	1011875	29224	1011232	29867
15-06-2021	1011837	29262	1011232	29867
16-06-2021	1011798	29301	1011231	29868
17-06-2021	1011761	29338	1011231	29868
18-06-2021	1011716	29383	1011231	29868
19-06-2021	1011684	29415	1011231	29868
20-06-2021	1011653	29446	1011231	29868
21-06-2021	1011623	29476	1011231	29868
22-06-2021	1011595	29504	1011231	29868
23-06-2021	1011556	29543	1011231	29868
24-06-2021	1011521	29578	1011231	29868
25-06-2021	1011491	29608	1011231	29868
26-06-2021	1011473	29626	1011231	29868
27-06-2021	1011445	29654	1011231	29868
28-06-2021	1011416	29683	1011231	29868
29-06-2021	1011393	29706	1011231	29868
30-06-2021	1011365	29734	1011231	29868

The above data was entered into Excel to give the following scatterplot.

In the following graph we compare the actual data and the data that we got from the simple SIR Model. While drawing the graph we can get a clear picture if our model was able to predict the spread of the virus.

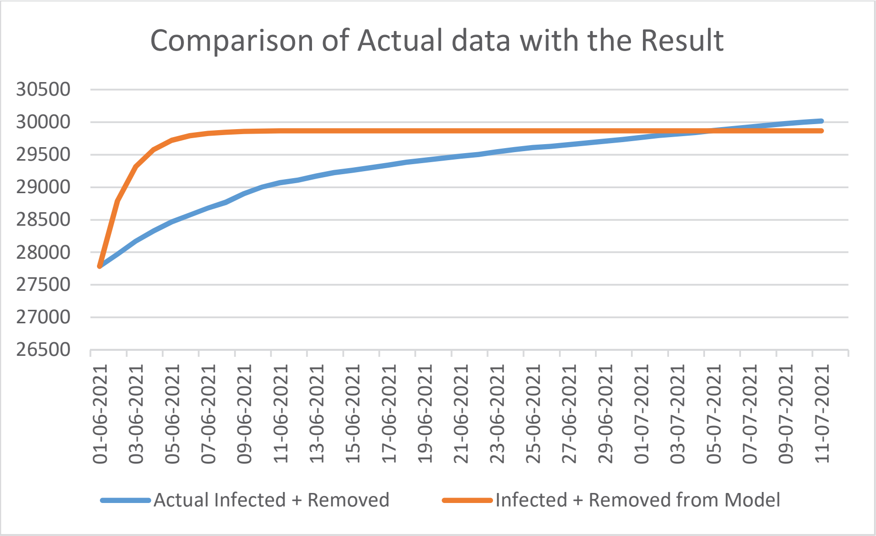

Figure 6 Comparison of the calculated **I** + **R** value with the actual **I** + **R** value.

Also a parameter called the basic reproduction number (β) determines if there will be a pandemic or not. It is the number of secondary infections that are caused by a single infected person during the course of their infection. The basic reproduction number is denoted by the quantity β, where

$$\begin{aligned} \beta &= Nr/\sigma \\ \rightarrow \beta &= N/\rho \end{aligned} \tag{13}$$

If $\beta > 1$, there will be an epidemic.
If $\beta < 1$, there won't be an epidemic.

Putting the values of N and ρ in (14) we get
$\beta = 0.53$, which is clearly < 1.

Which means an epidemic won't occur that is not actually true.

5 Conclusion

The goal of this paper was to develop a simple SIR Model and see if the model is effective to predict the spread of COVID-19 in Sambalpur. We first explained the method on how to develop the model, and then we used the real data to find the value of the relative removal rate (ρ). Then we used the value of ρ to predict the values of **S** and $\mathbf{I} + \mathbf{R}$ and then compare the values with the real data.

As we can see from the figure 6 we can conclude that

- SIR Model was approximately able to predict the spread till 05/06/2021.
- SIR Model predicted that the spread of the disease will stop after 05/06/2021.
- The peak of $I + R$ will reach on 05/06/2021.
- But as we can see from the graph that the cases still increased after 05/06/2021.

The SIR Model predicted that the number of $I + R$ will increase day by day until 05/06/2021. So, if we consider the data only till 05/06/2021 we can say that the model was able to predict the spread of COVID-19. The delay in attaining the peak may be due to the preventive measures taken by the Government such as imposing strict lockdown, wearing masks and following social distancing etc.

But if we consider the data after the date 05/06/2021, we can see that the prediction says that after 05/06/2021 there will be no more new cases. But from the actual data we can see that the number of cases continue to increase.

Also we saw that the value of the basic reproduction number suggested that the pandemic won't occur, but we can see that it is not the case.

So we can conclude that the simple SIR Model is efficient in predicting the spread of the COVID-19 till 05/06/21.

5.1 Limitations of the Model

- It could not predict the spread of the disease after 05/06/21.
- It cannot predict the value of **I** and **R** individually, which is a major setback of the simple model using recursive formula.
- It cannot predict the end of the pandemic correctly.

It may be because we have assumed a very simple Model.

We have assumed that the value of **N** is constant, which is not the case. As people are still migrating in and out of Sambalpur. So the value of **N** keeps changing.

5.2 Further Scope of Research

- Why the SIR model could not predict the spread after 05/06/2021 when the recurrence formula approach was used?
- What more parameters can be used or what changes can be done to the model so that it can predict the spread of the virus.

References

[1] COVID-19 (2021), Retrieved from https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963

[2] Coronavirus Disease (COVID-19), WHO, Retrieved from http://www.who.int/health-topics/coronavirus#tab=tab_1

[3] About Sambalpur District (2021), Retrieved from https://sambalpur.nic.in/about-district/

[4] Wikipedia (2021), Mathematical Model, Retrieved from https://en.wikipedia.org/wiki/Mathematical_model

[5] Le Beau, Alexandria, "Modeling the spread of Measles" (2020), *Electronic Thesis, Projects and Dissertations*. 1110, California. https://scholarworks.lib.csusb.edu/etd/1110

[6] Alexei Sharov (1997), Exponential Model, Retrieved from http://alexei.nfshot.com/PopEcol/lec5/exp.html

[7] Charlotte Alger, Kaitlin Todd (2015), The SIR Model of Disease Spread, Boston, Retrieved from http://web.simmons.edu/~grigorya/390/projects/Charlotte-Kaitlin_Epidemics.pdf

[8] Anwar Zeb, Ebraheem Alzahrani, Vedat Suat Erturk, Gul Zaman, "Mathematical Model for Coronavirus Disease 2019 (COVID-19) Containing Isolation Class", *Biomed Research International,* vol. 2020, Article ID 3452402, 2020 https://doi.org/10.1155/2020/3452402

[9] Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Flippo, Angela Di Matteo and Marta Colaneri , "Modelling the COVID-19 epidemic and implementation of population wide interventions in Italy", *Nature Medicine*, vol. 26, pp. 850–860 (2020), Italy https://doi.org/10.1038/s41591-020-0883-7

[10] Pakwan Riyapan, Sherif Eneye Shuaib, and Arthit Intarasit, "A Mathematical model of COVID-19 Pandemic: a case study of Bangkok, Thailand", *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6664483, 2021. https://doi.org/10.1155/2021/6664483

[11] Saad Awadh Alanazi, M. M. Kamruzzzaman, Madallah Alruwaili, Nasser Alshammari, Salman Ali Alqahtani, and Ali Karime , "Measuring and preventing COVID-19 using the SIR Model and Machine Learning in Smart Health Care", *Journal of Healthcare Engineering*. Vol. 2020, Article ID 8857346, 2020, Saudi Arabia http://doi.org/10.1155/2020/8857346

[12] Muhammad Abdy, Syfruddin Side, Suwardi Annas, Wahyuddin Nur and Wahidah Sanusi (2021), "An SIR Epidemic Model for COVID-19 spread with fuzzy parameter: the case of Indonesia", *Advances in Difference Equations*, vol. 2021, Article ID s13662-021-03263-6, 2021, Indonesia https://doi.org/10.1186/s13662-021-03263-6

[13] Omar Zakary, Sara Bidah, Mostafa Rachik, Hanane Ferjouchia, "Mathematical Model to Estimate and Predict the COVID-19 Infections in Morocco: Optimal Control Strategy", *Journal of Applied mathematics*, vol. 2020, Article ID 9813926, 2020, Morocco http://doi.org/10.1155/2020/9813926

[14] District Dashboard (2021), COVID-19, Odisha, Retrieved from https://statedashboard.odisha.gov.in/Pages/district

Biographies

S. Kapoor is an Assistant Professor of Mathematics at the Regional Institute of Education, Bhubaneswar (A Unit of National Council of Educational Research and Training, Delhi). He completed his Ph. D in Mathematics at IIT Roorkee.

Dr Kapoor is presently working in the fields of Computational Fluid Dynamics, Applied Numerical Method, Hydrodynamics Stability in Porous Media.

Roushnee Naik is a student of Mathematics. She has completed her B.Sc.(Mathematics) B.Ed from Regional Institute of Education Bhubaneswar. She is currently pursuing M.Sc.(Mathematics) at Govt. Autonomous College, Rourkela.