



Anton de Kom Universiteit van Suriname

Bibliotheek

Universiteitscomplex, Leysweg 86, Paramaribo, Suriname, Postbus 9212
Telefoon (597)464547, Fax (597)434211, E-mail: adekbib@uvs.edu

APPROVAL

NAAM: *Preka Soelèma*

verleent aan de AdeKUS kosteloos de niet-exclusieve toestemming om haar/zijn afstudeerscriptie via de catalogus en de institutional repository full-text beschikbaar te stellen aan gebruikers binnen en buiten de AdeKUS.

Plaats en datum, *Paramaribo, 11-febr-2022*

Handtekening *Preka S*



Anton de Kom Universiteit van Suriname

Faculty of Mathematical and Natural Sciences

Field of study: Mathematics

Modeling the COVID-19 outbreak in Suriname:
A comparison between a Stochastic model and a system of
Differential equations.

Thesis submitted in fulfillment of the requirements for the Bachelor of
Science degree in Mathematics (BSc.)

Name: Soelèma R. Pieka
Supervisor: R.E.J. Neslo
Date: August 23rd 2021

Preface

In front of you is the thesis “Modeling the COVID-19 outbreak in Suriname: A comparison between a Stochastic model and a system of Differential equations”. It was written to fulfill the graduation requirements of the Faculty of Mathematical and Natural Sciences’ Bachelor of Science program in Mathematics, at the Anton de Kom University in Suriname.

My research hypothesis was raised with my supervisor, Dr. R.E.J. Neslo. The research was difficult but conducting a thorough research allowed me to achieve the desired result.

I would like to thank my supervisor for his great advice and support throughout this process. I would also like to thank Dr. Venetiaan, who had given her time to help me with proofreading my thesis.

My mother deserves a special thanks: if not for her support and motivation, I would not have come this far.

Abstract

The ability to predict the future number of infected individuals during the COVID-19 outbreak plays a very important role in fighting the spread within a country. In this thesis we constructed a mathematical model for the COVID-19 pandemic during the 2020 outbreak in Suriname.

We created two compartmental models based on publicly available data from Suriname. We show that a stochastic model behaves just as well as a model based on differential equations. Due to the close fit of the stochastic compartmental model, we also assumed that this model could be used to predict future number of cases. An assumption we verified with data outside of the modeling data.

For a more accurate prediction model we recommend retrieving the data from the COVID-19 management team directly and dividing some of the compartments into more detailed compartments.

Contents

Introduction	5
1 Compartmental models	7
1.1 Compartmental models based on a system of differential equations	8
1.2 Stochastic models	11
2 Data	14
3 The Model	19
3.1 Compartmental model based on a system of differential equations.....	20
3.1.1 Differential Model version 1 (30-day average)	22
3.1.2 Differential Model version 2 (14-day average)	22
3.1.3 Differential Model version 3 (7-day average)	23
3.2 Stochastic compartmental models.....	25
3.2.1 Stochastic Model version 1 (30-day average).....	27
3.2.2 Stochastic Model version 2 (14-day average).....	28
3.2.3 Stochastic Model version 3 (7-day average).....	29
4 Results	32
4.1 30-day average comparison.....	32
4.2 14-day average comparison.....	36
4.3 7-day average comparison.....	39
4.4 Future predictions.....	42
5 Conclusion	48
Appendices	50
A. Data	50
B. Python script.....	58
C. References	71

Introduction

From December 31st of 2019 through January 3rd of 2020, the World Health Organization (WHO) received reports of cases of pneumonia, where the causal agent was yet to be identified. On January 7th of 2020, a new type of coronavirus, dubbed 2019-novel coronavirus (2019-nCoV), was isolated. By January 20th there had been reports of a total of 282 confirmed cases of 2019-nCoV. Of these cases, 278 were from China, 2 from Thailand, 1 from Japan and 1 from the Republic of Korea. Among the 282 cases, 6 deaths had already been reported (World Health Organization, 2020). Within one week the virus had spread to 11 additional countries. This brought the total number of confirmed cases to 4593, which included 106 deaths (World Health Organization, 2020). According to the data reported in the WHO's situation report on February 11th of 2020, there had been a total number of 43103 confirmed cases and 1018 deaths across 25 countries. By this time, the WHO had named the disease Coronavirus disease 2019, COVID-19 for short. At the time of writing there had been over 186.2 million confirmed cases with over 4 million deaths (World Health Organization, 2021).

Such as any other disease COVID-19 has several symptoms. These symptoms range from mild symptoms like a fever, to life threatening symptoms like respiratory problems. Some of these symptoms require constant medical care, which result in hospital admission. There are some cases where the patient must be admitted to the intensive care unit (ICU). This led to the following realization: if the number of COVID-19 cases increases, the number of patients who need to be admitted to the hospital will also increase. This poses a problem for countries with limited hospital resources. One of these countries is the republic of Suriname.

On March 13th of 2020 the first confirmed case of COVID-19 was reported in Suriname. By April 4th there had been 10 confirmed cases and 1 death. On May 19th of 2020, the number of confirmed cases started to increase, and the number of deaths increased on June 9th. As of July 1st of 2020, there had been 21,936 confirmed cases and 531 deaths (Suriname: Coronavirus Pandemic Country Profile, 2021). As these numbers kept growing every day, the strain on the medical sector of Suriname became noticeable.

In most cases of a disease outbreak policy makers enforce different kinds of policies to reduce the infection rate of the disease. Some of these policies include social distancing, mandatory quarantine for contacts of infected individuals, partial- and full lockdowns. Mathematical models are used to predict the effectiveness of these policies beforehand. These models help predict the progression of the epidemic caused by the disease. Which in turn can give some insight into the number of resources (medicine, doctors, hospital beds, finances, etc.) necessary to handle the epidemic.

Constructing a mathematical model that closely describes the COVID-19 pandemic in Suriname can be helpful in analyzing the behavior of the pandemic. It can also be used to predict future behavior of the disease. These predictions can in turn be used to see the effects certain policies could have on the pandemic.

Aleman, Wibisono and Schwartz (2009) presented a non-homogeneous agent-based simulation model which accounted for individual behavior and transmission rates at the Winter Simulation Conference. Their model considered the effect the disease had on the different age groups, the rate of infection for the different types of contact between individuals and the type of transportation. They argued that adding these factors made their model more accurate than the traditional non-homogeneous mixing models. However, the downside to their model was the time needed for the simulation (Aleman, Wibisono, & Schwartz, 2009).

Kretzschmar, Rozhnova and van Boven (2020) wrote a report that discussed the effect that isolation and contact tracing have on slowing down the COVID-19 epidemic. They constructed a stochastic model that described the early stages of an epidemic as a branching process. Their research showed that controlling the COVID-19 outbreak not only depended on isolating the infected but also tracking down and isolating known contact of the infected cases. This theory does however have some complications when it comes to using this in practice. This model can only be used while the situation can be described by a branching process and has exponential growth. On the plus side this model was constructed so that quantitative information can easily be inputted into the model for analysis (Kretzschmar, Rozhnova, & van Boven, 2020).

This thesis presents a compartmental model for the COVID-19 spread in Suriname. These models are commonly described using a system of differential equations, but for this thesis we also created a stochastic compartmental model. We did this under the hypothesis that ***modeling the COVID-19 pandemic in Suriname using a stochastic compartmental model will give similar results as modeling the pandemic using a compartmental model based on a system of differential equations.***

The first chapter gives a brief overview of background information regarding the mathematical models used in this thesis. In the second chapter we discuss the data and the process of cleaning the data used for this model. The third chapter focusses on the construction of the models. In the first section we discuss the construction of the model based on a system of differential equations and the second section presents the stochastic compartmental model. The fourth chapter presents the results and a comparison between both models. In the last section of this chapter, we did some future prediction using data outside the modeling data. In the last chapter of this thesis, we present our conclusions and recommendations for future research into this topic.

1 Compartmental models

This chapter focusses on the background information for the models used in this thesis. The base model used in this thesis is a compartmental model. Compartmental models are models that describe the relationship between compartments within a system.

When modeling a disease outbreak, the most used mathematical models are compartmental models. The population is divided into compartments. In most cases two of the compartments are the susceptible and the infected. The susceptible S are the part of the population who have not yet been infected by the disease and are at risk of becoming infected. The infected I are the part of the population who have been infected by the disease and this group can also infect individuals from the susceptible group. In some cases, there is a group named removed R . This group is the part of the population who have already recovered from the disease and are immune to reinfection or have died from the disease.

Depending on the properties of the disease, it can be modeled as an SIS or an SIR compartmental model. The former is used for nonfatal diseases where people who recover from the disease do not develop immunity against reinfection. In other words, a recovered individual becomes susceptible again. A possible path one can walk within this system is susceptible, infected, and then susceptible again. In the case of an SIR model, the disease can be nonfatal or recovered individual develop immunity against reinfection or it could be a combination of both. The parameters detailing the relationship between these compartments can be described using differential equations or stochastic probability matrices.

1.1 Compartmental models based on a system of differential equations

A differential equation is an equation that describes the connections between a quantity and the rate with which the quantity changes in a certain timeframe. The differential equations in compartmental models describe the connection between each compartment and the rate of change between these compartments. This rate of change can be expressed as the rate at which the disease is transferred.

The most used mathematical model for a disease outbreak is the *SIR* compartmental model. This model is usually described with a system of differential equations. In this model the population is split into three compartments. The compartment S_t consists of the susceptible people within the population at time. These people have not yet been infected. The risk of infection is assumed to be equal for everyone in this category. The infected people in the population are put into the infected I_t category. This group is also infective, because if they encounter individuals from S , they can infect these individuals. The last group R_t are the infected individuals who are removed. This removal can be due to vaccination, recovery with immunity against reinfection and in some cases death.

To determine describe how the compartments interact with each other, there are a few assumptions made for the *SIR* compartmental model. One of these assumptions is that the rate of infection is depended on the population N and that on average an individual encounters αN individuals per time unit with α being a function of the population size. This is called mass action incidence. The probability of an infected individual encountering a susceptible individual is equal to S/N . Putting these two numbers together we get that the number of new infections per time unit can be calculated with $(\alpha N)(S/N)$ which becomes αS . Which brings the rate of new infections to αSI . The rate at which the infected leave the infected group is assumed to be constant and equal to βI with β being the recovery rate (Brauer, Castillo-Chavez, & Feng, 2019). The following assumptions are made for the *SIR* compartmental model:

- (i) The rate of new infections, $\alpha S_t I_t$ is given by mass action incidence.
- (ii) At a rate of βI_t , infected leave the infective group and go to the removed group.
- (iii) N remains constant i.e., there is no entry into or departure from the population.
- (iv) The disease is not fatal.

(1.1.1)

Taking these assumptions into count, we get the following system of differential equations for the *SIR* compartmental model:

$$\begin{aligned}
S'_t &= -\alpha S_t I_t \\
I'_t &= \alpha S_t I_t - \beta I_t \\
R'_t &= \beta I_t
\end{aligned}
\tag{1.1.2}$$

With initial condition S_0 , I_0 and R_0 . S_t denotes the number of susceptible individuals at time t ; I_t denotes the number of infected individuals at time t ; R_t denotes the number of removed individuals at time t ; S'_t denotes the change within the susceptible group from time $t - 1$ to t , i.e. the number of new infections per time unit; I'_t denotes the change within the infective group from time $t - 1$ to t ; R'_t denotes the change within the removed group from time $t - 1$ to t i.e. the rate of recovery per time unit; α denotes the transmission rate and β denotes the recovery rate.

The results from the differential equations in (1.1.2) can be used to calculate the number of people in each compartment at any time using the following equations:

$$\begin{aligned}
S_t &= S_{t-1} + S'_{t-1} \\
I_t &= I_{t-1} + I'_{t-1} \\
R_t &= R_{t-1} + R'_{t-1}
\end{aligned}
\tag{1.1.3}$$

To further explain *SIR* compartmental model, we will present an example.

Consider community with a population of $N = 50000$ and a disease X with the following properties: a transmission rate α of 0.00001, a recovery rate β of 0.0715 and initial conditions $S_0 = 45400$, $I_0 = 2500$ and $R_0 = 2100$. This results in the following model:

$$\begin{aligned}
S'_t &= -(0.00001)S_t I_t \\
I'_t &= (0.00001)S_t I_t - (0.0715)I_t \text{ where } t \text{ is given in days.} \\
R'_t &= (0.0715)I_t
\end{aligned}
\tag{1.1.4}$$

The number of individuals in each compartment for 3 consecutive days can be calculated using (1.1.3) and (1.1.4). For day one at $t = 1$ we calculate S_1 , I_1 and R_1 the following way:

$$\begin{cases}
S_1 = S_0 + S'_0 = 45400 - 0.00001 S_0 I_0 \\
I_1 = I_0 + I'_0 = 2500 + 0.00001 S_0 I_0 - 0.0715 I_0 = \\
R_1 = R_0 + R'_0 = 2100 + 0.0715 I_0
\end{cases}$$

$$\begin{cases}
S_1 = 45400 - 0.00001 * 45400 * 2500 = 45400 - 1135 \\
I_1 = 2500 + 0.00001 * 45400 * 2500 - 0.0715 * 2500 = 2500 + 1135 - 178.75 \\
R_1 = 2100 + 0.0715 * 2500 = 2100 + 178.75
\end{cases}$$

Now we get $S_1 = 44265$, $I_1 = 3456.25$ and $R_1 = 2278.75$. Rounding these values, we get that at time $t = 1$ the number of susceptible individuals is 44265, the number of infected individuals is 3456 and the number of recovered individuals is equal to 2279. In the same way, the values for day two can be calculated.

$$\begin{cases} S_2 = S_1 + S'_1 = 44265 - 0.00001S_1I_1 \\ I_2 = I_1 + I'_1 = 3456 + 0.00001S_1I_1 - 0.0715I_1 = \\ R_2 = R_1 + R'_1 = 2279 + 0.0715I_1 \end{cases}$$

$$\begin{cases} S_2 = 44265 - 0.00001 * 44265 * 3456 = 44265 - 1529.7984 \\ I_2 = 3456 + 0.00001 * 44265 * 3456 - 0.0715 * 3456 = 3456 + 1529.7984 - 247.104 \\ R_2 = 2279 + 0.0715 * 3456 = 2279 + 247.104 \end{cases}$$

After rounding we get the following values: we get $S_2 = 42735$, $I_2 = 4739$ and $R_2 = 2526$. These values are then used to calculate the situation for the third day.

$$\begin{cases} S_3 = S_2 + S'_2 = 42735 - 0.00001S_2I_2 \\ I_3 = I_2 + I'_2 = 4739 + 0.00001S_2I_2 - 0.0715I_2 = \\ R_3 = R_2 + R'_2 = 2526 + 0.0715I_2 \end{cases}$$

$$\begin{cases} S_3 = 42735 - 0.00001 * 42735 * 4739 = 42735 - 2025.21165 \\ I_3 = 4739 + 0.00001 * 42735 * 4739 - 0.0715 * 4739 = 4739 + 2025.21165 - 338.8385 \\ R_3 = 2526 + 0.0715 * 4739 = 2526 + 338.8385 \end{cases}$$

After rounding we get the following values: we get $S_3 = 40710$, $I_3 = 6425$ and $R_3 = 2865$. It should be noted that for all three days, the number of people for each compartment sums up to the population. The next section gives a brief overview of stochastic compartmental models.

1.2 Stochastic models

A stochastic process is a sequence of variables. These variables represent events whose outcome happen by chance. For instance, if some dice is tossed 6 times, then the stochastic process is the collection of the outcome for each of the 6 throws. So, a stochastic process $X = \{X_t, t \in T\}$ is a collection of random variables. The time is given by t and $X(t)$ is the outcome of the process at time t .

Stochastic processes can be discrete or continuous time processes. If a process is a continuous time process, then it means that the time t can be any positive real number. The process then looks like $X = \{X_t, t \geq 0\}$. An example a of continuous time processes is the intermediate arrival time of clients of a department store in minutes (2.4, 5, 0.75, 4.25, etc.).

In a discrete time process the time is given by any positive natural number. The process is observed during discrete time periods. In this process $X = \{X_t, t = 0, 1, 2, \dots\}$ the time can represent any discrete period (hours, days, months, years, etc.). In this thesis the COVID-19 pandemic was described as a discrete time stochastic process because it was observed daily. The possible outcomes of a discrete time stochastic process are called states. The jump from state X_t to state X_{t-1} is called a transition.

The probability of transitioning between the possible states within a discrete time stochastic process is called a transition probability. The collection of all these probabilities can be written in a matrix transition probability matrix.

In general, discrete time stochastic processes that have s states, have an $s \times s$ transition probability matrix:

$$\begin{pmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,s} \\ p_{2,1} & p_{2,2} & \dots & p_{2,s} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ p_{s,1} & p_{s,2} & \dots & p_{s,s} \end{pmatrix} \quad (1.2.1)$$

Where the probabilities $p_{i,j}$ have the following properties:

- (i) $p_{i,j} = P(X_{t+1} = j | X_t = i)$: the probability for the transition i to j from time t to $t + 1$.
- (ii) $0 \leq p_{i,j} \leq 1, \forall i, j \in \mathcal{S}$, the set containing all the states.
- (iii) $\sum_{j \in \mathcal{S}} p_{i,j} = 1, \forall i \in \mathcal{S}$

The probability $p_{i,j}$ is the probability that if the system is in state i at time t , it will be in state j at time $t + 1$. If a state is equal to i at time t , the process must be in one of the states within the system at time $t + 1$. Hence every probability $p_{i,j}$ is either equal to 0 or a positive

number between 0 and 1. Also all the probabilities in each row must sum up to 1 (Winston, 2004).

The transition probability matrix shown above can be used to calculate the likelihood that a process will be in a certain state at the next period. Suppose that we would be interested in knowing in what state a process would be after n periods and the transition probability matrix is defined as P then we would have to construct $P(n)$. This matrix has the following properties:

- (i) $P(n) = (P)^n$
 - (ii) $p^n_{i,j} = P(X_{t+n} = j | X_t = i)$: the probability that the process will be in state j after n periods, given that it was in state i at time t .
 - (iii) $0 \leq p^n_{i,j} \leq 1, \forall i, j \in \mathcal{S}$, the set containing all the states.
 - (vi) $\sum_{j \in \mathcal{S}} p^n_{i,j} = 1, \forall i \in \mathcal{S}$
- (1.2.3)**

The transition probability matrix can be used to answer different kinds of questions. The transition probability matrix in its original form can tell us how likely it is that a system in a certain state will be in another state after one period. Given a quantity of elements in each state, it can also tell us the number of elements one could expect in each state after one period. To further explain this, we will present an example.

Consider a system with the states $\mathcal{S} = \{X, Y\}$ and a transition probability matrix P :

	X	Y
X	0.3	0.7
Y	0.8	0.2

	X	Y
X	$p_{X,X}$	$p_{X,Y}$
Y	$p_{Y,X}$	$p_{Y,Y}$

This transition probability matrix in this form gives us the following information: Given that the system is currently ($t = 0$) in state X , the probability that the system will be in state Y at the next period ($t = 1$) is equal to 0.7.

If we would like to know how the probability distribution will be after 3 periods, i.e., how will the transition probabilities look at time $t = 3$, then we will have to calculate $(P)^3$. This is done in the following manner:

$$(P)^3 = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}^3 = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix} \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix} \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}^2 \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.65 & 0.35 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}$$

Now we end up with $(P)^3 = \begin{bmatrix} 0.475 & 0.525 \\ 0.6 & 0.4 \end{bmatrix}$. Before we continue, it should be noted that properties (1.2.3) (iii) and (vi) are true for the matrix $(P)^3$. The matrix $(P)^3$ gives us the following information: If an element within the system starts in state Y at time $t = 0$, there is 40% chance that it will still be in state Y after 3 periods i.e., at time $t = 3$.

Given the information that at $t = 0$ there are 120 elements in state X ($X_0 = 120$) and 180 elements in state Y ($Y_0 = 180$), we can use this information in combination with the transition probability matrix to predict the number of elements in each state after 1 period (X_1 and Y_1). The values for X_0 and Y_0 are called the initial condition. X_1 will consist of a portion of the elements from X_0 that stay in X and a portion of the elements of Y_0 that transition to X . So

$X_1 = p_{X,X}X_0 + p_{Y,X}Y_0$. Y_1 will consist of a portion of the elements from Y_0 that stay in Y and a portion of the elements of X_0 that transition to Y . So $Y_1 = p_{X,Y}X_0 + p_{Y,Y}Y_0$.

Putting these two formulas together we get:

$$X_1 = p_{X,X}X_0 + p_{Y,X}Y_0$$

$$Y_1 = p_{X,Y}X_0 + p_{Y,Y}Y_0$$

If we write these equations as the result of matrix calculations, we get:

$$\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = \begin{bmatrix} p_{X,X}X_0 + p_{Y,X}Y_0 \\ p_{X,Y}X_0 + p_{Y,Y}Y_0 \end{bmatrix} = \begin{bmatrix} p_{X,X} & p_{Y,X} \\ p_{X,Y} & p_{Y,Y} \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} = \begin{bmatrix} p_{X,X} & p_{Y,X} \\ p_{X,Y} & p_{Y,Y} \end{bmatrix}^T \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} = P^T \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$$

This gives us the formula for calculating X_1 and Y_1 and using these gives us the following result:

$$\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = P^T \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{bmatrix}^T \begin{bmatrix} 120 \\ 180 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.8 \\ 0.7 & 0.2 \end{bmatrix} \begin{bmatrix} 120 \\ 180 \end{bmatrix} = \begin{bmatrix} 180 \\ 120 \end{bmatrix}$$

These results tell us that if at time $t = 0$ X has 120 elements and Y has 180 elements, we can estimate that at time $t = 1$, X will have 180 elements and Y 120 elements.

This example is a small representation of how the model in this thesis will operate. Now that we have given a brief overview of stochastic compartmental models, in the next chapter we will discuss the data used and the method used in cleaning the data used for the models.

2 Data

The data used in this research was observed from August 1st to September 30th, 2020. This choice was made because acquiring the necessary data directly from the COVID-19 management team had proven difficult. And since attempts at contacting the team were unsuccessful, we had to retrieve data from the Suriname COVID-19 website. Before August 1st, 2020, the website did not show detailed data concerning the COVID-19 numbers in Suriname, which is why we did not start our data collection before this time. By the time the management team had published more detailed data on the website, we were well within the month of July, so we decided to start data collection on August 1st. The data from August 1st to August 30st was used to train the model and the remaining data from September was used to test the model.

From August 1st of 2020 to September 30th of 2020 COVID-19 data was retrieved every day from the Suriname COVID-19 website. The retrieved data consisted of the categories: new positive cases over the last 24 hours, new deceased cases over the last 24 hours, new recovered cases over the last 24 hours, total number of positive cases, total number of deceased cases, total number of recovered cases, number of positive cases in isolation, number of positive cases in the nursing ward, number of positive cases in the ICU and the number of active cases. According to the General Bureau of Statistics of Suriname as of July 2018, the estimated year population of Suriname was equal to 590100. At the time of writing this was the most recent population estimate, hence we assumed this number to be a good approximation to the actual population total. The susceptible category was calculated using the population number retrieved from this report. This data was defined using the following variables:

- S : The group of individuals who are susceptible to the disease.
 - J : The infected individuals who show minimal to no symptoms.
 - N : The infected individuals admitted to the nursing ward.
 - C : This group consist of infected individuals admitted to the ICU.
 - G : The total number of the people who recovered from the disease.
 - ΔG : The number of newly recovered individuals.
 - D : Deceased. The total number of people who died from the disease.
 - ΔD : The number of newly deceased individuals.
 - P : The total number of positive cases.
 - ΔP : The number of newly tested positive individuals.
 - A : The number of active cases.
 - t : The time in days with $t = 0$ equivalent to August 1st, 2020.
- (2.1)

The collected data can be viewed in the tables on the next page.

(2.2) Observed COVID-19 data.

t	Date	ΔP	P	ΔD	D	A	ΔG	G	C	N	J
0	01-08	54	1760	0	26	573	43	1161	12	88	561
1	02-08	89	1849	1	27	622	33	1194	15	80	613
2	03-08	44	1893	0	27	639	32	1227	12	71	556
3	04-08	88	1981	0	27	676	51	1278	10	90	576
4	05-08	69	2050	0	27	696	49	1327	13	88	595
5	06-08	46	2096	2	29	621	119	1446	9	91	521
6	07-08	107	2203	0	29	669	59	1505	12	99	558
7	08-08	103	2306	0	29	719	54	1558	7	82	630
8	09-08	85	2391	0	29	727	77	1635	13	82	632
9	10-08	98	2489	2	31	785	39	1674	13	90	682
10	11-08	70	2559	8	39	808	38	1712	14	116	678
11	12-08	94	2653	0	39	825	77	1789	15	118	692
12	13-08	108	2761	1	40	891	41	1830	12	120	759
13	14-08	77	2838	1	41	893	64	1894	15	153	735
14	15-08	123	2961	1	42	929	102	1990	15	132	820
15	16-08	55	3016	5	47	933	67	2063	11	116	852
16	17-08	61	3077	1	48	892	118	2138	13	132	776
17	18-08	139	3216	6	54	966	63	2196	13	126	859
18	19-08	79	3295	0	54	1014	43	2239	17	143	910
19	20-08	71	3366	1	55	928	169	2383	19	129	832
20	21-08	94	3460	1	56	908	38	2498	20	188	752
21	22-08	109	3569	1	57	951	61	2559	20	210	777
22	23-08	38	3607	1	58	876	114	2673	20	145	711
23	24-08	25	3632	2	60	814	70	2758	21	190	638
24	25-08	66	3698	1	61	819	60	2818	20	166	655
25	26-08	26	3724	1	62	799	44	2862	23	165	635
26	27-08	69	3793	4	66	834	30	2893	23	133	702
27	28-08	55	3848	0	66	811	31	2971	24	157	660
28	29-08	106	3954	1	67	896	20	2991	24	173	734
29	30-08	55	4009	0	67	869	82	3073	23	158	718
30	31-08	25	4034	4	71	823	67	3140	18	145	634
31	01-09	55	4089	1	72	846	27	3167	19	153	703
32	02-09	60	4149	0	72	805	106	3273	19	114	712
33	03-09	66	4215	1	73	824	46	3319	17	112	746
34	04-09	37	4252	2	75	811	48	3367	17	120	718
35	05-09	68	4320	1	76	818	60	3426	15	120	723
36	06-09	26	4346	9	85	767	68	3494	12	116	674
37	07-09	14	4360	6	91	725	50	3544	8	108	635
38	08-09	59	4419	0	91	733	51	3595	9	106	640
39	09-09	28	4447	1	92	722	38	3633	8	107	626
40	10-09	30	4477	1	93	678	73	3706	9	73	589

(2.3) Observed COVID-19 data (continued).

t	Date	ΔP	P	ΔD	D	A	ΔG	G	C	N	J
41	11-09	52	4529	0	93	689	41	3747	8	89	610
42	12-09	50	4579	0	93	698	41	3788	7	93	613
43	13-09	3	4582	0	93	639	62	3850	7	99	548
44	14-09	29	4611	2	95	571	85	3935	9	85	489
45	15-09	14	4625	0	95	534	61	3996	7	69	472
46	16-09	20	4645	0	95	461	92	4089	7	54	414
47	17-09	26	4671	0	95	416	71	4160	7	57	364
48	18-09	20	4691	1	96	315	120	4280	7	65	282
49	19-09	18	4709	1	97	229	78	4383	7	66	181
50	20-09	14	4723	0	97	138	105	4488	6	66	96
51	21-09	17	4740	0	97	118	37	4525	4	22	49
52	22-09	19	4759	3	100	118	16	4541	4	60	64
53	23-09	20	4779	1	101	118	19	4560	5	54	71
54	24-09	10	4789	1	102	109	18	4578	5	66	76
55	25-09	28	4817	0	102	119	18	4596	6	58	72
56	26-09	14	4831	0	102	109	24	4620	4	44	51
57	27-09	4	4835	0	102	72	41	4661	6	39	42
58	28-09	1	4836	0	102	67	6	4667	7	31	41
59	29-09	27	4863	2	104	83	8	4675	5	33	37
60	30-09	14	4877	0	104	78	12	4695	5	44	31

To validate this data, we made some assumptions based on news reports about the data we collected and then checked to see whether the collected data was in accordance with the assumptions. The assumptions were:

$$\begin{aligned}
 & \text{(i) } A_t = J_t + N_t + C_t \text{ at any time; } t = 0,1,2, \dots, 60 \\
 & \text{(ii) } P_t = G_t + D_t + A_t \text{ at any time; } t = 0,1,2, \dots, 60
 \end{aligned}
 \tag{2.4}$$

The first equation addresses the assumption that the sum of the cases in isolation, the cases in the nursing ward and the ICU is equal to the number of active cases. The second equation describes the composition of the total number of positive cases. While checking the data with the calculations above we discovered that the data was not accurate. The reason for the inconsistency was not reported in the media, but we believe that this could be due to an error during the data entry process. After seeing that the data was not accurate, we started the data cleaning process.

It was reported in the media that the COVID-19 management team had lost contact with several people in isolation which led us to assume that the reported number of isolated cases could be incorrect. We also assumed that the data concerning the ICU and the nursing ward were correct, because these numbers were reported by the hospital staff and the chance of these numbers being accurate was greater than the chance of them being incorrect. Since the data concerning the new positive cases, the total positive cases, the new deceased cases, and the total

number of deceased cases were accurate we assumed that they were correct. We also assumed that the number of active cases is correct. The numbers for the remaining categories were corrected using these assumptions and the following formulas:

$$\begin{aligned}
 & \text{(i) } J_t = A_t - C_t - N_t ; t = 0 \dots 30 \\
 & \text{(ii) } G_t = P_t - D_t - A_t ; t = 0 \dots 30 \\
 & \text{(iii) } \Delta G_t = G_t - G_{t-1} ; t = 1 \dots 30 \\
 & \text{(iv) } S_t = 590100 - P_t ; t = 0 \dots 30
 \end{aligned}
 \tag{2.5}$$

The first equation uses the active cases, the cases in the hospital nursing ward and the ICU to calculate the number of cases in isolation at any given time. The second equation uses the total number of the total number of positive cases, the total number of deceased cases and the active cases to calculate the total number of recovered cases. The third equation calculates the newly recovered cases. The fourth equation calculates the number of susceptible individuals at any time. The data cleaning calculations were done using Microsoft excel. After cleaning up the data, we came up with the following dataset:

(2.6) Corrected COVID-19 data.

<i>t</i>	Date	<i>S*</i>	ΔP	<i>P</i>	ΔD	<i>D</i>	<i>A</i>	ΔG	<i>G</i>	<i>C</i>	<i>N</i>	<i>J</i>
0	01-08	588340	54	1760	0	26	573	0	1161	12	88	473
1	02-08	588251	89	1849	1	27	622	39	1200	15	80	527
2	03-08	588207	44	1893	0	27	639	27	1227	12	71	556
3	04-08	588119	88	1981	0	27	676	51	1278	10	90	576
4	05-08	588050	69	2050	0	27	696	49	1327	13	88	595
5	06-08	588004	46	2096	2	29	621	119	1446	9	91	521
6	07-08	587897	107	2203	0	29	669	59	1505	12	99	558
7	08-08	587794	103	2306	0	29	719	53	1558	7	82	630
8	09-08	587709	85	2391	0	29	727	77	1635	13	82	632
9	10-08	587611	98	2489	2	31	785	38	1673	13	90	682
10	11-08	587541	70	2559	8	39	808	39	1712	14	116	678
11	12-08	587447	94	2653	0	39	825	77	1789	15	118	692
12	13-08	587339	108	2761	1	40	891	41	1830	12	120	759
13	14-08	587262	77	2838	1	41	893	74	1904	15	153	725
14	15-08	587139	123	2961	1	42	929	86	1990	15	132	782
15	16-08	587084	55	3016	5	47	933	46	2036	11	116	806
16	17-08	587023	61	3077	1	48	892	101	2137	13	132	747
17	18-08	586884	139	3216	6	54	966	59	2196	13	126	827
18	19-08	586805	79	3295	0	54	1014	31	2227	17	143	854
19	20-08	586734	71	3366	1	55	928	156	2383	19	129	780
20	21-08	586640	94	3460	1	56	908	113	2496	20	188	700
21	22-08	586531	109	3569	1	57	951	65	2561	20	210	721

* This data field was calculated using (2.5) and added to the dataset.

● These values have been corrected.

(2.7) Corrected COVID-19 data.

t	Date	S^*	ΔP	P	ΔD	D	A	ΔG	G	C	N	J
22	23-08	586493	38	3607	1	58	876	112	2673	20	145	711
23	24-08	586468	25	3632	2	60	814	85	2758	21	190	603
24	25-08	586402	66	3698	1	61	819	60	2818	20	166	633
25	26-08	586376	26	3724	1	62	799	45	2863	23	165	611
26	27-08	586307	69	3793	4	66	834	30	2893	23	133	678
27	28-08	586252	55	3848	0	66	811	78	2971	24	157	630
28	29-08	586146	106	3954	1	67	896	20	2991	24	173	699
29	30-08	586091	55	4009	0	67	869	82	3073	23	158	688
30	31-08	586066	25	4034	4	71	823	67	3140	18	145	660
31	01-09	586011	55	4089	1	72	846	31	3171	19	153	674
32	02-09	585951	60	4149	0	72	805	101	3272	19	114	672
33	03-09	585885	66	4215	1	73	824	46	3318	17	112	695
34	04-09	585848	37	4252	2	75	811	48	3366	17	120	674
35	05-09	585780	68	4320	1	76	818	60	3426	15	120	683
36	06-09	585754	26	4346	9	85	767	68	3494	12	116	639
37	07-09	585740	14	4360	6	91	725	50	3544	8	108	609
38	08-09	585681	59	4419	0	91	733	51	3595	9	106	618
39	09-09	585653	28	4447	1	92	722	38	3633	8	107	607
40	10-09	585623	30	4477	1	93	678	73	3706	9	73	596
41	11-09	585571	52	4529	0	93	689	41	3747	8	89	592
42	12-09	585521	50	4579	0	93	698	41	3788	7	93	598
43	13-09	585518	3	4582	0	93	639	62	3850	7	99	533
44	14-09	585489	29	4611	2	95	571	95	3945	9	85	477
45	15-09	585475	14	4625	0	95	534	51	3996	7	69	458
46	16-09	585455	20	4645	0	95	461	93	4089	7	54	400
47	17-09	585429	26	4671	0	95	416	71	4160	7	57	352
48	18-09	585409	20	4691	1	96	315	120	4280	7	65	243
49	19-09	585391	18	4709	1	97	229	103	4383	7	66	156
50	20-09	585377	14	4723	0	97	138	105	4488	6	66	66
51	21-09	585360	17	4740	0	97	118	37	4525	4	22	92
52	22-09	585341	19	4759	3	100	118	16	4541	4	60	54
53	23-09	585321	20	4779	1	101	118	19	4560	5	54	59
54	24-09	585311	10	4789	1	102	109	18	4578	5	66	38
55	25-09	585283	28	4817	0	102	119	18	4596	6	58	55
56	26-09	585269	14	4831	0	102	109	24	4620	4	44	61
57	27-09	585265	4	4835	0	102	72	41	4661	6	39	27
58	28-09	585264	1	4836	0	102	67	6	4667	7	31	29
59	29-09	585237	27	4863	2	104	83	9	4676	5	33	45
60	30-09	585223	14	4877	0	104	78	19	4695	5	44	29

* This data field was calculated using (2.5) and added to the dataset.

● These values have been corrected.

After cleaning up the data, we continued with the construction of the models which we present in the next chapter.

3 The Model

After cleaning up the data, we started with the construction of the mathematical model. The first step was defining the compartments. We used the S , I and R compartments for this disease. These compartments are defined as follows:

- S : The individuals who are susceptible to the disease.
 - I : The individuals who were infected i.e., the individuals in the hospital nursing ward and the ICU and in isolation
 - R : The individuals who have been removed i.e., the individuals who died from the disease and the individuals who recovered.
- (3.1)**

The compartments in our model had the following composition:

- (i) $S_t = S_t$; $t = 0 \dots 60$
 - (ii) $I_t = A_t$; $t = 0 \dots 60$
 - (iii) $R_t = G_t + D_t$; $t = 1 \dots 60$
- (3.2)**

Using these new compartment definitions, we proceeded to the next step in constructing the model. We constructed a compartmental model with differential equations and a stochastic compartmental model for this disease. The results from both models were then compared to see whether the results from the stochastic compartmental model approximated the observed values as well as the results from the differential equations. The first section of this chapter focusses on the model based on a system of differential equations.

3.1 Compartmental model based on a system of differential equations

The first model we constructed was the model based on a system of differential equations. For this model we used the same methods and assumptions as discussed earlier in section 1.1. Some of these assumptions were adjusted to fit our disease. Our model operated under the following assumption:

- (i) The rate of new infections, $\alpha S_t I_t$ is given by mass action incidence.
 - (ii) At a rate of βI_t , infected leave the infective group and go to the removed group. In this case the removed group consists of recovered and deceased individuals.
 - (iii) N remains constant i.e., there is no entry into or departure from the population.
- (3.1.1)

The reason we removed the assumption (iv) of (1.1.1) from our model is because the disease that our model is based on is fatal. And our removed category is comprised of recovered and diseased individuals. Taking these assumptions into a count, our general model becomes:

$$\begin{aligned}
 S'_t &= -\alpha S_t I_t \\
 I'_t &= \alpha S_t I_t - \beta I_t \\
 R'_t &= \beta I_t \\
 S_t &= S_{t-1} + S'_{t-1} \\
 I_t &= I_{t-1} + I'_{t-1} \\
 R_t &= R_{t-1} + R'_{t-1}
 \end{aligned}$$

(3.1.2)

With initial condition S_0, I_0 and R_0 . S_t denotes the number of susceptible individuals at time t . I_t denotes the number of infected individuals at time t . R_t denotes the number of removed individuals at time t . S'_t denotes the change within the susceptible group from time $t - 1$ to t , i.e., the number of new infections per time unit. I'_t denotes the change within the infective group from time $t - 1$ to t ; R'_t denotes the change within the removed group from time $t - 1$ to t i.e., the rate of recovery per time unit. t is the time given in days, α denotes the transmission rate and β denotes the recovery rate. The transmission rate α and the recovery rate β are unknown at this point and must be estimated. To estimate these, we used the formulas in (3.1.2) and the corrected data from chapter 2. Deriving the formula for estimating these parameters was done in the following way.

Derivation of α

S_i is the number of susceptible individuals at time $t = i$.
 $S'_{i-1} = (S_i - S_{i-1})$ is the change within the S category from time $i - 1$ to i . Using the first formula in (3.1.2) we get the equality:
 $-\alpha S_{i-1} I_{i-1} = (S_i - S_{i-1})$. From this equality we can estimate the transmission rate α_i from time $i - 1$ to i by $\alpha_i = \frac{-(S_i - S_{i-1})}{S_{i-1} I_{i-1}}$. If we assume that the transmission rate α can be approximated with the average of α_i , we get that $\alpha = \frac{\sum_{i=1}^k \alpha_i}{k}$.

Derivation of β

R_i is the number of removed individuals at time $t = i$.
 $R'_{i-1} = (R_i - R_{i-1})$ is the change within the R category from time $i - 1$ to i . Using the third formula in (3.1.2) we get the equality:
 $\beta I_{i-1} = (R_i - R_{i-1})$. From this equality we can estimate the recovery rate β_i from time $i - 1$ to i by $\beta_i = \frac{(R_i - R_{i-1})}{I_{i-1}}$. If we assume that the recovery rate β can be approximated with the average of β_i , we get that $\beta = \frac{\sum_{i=1}^k \beta_i}{k}$.

After deriving transmission and recovery rate, the next step was finding out how the model would operate. The calculations for estimating the rates depended on the average of the daily rates. The next challenge was to determine the length of the time interval needed to calculate average rates. To determine this, we looked at 3 lengths: 30 days, 14 days, and 7 days. We calculated the rates and the fitted values for each length. In the next three subsections we present the models for each of the different lengths. Before we do this, we define the following variables for both the stochastic model and the model based on differential equations:

- S_f : The fitted values for the S compartment.
 - I_f : The fitted values for the I compartment.
 - R_f : The fitted values for the R compartment.
 - Se : The absolute error terms for the S compartment.
 - Ie : The absolute error terms for the I compartment.
 - Re : The absolute error terms for the R compartment.
- (3.1.3)

And the corresponding calculations:

- $Se_i = S_i - S_f_i, i = 0,1,2, \dots, 30$
 - $Ie_i = I_i - I_f_i, i = 0,1,2, \dots, 30$
 - $Re_i = R_i - R_f_i, i = 0,1,2, \dots, 30$
- (3.1.4)

Now that we defined a few new variables, and we can start with the first version of the model. Every calculation done in the remaining parts of this chapter was done using Python programming language. The script for these calculations can be viewed in appendix B.

3.1.1 Differential Model version 1 (30-day average)

The first version of our model calculated the rates using a length of 30 days. In this case we took

$$\begin{aligned}\alpha_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (1, 2, \dots, 30), \alpha = \frac{\sum_{t=1}^{30} \alpha_t}{30} \\ \beta_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (1, 2, \dots, 30), \beta = \frac{\sum_{t=1}^{30} \beta_t}{30}\end{aligned}\tag{3.1.5}$$

These calculations resulted in the following system of differential equations:

$$\begin{aligned}S'_t &= -(1.59 \times 10^{-7})S_t I_t \\ I'_t &= (1.59 \times 10^{-7})S_t I_t - 0.08I_t \\ R'_t &= 0.08I_t\end{aligned}$$

After constructing this system, we used it to calculate the fitted values for each compartment with the initial condition equal to the data observed on August 1st. We also calculated the error terms using (3.1.4). Calculating the fitted values was done using the last three formulas in (3.1.2) and the method in the example from section 1.1.

3.1.2 Differential Model version 2 (14-day average)

In our next version, we calculated the transition probability using a length of 14 days. Here we took

$$\begin{aligned}\alpha(1)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (1, 2, \dots, 14), \alpha(1) = \frac{\sum_{t=1}^{14} \alpha(1)_t}{14} \\ \beta(1)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (1, 2, \dots, 14), \beta(1) = \frac{\sum_{t=1}^{14} \beta(1)_t}{14} \\ \alpha(2)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (15, 16, \dots, 28), \alpha(2) = \frac{\sum_{t=15}^{28} \alpha(2)_t}{14} \\ \beta(2)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (15, 16, \dots, 28), \beta(2) = \frac{\sum_{t=15}^{28} \beta(2)_t}{14} \\ \alpha(3)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (29, 30), \alpha(3) = \frac{\sum_{t=29}^{30} \alpha(3)_t}{2} \\ \beta(3)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (29, 30), \beta(3) = \frac{\sum_{t=29}^{30} \beta(3)_t}{2}\end{aligned}\tag{3.1.6}$$

These calculations resulted in 3 different systems of differential equations, one for each 14-day length part of the data. The system was calculated using only 2 data points. This, because

these two data points were left after calculating the first two matrices. We got the following systems:

$$\begin{aligned} S'_t &= -(1.95 \times 10^{-7})S_t I_t \\ \text{System 1} = I'_t &= (1.95 \times 10^{-7})S_t I_t - 0.08I_t \\ R'_t &= 0.08I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(1.34 \times 10^{-7})S_t I_t \\ \text{System 2} = I'_t &= (1.34 \times 10^{-7})S_t I_t - 0.08I_t \\ R'_t &= 0.08I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(7.99 \times 10^{-8})S_t I_t \\ \text{System 3} = I'_t &= (7.99 \times 10^{-8})S_t I_t - 0.09I_t \\ R'_t &= 0.09I_t \end{aligned}$$

After estimating these systems, we calculated the fitted values for each compartment with the initial condition for system 1, 2 and 3 set to the data observed on August 1st, August 15th, and August 29th respectively.

3.1.3 Differential Model version 3 (7-day average)

The last version of the model we constructed, had rates that were calculated using an average of 7 days.

$$\begin{aligned} \alpha(1)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (1, 2, \dots, 7), \alpha(1) = \frac{\sum_{t=1}^7 \alpha(1)_t}{7} \\ \beta(1)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (1, 2, \dots, 7), \beta(1) = \frac{\sum_{t=1}^7 \beta(1)_t}{7} \\ \alpha(2)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (8, 9, \dots, 14), \alpha(1) = \frac{\sum_{t=8}^{14} \alpha(2)_t}{7} \\ \beta(2)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (8, 9, \dots, 14), \beta(1) = \frac{\sum_{t=8}^{14} \beta(2)_t}{7} \\ \alpha(3)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (15, 16, \dots, 21), \alpha(1) = \frac{\sum_{t=15}^{21} \alpha(3)_t}{7} \\ \beta(3)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (15, 16, \dots, 21), \beta(1) = \frac{\sum_{t=15}^{21} \beta(3)_t}{7} \\ \alpha(4)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (22, 23, \dots, 28), \alpha(1) = \frac{\sum_{t=22}^{28} \alpha(4)_t}{7} \\ \beta(4)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (22, 23, \dots, 28), \beta(1) = \frac{\sum_{t=22}^{28} \beta(4)_t}{7} \\ \alpha(5)_t &= \frac{-(S_t - S_{t-1})}{S_{t-1}} ; t \in (29, 30), \alpha(1) = \frac{\sum_{t=29}^{30} \alpha(5)_t}{2} \\ \beta(5)_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (29, 30), \beta(1) = \frac{\sum_{t=29}^{30} \beta(5)_t}{2} \end{aligned} \tag{3.1.7}$$

These calculations resulted in 5 different systems of differential equations, one for each 7-day length part of the data.

$$\begin{aligned} S'_t &= -(1.99 \times 10^{-7})S_t I_t \\ \text{System 1} = I'_t &= (1.99 \times 10^{-7})S_t I_t - 0.09I_t \\ R'_t &= 0.09I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(1.9 \times 10^{-7})S_t I_t \\ \text{System 2} = I'_t &= (1.9 \times 10^{-7})S_t I_t - 0.08I_t \\ R'_t &= 0.08I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(1.57 \times 10^{-7})S_t I_t \\ \text{System 3} = I'_t &= (1.57 \times 10^{-7})S_t I_t - 0.09I_t \\ R'_t &= 0.09I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(1.11 \times 10^{-7})S_t I_t \\ \text{System 4} = I'_t &= (1.11 \times 10^{-7})S_t I_t - 0.07I_t \\ R'_t &= 0.07I_t \end{aligned}$$

$$\begin{aligned} S'_t &= -(7.99 \times 10^{-8})S_t I_t \\ \text{System 5} = I'_t &= (7.99 \times 10^{-8})S_t I_t - 0.09I_t \\ R'_t &= 0.09I_t \end{aligned}$$

After estimating these systems, we calculated the fitted values for each compartment with the initial condition for system 1, 2, 3, 4 and 5 set to the data observed on August 1st, August 8th, August 15th, August 22nd, and August 29th respectively. The next section presents the construction of the stochastic compartmental model.

3.2 Stochastic compartmental models

The stochastic compartmental model we used is like the model developed in “Markov Chain Modeling of HIV, Tuberculosis, and Hepatitis B Transmission in Ghana” (Twumasi, Asiedu, & Nortey, 2019). In this article the writers developed a stochastic *SIR* model for a few diseases. The most important difference between this model and our model is the method with which the transition probabilities were calculated. The method they used to calculate the transition probabilities is the maximum likelihood estimation. In our model we used a different method, which we explain in this section. Our model operated under the following assumptions:

- (i) A susceptible individual can become infected or stay susceptible.
 - (ii) An infected individual can stay infected or move to the removed group.
 - (iii) A removed individual will always stay in the removed group.
 - (iv) No one enters the population, i.e., there are no births and arrivals from other countries.
 - (v) Besides the COVID-19 deceased individuals, no one leaves the population, i.e., no one leaves the country or dies from anything other than COVID-19.
 - (vi) If the population is denoted by N then $N = S + I + R$ is constant at any time t .
- (3.2.1)

These assumptions help make the model solvable. Without these assumptions the model could become too complex to solve. After the model is constructed, the assumptions can be adjusted to make the model more realistic. Taking these assumptions into count we get the following stochastic model:

- $X = \{X_t, t = 0, 1, 2, 3, \dots\}$ the current state at time t .
- $E = \{S, I, R\}$ the set containing all the possible states.
- Probability transition matrix P ; $P_{i,j} = P\{X_{t+1} = j | X_t = i\}$;
 $0 \leq P_{i,j} \leq 1; \sum_{j \in E} P_{i,j} = 1; i, j \in E$

With the matrix P becoming:

	S	I	R
S	$1 - u$	u	0
I	0	$1 - v$	v
R	0	0	1

(3.2.2)

With

- u : the probability of transitioning from S at time t , to I at time $t + 1$.
- v : the probability of transitioning from I at time t , to R at time $t + 1$.
- $1 - u$: the probability staying in S from time t to time $t + 1$.
- $1 - v$: the probability of staying in I from time t to time $t + 1$.

Before moving on to the next step in the model construction, we will talk about the other elements in matrix P . The probabilities that are equal to 0 are a direct consequence of (3.2.1). The transitions with probabilities equal to 0, represent transitions that are not possible according to our assumptions. The transition with a probability equal to 1, means that once the system is in the state R it will remain in this state. After defining the model, we derived the transition probabilities in the following way:

Derivation of u

S_i is the number of susceptible individuals at time $t = i$ and S_j at time $t = j$. From the assumptions made earlier in (3.2.1), we know that

$$(S_i \geq S_j) \forall i, j \in (0, 1, 2, \dots, 60) \wedge i \leq j.$$

$(S_i - S_j)$ is the total number of individuals who moved from the S category to the I category during the time interval i to j .

This follows from the assumptions made earlier in (3.2.1).

The proportion of individuals who transitioned from S to I during this time interval is equal to $\frac{S_i - S_j}{S_i}$.

We define u_i as the number of individuals who transitioned from S to I during the time interval $i - 1$ to i proportional to the number of individuals in S at time $i - 1$, so $u_i = \frac{S_{i-1} - S_i}{S_{i-1}}$.

If we assume that the transition probability from S to I (u) can be approximated with the average of u_i , we get that $u = \frac{\sum_{i=1}^k u_i}{k}$.

Derivation of v

R_i is the number of removed individuals at time $t = i$ and R_j at time $t = j$. From the assumptions made earlier in (3.2.1), we know that

$$(R_i \leq R_j) \forall i, j \in (0, 1, 2, \dots, 60) \wedge i \leq j.$$

$(R_j - R_i)$ is the total number of individuals who moved from the I category to the R category during the time interval i to j .

This follows from the assumptions made earlier in (3.2.1).

The proportion of individuals who transitioned from I to R during this time interval is equal to $\frac{R_j - R_i}{I_i}$.

We define v_i as the number of individuals who transition from I to R during the time interval $i - 1$ to i proportional to the number of individuals in I at time $t - 1$, so $v_i = \frac{R_i - R_{i-1}}{I_{i-1}}$.

If we assume that the transition probability from I to R (v) can be approximated with the average of v_i , we get that $v = \frac{\sum_{i=1}^k v_i}{k}$.

If you look at both derivations above, you can see that we only used the difference in the R and S daily. The reason for doing this was because when we looked at the data, we saw that the R and S increased and decreased in numbers, respectively. The reason why we did not use the daily difference of the I compartment is because the numbers of this compartment both increased and decreased over time. Using the daily difference of this compartment would lead to negative transition probabilities, which is not allowed.

After deriving the calculations to determine the transition probabilities, the next step was constructing the models. We used the same intervals we used earlier for the model with differential equations. In the next three subsections we present the model for each of the different lengths.

3.2.1 Stochastic Model version 1 (30-day average)

The first version of our model calculated the transition probability using a length of 30 days. In this case we took

$$\begin{aligned} u_t &= \frac{S_{t-1} - S_t}{S_{t-1}} ; t \in (1, 2, \dots, 30), u = \frac{\sum_{t=1}^{30} u_t}{30} \\ v_t &= \frac{R_t - R_{t-1}}{I_{t-1}} ; t \in (1, 2, \dots, 30), v = \frac{\sum_{t=1}^{30} v_t}{30} \end{aligned} \quad (3.2.3)$$

These calculations resulted in the following matrix P :

	S	I	R
S	0.99987	0.00013	0
I	0	0.91775	0.08225
R	0	0	1

After estimating this probability transition matrix, we used it to calculate the fitted values for each compartment with the initial condition equal to the data observed on August 1st using (3.1.3). We also calculated the error terms using (3.1.4). Calculating the fitted values was done in the following way:

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 1st.
 The fitted values then become:

$$\begin{bmatrix} S_f_t \\ I_f_t \\ R_f_t \end{bmatrix} = (P^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix} ; t = 1, 2, \dots, 30$$

After calculating the fitted values and the error terms, we compared the fitted values with the observed values to determine how well the model fitted to the observed values.

3.2.2 Stochastic Model version 2 (14-day average)

In our next version, we calculated the transition probability using a length of 14 days. In this case we took

$$\begin{aligned}
 u(1)_t &= \frac{S_{t-1}-S_t}{S_{t-1}} ; t \in (1,2, \dots,14), u(1) = \frac{\sum_{t=1}^{14} u(1)_t}{14} \\
 v(1)_t &= \frac{R_t-R_{t-1}}{I_{t-1}} ; t \in (1,2, \dots,14), v(1) = \frac{\sum_{t=1}^{14} v(1)_t}{14} \\
 u(2)_t &= \frac{S_{t-1}-S_t}{S_{t-1}} ; t \in (15,16, \dots,28), u(2) = \frac{\sum_{t=15}^{28} u(2)_t}{14} \\
 v(2)_t &= \frac{R_t-R_{t-1}}{I_{t-1}} ; t \in (15,16, \dots,28), v(2) = \frac{\sum_{t=15}^{28} v(2)_t}{14} \\
 u(3)_t &= \frac{S_{t-1}-S_t}{S_{t-1}} ; t \in (29,30), u(3) = \frac{\sum_{t=29}^{30} u(3)_t}{2} \\
 v(3)_t &= \frac{R_t-R_{t-1}}{I_{t-1}} ; t \in (29,30), v(3) = \frac{\sum_{t=29}^{30} v(3)_t}{2}
 \end{aligned} \tag{3.2.4}$$

These calculations resulted in 3 different transition probability matrices, one for each 14-day length part of the data. The last matrix was calculated using only 2 data points. This, because these two data points were left after calculating the first two matrices. We got the following matrices $P(1)$, $P(2)$ and $P(3)$:

$P(1)$

	S	I	R
S	0.99985	0.00015	0
I	0	0.9169	0.0831
R	0	0	1

$P(2)$

	S	I	R
S	0.99988	0.00012	0
I	0	0.91923	0.08077
R	0	0	1

$P(3)$

	S	I	R
S	0.99993	0.00007	0
I	0	0.91923	0.08077
R	0	0	1

After estimating these matrices, we calculated the fitted values for each compartment. Calculating the fitted values was done in the following way:

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 1st.

The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(1)^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1, 2, \dots, 14$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 15th.

The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(2)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1, 2, \dots, 14 \text{ and } t = i + 14$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 29th.

The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(3)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1, 2 \text{ and } t = i + 28$$

Just like in the previous subsection, we compared the fitted values with the observed values.

3.2.3 Stochastic Model version 3 (7-day average)

The last version of the model we constructed, had transition probabilities that were calculated using an average of 7 days.

$$\begin{aligned} u(1)_t &= \frac{S_{t-1} - S_t}{S_{t-1}}; t \in (1, 2, \dots, 7), u(1) = \frac{\sum_{t=1}^7 u(1)_t}{7} \\ v(1)_t &= \frac{R_t - R_{t-1}}{I_{t-1}}; t \in (1, 2, \dots, 7), v(1) = \frac{\sum_{t=1}^7 v(1)_t}{7} \\ u(2)_t &= \frac{S_{t-1} - S_t}{S_{t-1}}; t \in (8, 9, \dots, 14), u(2) = \frac{\sum_{t=8}^{14} u(2)_t}{7} \\ v(2)_t &= \frac{R_t - R_{t-1}}{I_{t-1}}; t \in (8, 9, \dots, 14), v(2) = \frac{\sum_{t=8}^{14} v(2)_t}{7} \\ u(3)_t &= \frac{S_{t-1} - S_t}{S_{t-1}}; t \in (15, 16, \dots, 21), u(3) = \frac{\sum_{t=15}^{21} u(3)_t}{7} \\ v(3)_t &= \frac{R_t - R_{t-1}}{I_{t-1}}; t \in (15, 16, \dots, 21), v(3) = \frac{\sum_{t=15}^{21} v(3)_t}{7} \\ u(4)_t &= \frac{S_{t-1} - S_t}{S_{t-1}}; t \in (22, 23, \dots, 28), u(4) = \frac{\sum_{t=22}^{28} u(4)_t}{7} \\ v(4)_t &= \frac{R_t - R_{t-1}}{I_{t-1}}; t \in (22, 23, \dots, 28), v(4) = \frac{\sum_{t=22}^{28} v(4)_t}{7} \\ u(5)_t &= \frac{S_{t-1} - S_t}{S_{t-1}}; t \in (29, 30), u(5) = \frac{\sum_{t=29}^{30} u(5)_t}{2} \\ v(5)_t &= \frac{R_t - R_{t-1}}{I_{t-1}}; t \in (29, 30), v(5) = \frac{\sum_{t=29}^{30} v(5)_t}{2} \end{aligned} \tag{3.2.5}$$

These calculations resulted in 5 different transition probability matrices, one for each 7-day length part of the data. The last matrix was calculated using only two data points. This, because these two data points were left after calculating the first 4 matrices. We got the following matrices $P(1)$, $P(2)$, $P(3)$, $P(4)$ and $P(5)$:

$P(1)$	S	I	R
S	0.99987	0.00013	0
I	0	0.91234	0.08766
R	0	0	1

$P(2)$	S	I	R
S	0.99984	0.00016	0
I	0	0.92146	0.07854
R	0	0	1

$P(3)$	S	I	R
S	0.99985	0.00015	0
I	0	0.91149	0.08851
R	0	0	1

$P(4)$	S	I	R
S	0.99991	0.00009	0
I	0	0.92697	0.07303
R	0	0	1

$P(5)$	S	I	R
S	0.99993	0.00007	0
I	0	0.91339	0.08661
R	0	0	1

After estimating these matrices, we calculated the fitted values for each compartment. We also calculated the error terms. Calculating the fitted values was done in the following way:

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 1st.
 The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(1)^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1, 2, \dots, 7$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 8th.
 The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(2)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1, 2, \dots, 7 \text{ and } t = i + 7$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 15th.
 The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(3)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1, 2, \dots, 7 \text{ and } t = i + 14$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 22nd.

The fitted values then become:

$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(4)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1, 2, \dots, 7 \text{ and } t = i + 21$$

At time $t = 0$ we have S_0, I_0 and R_0 ; the data on August 29th.

The fitted values then become:

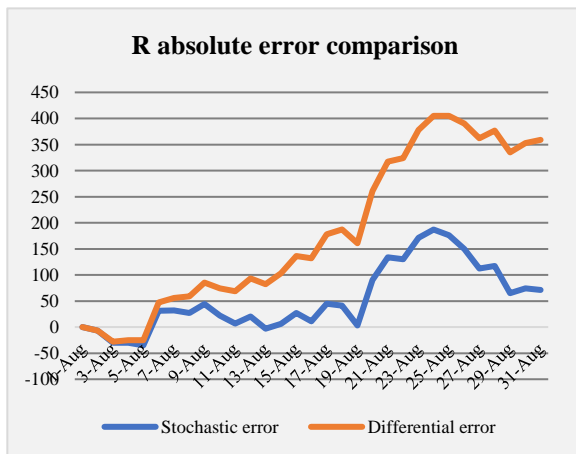
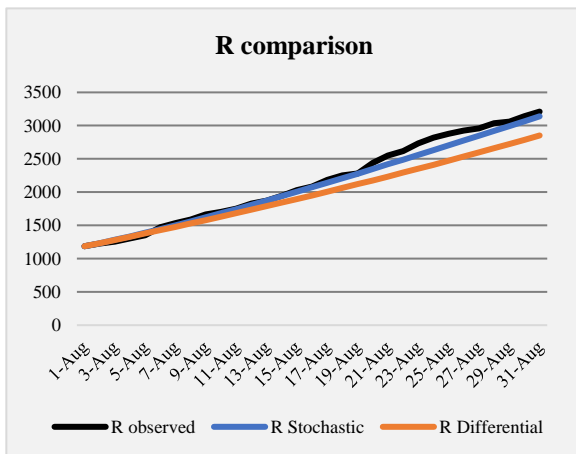
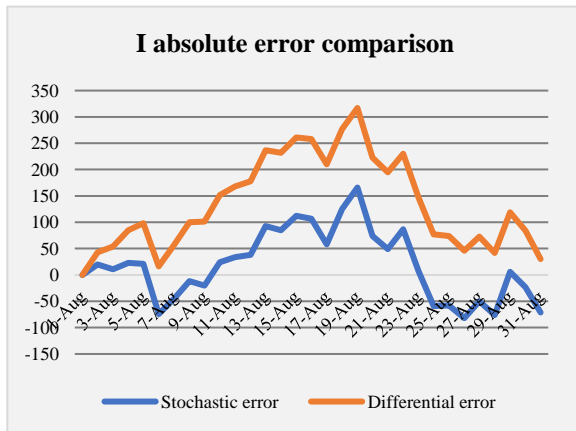
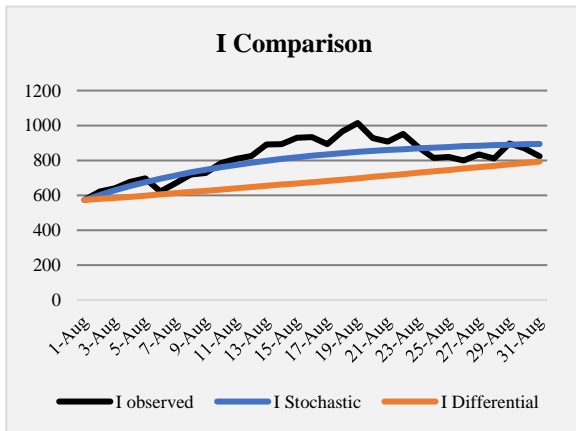
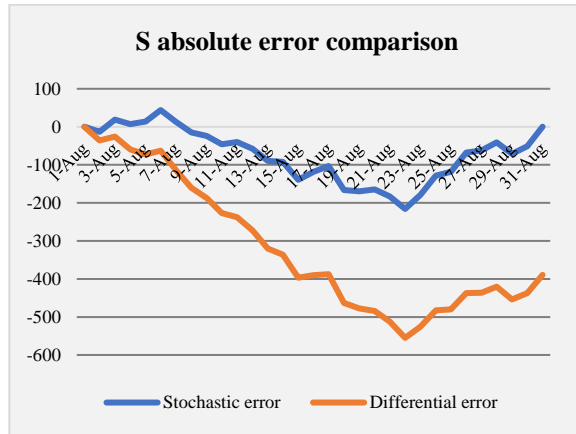
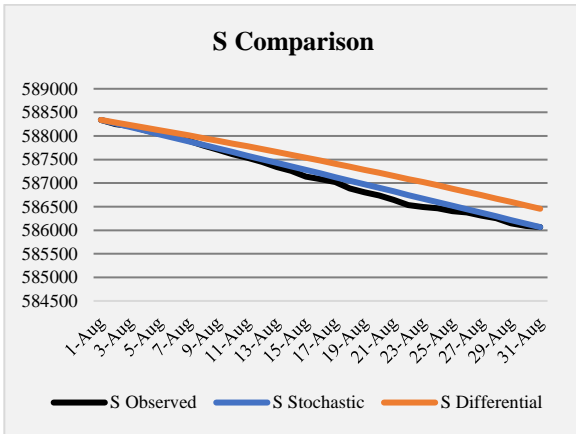
$$\begin{bmatrix} Sf_t \\ If_t \\ Rf_t \end{bmatrix} = (P(5)^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1, 2 \text{ and } t = i + 28$$

After constructing these versions, we compared the fitted values from both the stochastic models and the differential models. The result from this comparison is presented in the next chapter.

4 Results

4.1 30-day average comparison

In this chapter, we compare the fitted values from both models using 30-day average for each model. The graphs below show the result of this comparison.



If we start with the S compartment (first graph on the left), we can see that for the first 7 days of the data, both models had a good fit to the observed data. After these days, we can see that the stochastic model had a better fit compared to the differential model. If we look at the first graph on the right side of the previous page, we can see the error for both models. This graph is a better view of the fit for both models. The model based on differential equations has error terms ranging from -555 to 0, whereas the error for the stochastic compartment model has a range of -216 to 44. The latter is a smaller range than the former.

The second graph on the left shows the comparison for the I compartment. In this case we can see that neither model has the same structure as the observed values. We can see that the graph of the stochastic compartmental model lies closer to the observed values than the graph of the differential equation model. If we look at the graph next to this one, we can see that the error terms for the stochastic compartmental model lies between -82 and 166. The error terms of the differential equation model have a range of 0 to 317. Just like with the S compartment, the range of the error terms of the I compartment is smaller for the stochastic compartmental model.

The last graph on the left shows that for the first 13 to 14 days of the data, the fitted values from both models have a good fit to the observed values. After these days we see that the stochastic compartmental model has a better fit than the model with differential equations. If we look at the error terms in the graph to the right, we see that the error terms of the differential equation model have a range of -28 to 405. And the error terms of the stochastic compartmental model have a range of -35 to 187. Here we see that the range is smaller for the error terms from the stochastic compartmental model.

To further analyze the comparison between both models we used the mean (μ) and the standard deviation (σ) of the error terms. Before presenting our findings, we define the following variables:

$$\begin{aligned}
 \mu_S &= \overline{Se} = \frac{\sum_{i=0}^{31} |Se_i|}{31}, \text{ the mean for the error terms of the } S \text{ compartment.} \\
 \mu_I &= \overline{Ie} = \frac{\sum_{i=0}^{31} |Ie_i|}{31}, \text{ the mean for the error terms of the } I \text{ compartment.} \\
 \mu_R &= \overline{Re} = \frac{\sum_{i=0}^{31} |Re_i|}{31}, \text{ the mean for the error terms of the } R \text{ compartment.} \\
 \sigma_S &= \sqrt{\frac{1}{31} \sum_{i=0}^{31} (|Se_i| - \mu_S)^2}, \text{ the standard deviation of the error terms in } S. \\
 \sigma_I &= \sqrt{\frac{1}{31} \sum_{i=0}^{31} (|Ie_i| - \mu_I)^2}, \text{ the standard deviation of the error terms in } I. \\
 \sigma_R &= \sqrt{\frac{1}{31} \sum_{i=0}^{31} (|Re_i| - \mu_R)^2}, \text{ the standard deviation of the error terms in } R.
 \end{aligned} \tag{4.1.1}$$

These calculations were also used for the comparison of the 14-day average and 7-day average versions of both models. The calculations done for μ and σ were done in Microsoft Excel. In the case of the 30-day average version we got the following results.

(4.1.2) mean (μ) and the standard deviation (σ) error terms (30-day average).

Model	S		I		R	
	μ_S	σ_S	μ_I	σ_I	μ_R	σ_R
Differential	317	172	135	87	188	142
Stochastic	79	62	55	39	61	56

Using this table, we could theorize that the stochastic model behaves better than the differential model. If we look at it, we can see that for each compartment the mean and the standard deviation of the stochastic compartmental model error is smaller than those of the differential equation model. This could also be seen in the graphs at the beginning. If we look at the 30-day average for both models, we can say that the stochastic compartmental model has a closer fit to the observed values compared to the model based on a system of differential equations.

The purpose of this thesis was to determine whether the stochastic compartmental model gave similar results as a differential model. So, this means that the difference of the error means from both models must be statistically insignificant. To accurately determine this, we use a two-sample t-test to determine whether the difference between the mean error from both models is significant. For this test, we used a confidence level of 0.95. The method used for these calculations can be found at Stat Trek (Hypothesis Test: Difference Between Means, 2021) and calculations were done using R studio.

The test was done separately for each compartment. The first step was stating the test hypothesis for each compartment. In this case we stated the null hypothesis and the alternative hypothesis in the following way:

S compartment

$$\text{Null hypothesis: } \overline{S_{e.diff}} - \overline{S_{e.stoch}} = 0$$

$$\text{Alternative hypothesis: } \overline{S_{e.diff}} - \overline{S_{e.stoch}} \neq 0$$

I compartment

$$\text{Null hypothesis: } \overline{I_{e.diff}} - \overline{I_{e.stoch}} = 0$$

$$\text{Alternative hypothesis: } \overline{I_{e.diff}} - \overline{I_{e.stoch}} \neq 0$$

R compartment

$$\text{Null hypothesis: } \overline{R_{e.diff}} - \overline{R_{e.stoch}} = 0$$

$$\text{Alternative hypothesis: } \overline{R_{e.diff}} - \overline{R_{e.stoch}} \neq 0$$

Each hypothesis was tested using a confidence level of 0.95 which puts the significance level α at 0.05. If the p-value of the test is smaller than α we should reject the null hypothesis. For our thesis hypothesis to be true, we would have to accept the null hypothesis. The test produced the following results for the p-value:

$$S \text{ compartment p-value} = 2.844e^{-6}$$

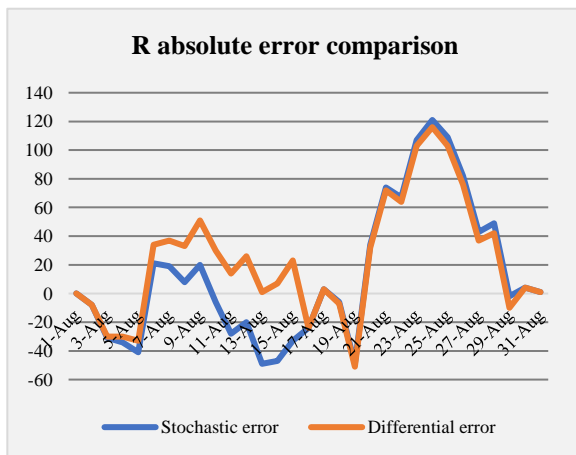
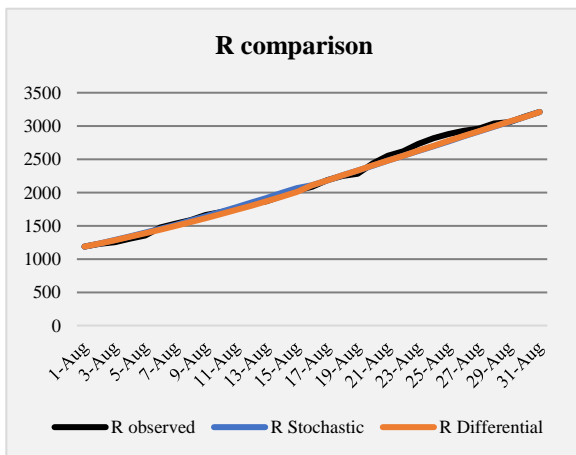
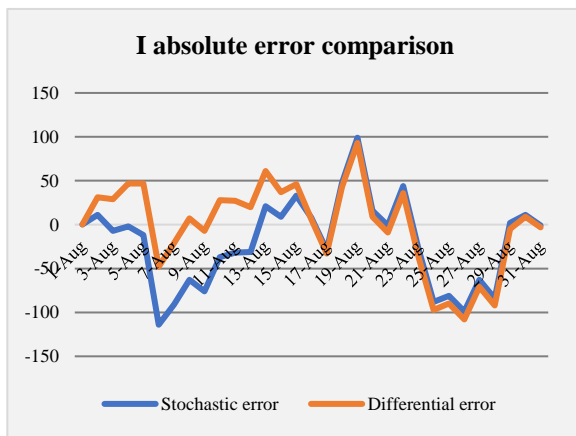
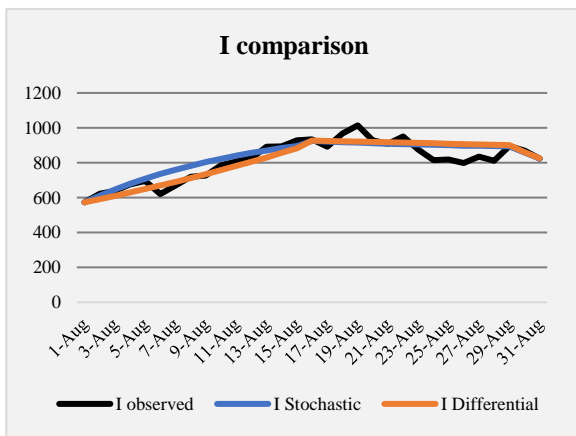
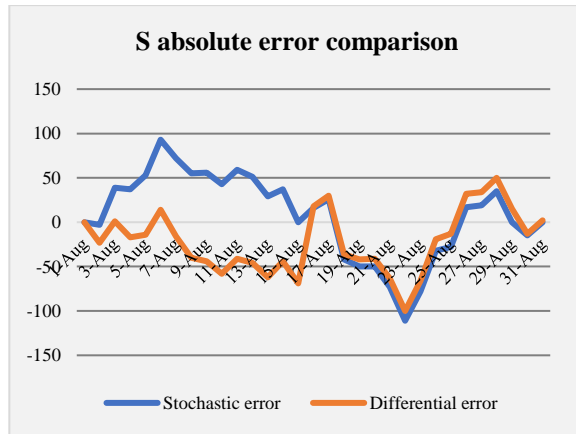
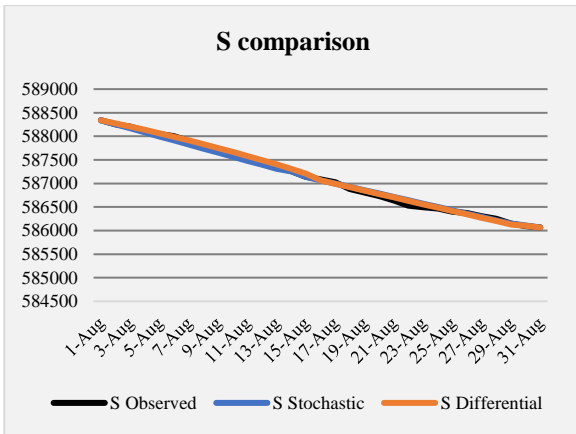
$$I \text{ compartment p-value} = 1.02e^{-7}$$

R compartment p-value = $4.43e^{-5}$

As you can see, for each compartment, the p-value is smaller than α , so this means that we should reject the null hypothesis and accept the alternative hypothesis. This means that the difference between the mean error of the models is statistically significant. This means that for the 30-day average versions of the models, the difference is too significant to determine whether the stochastic model behave as well as the differential model. In the next section we present the comparison of the 14-day version for each model.

4.2 14-day average comparison

In this chapter, we compare the fitted values from both models using 14-day average for each model. The graphs below show the result of this comparison.



If we start with the first graph on the left, the S compartment, we can see that both models have an almost perfect fit to the observed values. If we look at the graph on the left, we see that error terms of the stochastic compartmental model ranges from -111 to 93 and those of the differential equation model has a range of -100 to 50. In this case we see that the differential equation model has a smaller range than the stochastic compartmental model

The second graph on the left shows the comparison for the I compartment. In this case we can see that for the first 15 days, the differential model has a better fit than the stochastic model. And for the rest of the data, we see that both models have the same fit to the observed values. neither model has the same structure as the observed values. If we look at the error terms, we can see that the error for the stochastic compartmental model ranges from -114 to 99 and for the differential model the range is -108 to 93. In this case the error terms of the differential model have a smaller range.

The last graph on the left shows that the fitted values from both models have a good fit to the observed values. Both models have the same fitted values compared to the observed values. If we look at the error terms in the graph to the right, we see that the error terms of the differential equation model have a range of -51 to 116. And the error terms of the stochastic compartmental model have a range of -49 to 121. Here we see that the range is slightly smaller for the error terms from the stochastic compartmental model.

We calculated the mean (μ) and the standard deviation (σ) of the error terms just like in the previous section using (4.1.1). These calculations gave us the following results.

(4.2.1) mean (μ) and the standard deviation (σ) error terms (14-day average).

Model	S		I		R	
	μ_S	σ_S	μ_I	σ_I	μ_R	σ_R
Differential	34	32	39	31	36	31
Stochastic	39	27	40	55	37	33

If we look at the table above, we can see that for each compartment the mean and the standard deviation of the stochastic compartmental model error and those of the differential equation model are close to each other. We see that the differential model has smaller mean error values for each compartment. But the stochastic model has smaller variance for only one compartment. In this case we can say that looking at the 14-day average for both models, the differential model has a closer fit to the observed values compared to the stochastic model.

To accurately determine which whether the stochastic model behave as well as the differential model, we did a t-test just like in the previous section. In this case the null hypothesis and the alternative hypothesis become:

S compartment

Null hypothesis: $\overline{S_{e.diff}} - \overline{S_{e.stoch}} = 0$

Alternative hypothesis: $\overline{S_{e.diff}} - \overline{S_{e.stoch}} \neq 0$

I compartment

Null hypothesis: $\overline{I_{e.diff}} - \overline{I_{e.stoch}} = 0$

Alternative hypothesis: $\overline{I_{e.diff}} - \overline{I_{e.stoch}} \neq 0$

R compartment

Null hypothesis: $\overline{R_{e.diff}} - \overline{R_{e.stoch}} = 0$

Alternative hypothesis: $\overline{R_{e.diff}} - \overline{R_{e.stoch}} \neq 0$

Each hypothesis was tested using a confidence level of 0.95 which puts the significance level α at 0.05. The test produced the following results for the p-value:

S compartment p-value = 0.008133

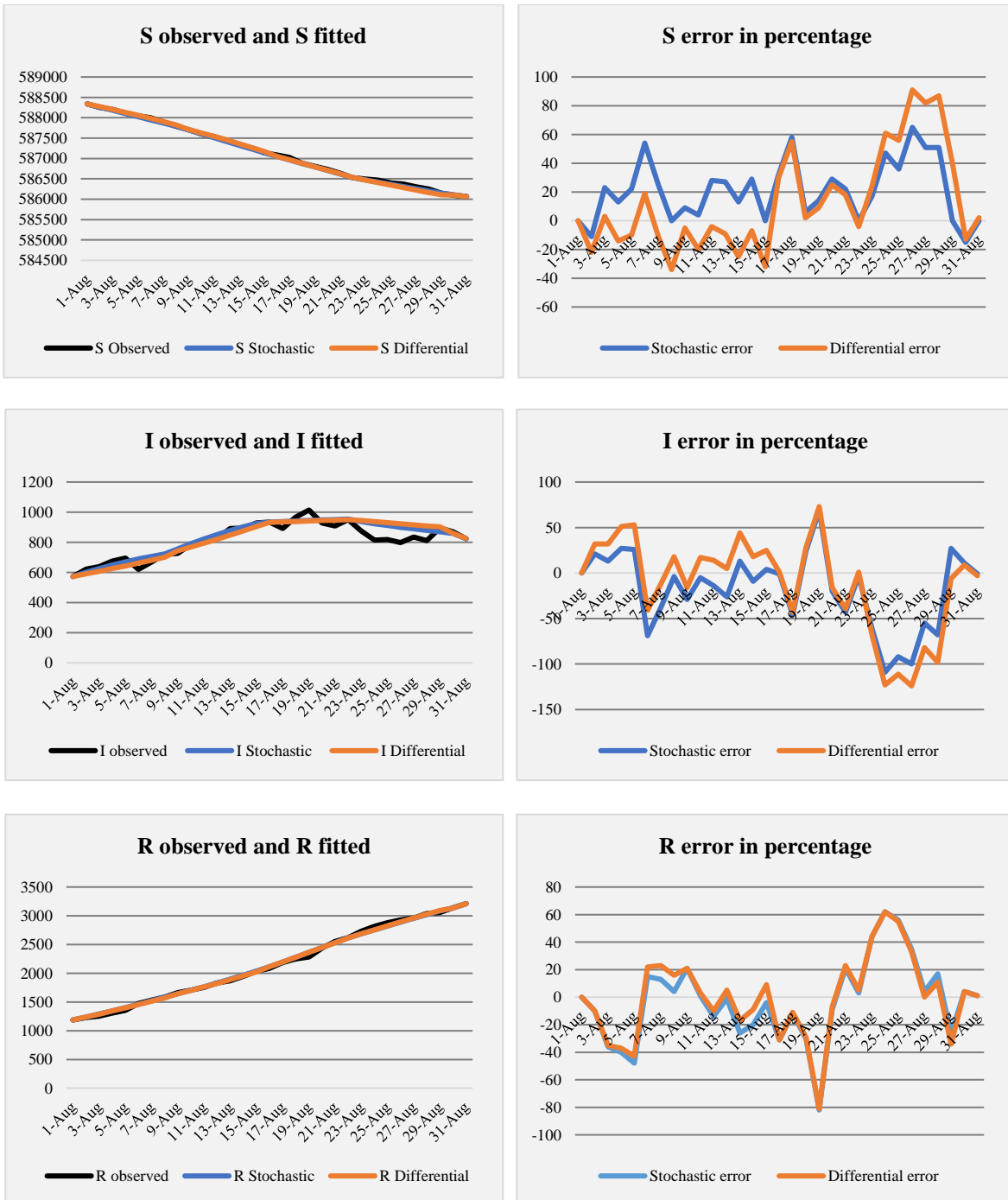
I compartment p-value = 0.1517

R compartment p-value = 0.3607

As you can see that except for the *S* compartment, each compartment has a p-value that is larger than α , so this means that we should accept the null hypothesis. This means that the difference between the mean error of the models is not statistically significant. This means that for the 14-day average versions of the models, the difference is not significant enough and thus, the stochastic model behaves statistically the same way as the differential model for the two compartments. The *S* compartment has a p-value that is smaller than α , which means that for this compartment the difference between the error for the stochastic model and the differential model is statistically significant. Since not all compartments have the same test results, we cannot say with absolute certainty that the stochastic model behaves the same as the differential model. In the next section we present the comparison of the 7-day version for each model.

4.3 7-day average comparison

In this chapter, we compare the fitted values from both models using 14-day average for each model. The graphs below show the result of this comparison



If we start with the first graph on the left, the S compartment, we can see that both models have an almost perfect fit to the observed values. If we look at the graph on the left, we see that error terms of the stochastic compartmental model ranges from -15 to 65 and those of the differential equation model has a range of -34 to 91. In this case we see that the stochastic model has a smaller range than the differential model.

The second graph on the left shows the comparison for the I compartment. In this case, just as the 14-day average comparison, we can see that for the first 15 days, the differential model has a better fit than the stochastic model. And for the rest of the data, we see that both models have the same fit to the observed values. Neither model has the same structure as the observed values. If we look at the error terms, we can see that the error for the stochastic compartmental model ranges from -109 to 69 and for the differential model the range is -124 to 73. In this case the error terms of the stochastic model have a smaller range.

The last graph on the left shows that the fitted values from both models have a good fit to the observed values. Both models have the same fitted values compared to the observed values. If we look at the error terms in the graph to the right, we see that the error terms of the differential equation model have a range of -81 to 62. And the error terms of the stochastic compartmental model have a range of -82 to 62. Here we see that the range is slightly smaller for the error terms from the differential model.

We calculated the mean (μ) and the standard deviation (σ) of the error terms just like in the previous sections using (4.1.1). These calculations gave us the following results.

(4.3.1) mean (μ) and the standard deviation (σ) error terms (7-day average).

Model	S		I		R	
	μ_S	σ_S	μ_I	σ_I	μ_R	σ_R
Differential	26	25	39	36	22	19
Stochastic	23	19	33	30	22	20

If we look at the table above, we can see that for each compartment the mean and the standard deviation of the stochastic compartmental model error and those of the differential equation model are close to each other. We see that the stochastic model has smaller mean error values for each compartment, except for the R compartment. In this case the mean error is the same for both models. The stochastic model has smaller variance for S and I . In this case we can say that looking at the 7-day average for both models, the stochastic model has a slightly closer fit to the observed values compared to the differential model.

To accurately determine which whether the stochastic model behaves as well as the differential model, we did a t-test just like in the previous sections. In this case the null hypothesis and the alternative hypothesis become:

S compartment

Null hypothesis: $\overline{S_{e.diff}} - \overline{S_{e.stoch}} = 0$

Alternative hypothesis: $\overline{S_{e.diff}} - \overline{S_{e.stoch}} \neq 0$

I compartment

Null hypothesis: $\overline{I_{e.diff}} - \overline{I_{e.stoch}} = 0$

Alternative hypothesis: $\overline{I_{e.diff}} - \overline{I_{e.stoch}} \neq 0$

R compartment

Null hypothesis: $\overline{R_{e.diff}} - \overline{R_{e.stoch}} = 0$

Alternative hypothesis: $\overline{R_{e.diff}} - \overline{R_{e.stoch}} \neq 0$

Each hypothesis was tested using a confidence level of 0.95 which puts the significance level α at 0.05. The test produced the following results for the p-value:

S compartment p-value = 0.2835

I compartment p-value = 0.59

R compartment p-value = 0.7699

As you can see each compartment has a p-value that is larger than α , so this means that we should accept the null hypothesis. This means that the difference between the mean error of the models is not statistically significant. This means that for the 7-day average versions of the models, the difference is not significant enough and thus, the stochastic model behaves statistically the same way as the differential model. So in for this version of the models, we have proven the hypothesis that the stochastic model behaves just as well as the differential model.

In two out of the three versions (14-day average and 7-day average), we saw that the stochastic model had a slightly better fit than the differential model. This observation gave the idea to go a step further with this research and determine how a stochastic model could be used to predict future data for COVID-19. We decided to use the 7-day average version of the stochastic compartmental model because this model had an overall closer fit to the observed values. Our findings are presented in the next section of this chapter.

4.4 Future predictions

After the construction of the model in the previous section, we used the second part of the data from September to determine how we could use this model to predict the future number of cases for each compartment. In the previous section we decided to use the version of the stochastic model based on 7-day average. Now we had to determine was how many days the model could make predictions for. We used the average of the most recent 7 days to predict the data for the next couple of days. To determine how many days ahead we could make predictions for we started with 30 days. We ran the model for 30 iterations, so it calculated the predicted numbers for every day in September. For this part of the process, we had to define the following variables:

- Sp : The predicted values for the S compartment.
 - Ip : The predicted values for the I compartment.
 - Rp : The predicted values for the R compartment.
 - Sd : The relative difference for the S compartment in percentage.
 - Id : The relative difference for the I compartment in percentage.
 - Rd : The relative difference for the R compartment in percentage.
- (4.4.1)

- $Sd_i = S_i - Sp_i, i = 1,2, \dots, 30$
 - $Id_i = I_i - Ip_i, i = 1,2, \dots, 30$
 - $Rd_i = R_i - Rp_i, i = 1,2, \dots, 30$
- (4.4.2)

This was done the following way:

With time $t = 0$, on August 31st, t in days we have S_0, I_0 and R_0 .
 The transition probability matrix P is given by:

	S	I	R
S	$1 - u$	u	0
I	0	$1 - v$	v
R	0	0	1

With:

$$u = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1,2, \dots, 30$$

After this we calculated the predicted values and compared them to the observed values and got the following results:



For these graphs we look at the observed values of September with the predictive values retrieved from the model to compare these two values. Next to each compartment we also plotted the difference between the observed and the predicted values.

When we look at the graphs containing the difference between the observed values and the predicted values, we see that for the S category, the difference started to increase after the 5th iteration. The difference in the other two compartments started increasing after later iterations. So based on S we ran the model again and after every 5th iteration we updated the transition probability matrix with transition probabilities calculated using the seven most recent observed data before that iteration, like this:

Prediction for September 1st through September 5th:

With time $t = 0$, on August 31st, t in days we have S_0, I_0 and R_0 .

The transition probability matrix $P(1)$ has the probabilities:

$$u(1) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(1) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; t = 1, 2, \dots, 5$$

Prediction for September 6th through September 10th:

With time $t = 0$, on September 5th, t in days we have S_0, I_0 and R_0 .

The transition probability matrix $P(2)$ has the probabilities:

$$u(2) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(2) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1, 2, \dots, 5 \text{ and } t = i + 5$$

Prediction for September 11st through September 15th:

With time $t = 0$, on September 10th, t in days we have S_0, I_0 and R_0 .

The transition probability matrix $P(3)$ has the probabilities:

$$u(3) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(3) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^t)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1,2, \dots, 5 \text{ and } t = i + 10$$

Prediction for September 16th through September 20th:

With time $t = 0$, on September 15th, t in days we have S_0, I_0 and R_0 .

The transition probability matrix $P(4)$ has the probabilities:

$$u(4) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(4) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1,2, \dots, 5 \text{ and } t = i + 15$$

Prediction for September 21st through September 25th:

With time $t = 0$, on September 20th, t in days we have S_0, I_0 and R_0 .

The transition probability matrix $P(5)$ has the probabilities:

$$u(5) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(5) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1,2, \dots, 5 \text{ and } t = i + 20$$

Prediction for September 26th through September 30th:

With time $t = 0$, on September 25th, t in days we have S_0, I_0 and R_0 .

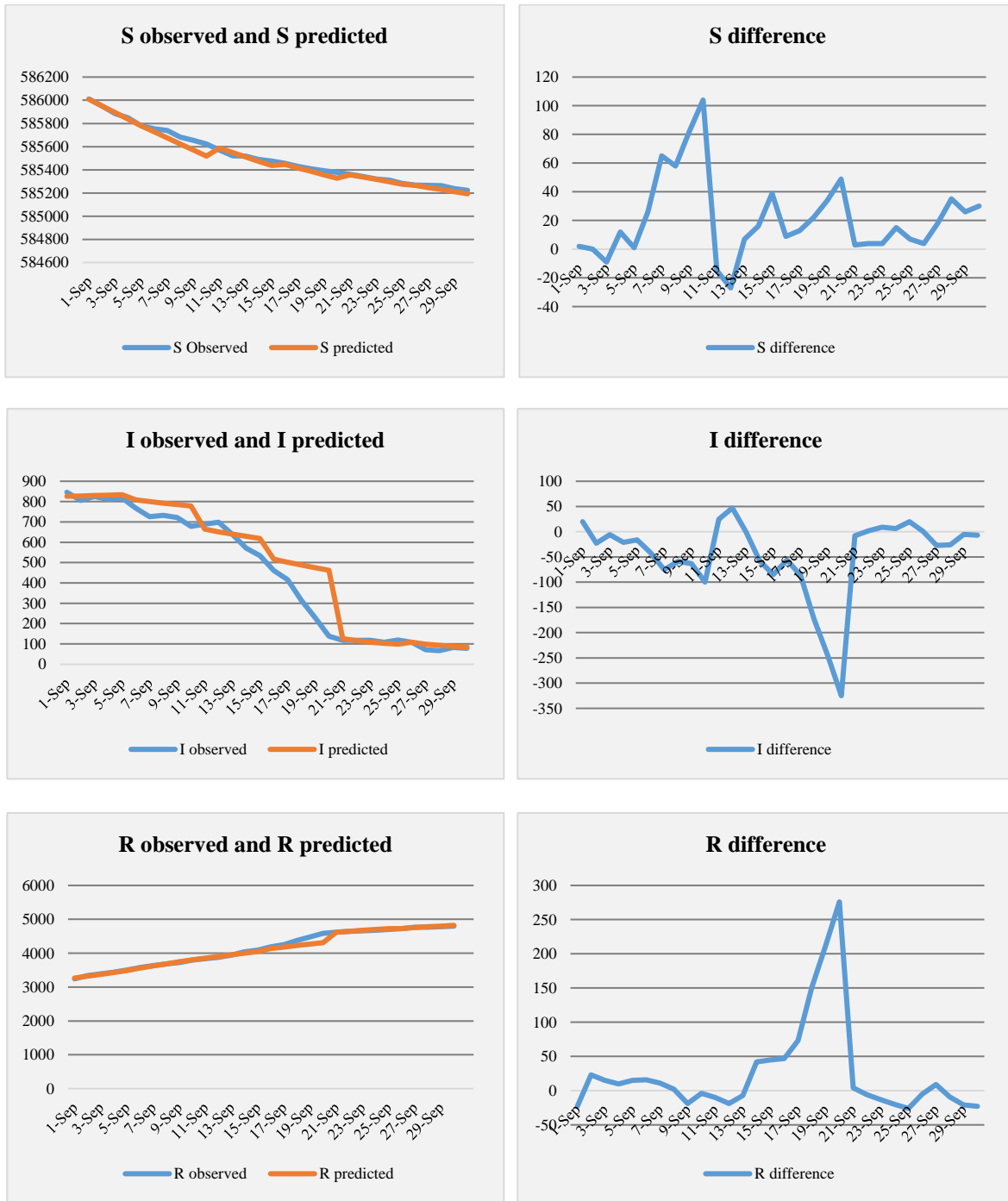
The transition probability matrix $P(6)$ has the probabilities:

$$u(6) = \frac{\sum_{j=t-6}^0 \frac{S_{j-1} - S_j}{S_{j-1}}}{7}$$

$$v(6) = \frac{\sum_{j=t-6}^0 \frac{R_j - R_{j-1}}{I_{j-1}}}{7}$$

$$\begin{bmatrix} Sp_t \\ Ip_t \\ Rp_t \end{bmatrix} = (P^i)^T \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}; i = 1,2, \dots, 5 \text{ and } t = i + 25$$

After running the model and adding new transition probabilities after every 5th iteration we got the following results:



After we got these results, we evaluated the results by looking at the mean (μ) and standard deviation (σ) of the difference between the observed values and the predicted values. The calculation for the mean and standard deviation of the difference terms was done in the same way as that of the error terms in the previous sections. Those results are visible in the table on the next page.

(4.4.3) mean (μ) and the standard deviation (σ) difference.

	<i>S</i>	<i>I</i>	<i>R</i>
μ	25	55	39
σ	25	73	62

The table above shows the mean and standard deviation of the difference between the observed values and the predicted values. Starting with *S*, we can see that the mean for the difference between the observed values and predicted values is equal to 25 with a standard deviation of 25. This means that if we run the prediction model, the mean difference between the predicted values and the actual values will be equal to 25 with a variation of 25, in other words the predicted number of susceptible people may be about 50 more or less than the actual number of susceptible individuals. For the *I* compartment we can see that the mean and the standard deviation is much larger than the other two compartments. If we run the prediction model the difference between the actual values and predicted values for the *I* compartment will be about 55 with a standard deviation of 73. For the last compartment, the *R* compartment we see that the mean difference between the predicted values and actual values is equal to 39 with a standard deviation of 62.

This concludes the construction of a mathematical model. In this chapter we focused on the construction of the model and the comparison between the stochastic model and the differential model and in the next chapter we will present the conclusions and recommendations for this thesis.

5 Conclusion

In this chapter we will present the conclusions from this study. The first thing we concluded from the results is that a stochastic compartmental model does approximate the COVID-19 outbreak in Suriname as well as a model based on differential equations. Starting with the construction of the differential model the first thing we did was estimate parameters by calculating the average rates over 30, 14 and 7 days. After doing this, we did the same thing for the stochastic compartmental model. We then used the graphs in chapter 4 to compare the two models with each other.

In Section 4.1 we saw that overall, the stochastic model had a better fit than the differential model when we looked at the graphs. And doing a t-test showed that the difference between the two models was statistically significant. So based on this, our hypothesis was not proven, when it came to using a 30-day average for the transition probabilities and the parameters. However, we believe that had the hypothesis been to determine whether the stochastic model behaved better than the differential model, these test results would have a different interpretation. But since we wanted to prove that there was no significant difference between the two models, these results had a different meaning.

In section 4.2 we saw that overall, the differential model had a better fit than the stochastic model. The t-test proved for two out of three compartments that the difference between the two models was not significant enough, meaning there was no significant difference between the results of these models. When it came to the 14-day average, our hypothesis was correct.

In section 4.3 we saw that both models had the same fit to the observed values. The stochastic model had a slightly better fit than the differential model. The t-test proved that for each compartment the stochastic model had the same results as the differential model. Looking back at the results from the first three sections of chapter 4, we can say that in two out of the three versions our hypothesis was correct. Using a stochastic compartmental model to approximate the COVID-19 outbreak in Suriname gave the same results as using a compartmental model based on differential equations.

However, it is possible that the fit of the stochastic model could be better if we had access to the actual data from the COVID-19 management team. As discussed in chapter 2, the data we retrieved from the website was not clean enough to use for our model. We believe that some of the error in both models was a direct consequence of the method we used in cleaning up the data.

In the fourth section of chapter 4 we looked at the possibility of using the model to predict future values. Here we saw that running the model for 30 days with the same transition probabilities gave us predictions that were far from the actual values. This, we believe it is also the result of a variance in the actual transition probabilities. Assuming that one transition probability could be used for a period of 30 days is a wrong assumption on our part. After seeing this we decided to rerun the model for 30 days but this time we added in a probability update

after every 5th iteration and in this case, we saw that the predicted values were closer to the actual values. This could be because every five days the actual transition probabilities of COVID-19 in Suriname change, which could be a direct consequence of preventive measures implemented by policy makers and the quality of the medical care in Suriname.

We saw that the overall predicted number of removed individuals, which consists of the deceased and recovered cases, was predicted with a mean difference of 39 individuals. For this category, such a number is relatively small because this category has numbers in thousands. On the other hand, the overall prediction of the infected individuals had a mean difference of 55 individuals. This number is relatively large because this category had numbers in hundreds. 39 in relation to a thousand is smaller than 55 in relation to a couple of hundred. We can see that the biggest difference for the infected is around September 20th, this suggests that the actual transition probability on that day was different than that of the week before. So, the assumption that predictions can be done based on the data from the previous week, is not accurate for every date. However, we believe that our model performs well enough to be able to predict the future behavior of COVID-19 Suriname. This in turn could be useful for developing and implementing preventive measures.

We believe that with the right data we could expand this model with more compartments. For example, the infected compartment could be divided in isolation and hospitalized. The latter can be divided nursing ward and ICU. The remove category could also be divided in recovered and deceased. We believe that with this expansion, the model would make more accurate predictions. It should be noted that we had intended to create a model containing all the compartments named previously, but due to the lack of detailed data this could not be achieved.

We believe that given more detailed data, a more accurate model could be constructed using the methods used in this research. This model could in turn be used to predict future numbers for each compartment. After which these numbers could indicate how many resources are needed for the epidemic.

Another thing to consider is that at the time this research concluded, a vaccine had already been developed and distributed to various individuals. For future research we recommend that an additional compartment is added which consists of vaccinated individuals. This compartment will connect directly to the susceptible compartment and the probability that an individual leaves the susceptible group and goes to the vaccinated group will be equal to the average fraction of people vaccinated daily. For a more detailed model the vaccinated compartment could be divided in individuals who received one dosage and individuals who received both dosages.

Appendices

A. Data

Section 4.1 30-day average comparison

Fitted values and error for differential model. (30-day average)

Date	<i>Sf</i>	<i>Se</i>	<i>If</i>	<i>Ie</i>	<i>Rf</i>	<i>Re</i>
01-08	588340	0	573	0	1187	0
02-08	588287	-36	579	43	1234	-7
03-08	588233	-26	585	54	1282	-28
04-08	588178	-59	591	85	1330	-25
05-08	588123	-73	598	98	1379	-25
06-08	588067	-63	605	16	1428	47
07-08	588011	-114	612	57	1478	56
08-08	587954	-160	619	100	1528	59
09-08	587896	-187	626	101	1579	85
10-08	587838	-227	633	152	1630	74
11-08	587779	-238	640	168	1682	69
12-08	587719	-272	647	178	1735	93
13-08	587659	-320	654	237	1788	82
14-08	587598	-336	661	232	1842	103
15-08	587536	-397	668	261	1896	136
16-08	587474	-390	675	258	1951	132
17-08	587411	-388	682	210	2007	178
18-08	587347	-463	689	277	2063	187
19-08	587283	-478	697	317	2120	161
20-08	587218	-484	705	223	2177	261
21-08	587152	-512	713	195	2235	317
22-08	587086	-555	721	230	2294	324
23-08	587019	-526	729	147	2353	378
24-08	586951	-483	737	77	2413	405
25-08	586882	-480	745	74	2474	405
26-08	586813	-437	753	46	2535	390
27-08	586743	-436	761	73	2597	362
28-08	586672	-420	769	42	2660	377
29-08	586600	-454	777	119	2723	335
30-08	586528	-437	785	84	2787	353
31-08	586455	-389	793	30	2852	359

Fitted values and error for stochastic model. (30-day average)

Date	<i>Sf</i>	<i>Se</i>	<i>If</i>	<i>Ie</i>	<i>Rf</i>	<i>Re</i>
01-08	588340	0	573	0	1187	0
02-08	588264	-13	602	20	1234	-7
03-08	588188	19	628	11	1284	-30
04-08	588112	7	653	23	1335	-30
05-08	588036	14	675	21	1389	-35
06-08	587960	44	695	-74	1444	31
07-08	587884	13	714	-45	1502	32
08-08	587809	-15	731	-12	1560	27
09-08	587733	-24	747	-20	1620	44
10-08	587657	-46	761	24	1682	22
11-08	587581	-40	774	34	1744	7
12-08	587505	-58	787	38	1808	20
13-08	587429	-90	798	93	1873	-3
14-08	587354	-92	808	85	1939	6
15-08	587278	-139	817	112	2005	27
16-08	587202	-118	826	107	2072	11
17-08	587126	-103	834	58	2140	45
18-08	587050	-166	841	125	2209	41
19-08	586975	-170	848	166	2278	3
20-08	586899	-165	854	74	2348	90
21-08	586823	-183	859	49	2418	134
22-08	586747	-216	864	87	2488	130
23-08	586672	-179	869	7	2560	171
24-08	586596	-128	873	-59	2631	187
25-08	586520	-118	877	-58	2703	176
26-08	586444	-68	881	-82	2775	150
27-08	586369	-62	884	-50	2847	112
28-08	586293	-41	887	-76	2920	117
29-08	586217	-71	890	6	2993	65
30-08	586142	-51	892	-23	3066	74
31-08	586066	0	894	-71	3140	71

Section 4.2 14-day average comparison

Fitted values and error for differential model. (14-day average)

Date	Sf	Se	If	Ie	Rf	Re
01-08	588340	0	573	0	1187	0
02-08	588274	-23	591	31	1235	-8
03-08	588206	1	610	29	1284	-30
04-08	588136	-17	629	47	1335	-30
05-08	588064	-14	649	47	1387	-33
06-08	587990	14	669	-48	1441	34
07-08	587913	-16	690	-21	1497	37
08-08	587834	-40	712	7	1554	33
09-08	587753	-44	734	-7	1613	51
10-08	587669	-58	757	28	1674	30
11-08	587582	-41	781	27	1737	14
12-08	587493	-46	805	20	1802	26
13-08	587401	-62	830	61	1869	1
14-08	587306	-44	856	37	1938	7
15-08	587208	-69	883	46	2009	23
16-08	587066	18	927	6	2107	-24
17-08	586993	30	925	-33	2182	3
18-08	586920	-36	923	43	2257	-7
19-08	586847	-42	921	93	2332	-51
20-08	586775	-41	919	9	2406	32
21-08	586703	-63	917	-9	2480	72
22-08	586631	-100	915	36	2554	64
23-08	586559	-66	913	-37	2628	103
24-08	586487	-19	911	-97	2702	116
25-08	586415	-13	909	-90	2776	103
26-08	586344	32	907	-108	2849	76
27-08	586273	34	905	-71	2922	37
28-08	586202	50	903	-92	2995	42
29-08	586131	15	901	-5	3068	-10
30-08	586104	-13	860	9	3136	4
31-08	586064	2	826	-3	3210	1

Fitted values and error for stochastic model. (14-day average)

Date	<i>Sf</i>	<i>Se</i>	<i>If</i>	<i>Ie</i>	<i>Rf</i>	<i>Re</i>
01-08	588340	0	573	0	1187	0
02-08	588254	-3	611	11	1235	-8
03-08	588168	39	646	-7	1285	-31
04-08	588082	37	678	-2	1339	-34
05-08	587997	53	708	-12	1395	-41
06-08	587911	93	735	-114	1454	21
07-08	587825	72	760	-91	1515	19
08-08	587739	55	782	-63	1579	8
09-08	587653	56	803	-76	1644	20
10-08	587568	43	822	-37	1710	-6
11-08	587482	59	840	-32	1779	-28
12-08	587396	51	856	-31	1848	-20
13-08	587310	29	870	21	1919	-49
14-08	587255	37	884	9	1992	-47
15-08	587139	0	896	33	2065	-33
16-08	587068	16	925	8	2107	-24
17-08	586997	26	921	-29	2182	3
18-08	586926	-42	918	48	2256	-6
19-08	586855	-50	915	99	2330	-49
20-08	586784	-50	912	16	2404	34
21-08	586713	-73	909	-1	2478	74
22-08	586642	-111	907	44	2551	67
23-08	586571	-78	904	-28	2624	107
24-08	586500	-32	902	-88	2697	121
25-08	586430	-28	900	-81	2770	109
26-08	586359	17	898	-99	2843	82
27-08	586288	19	897	-63	2916	43
28-08	586217	35	895	-84	2988	49
29-08	586146	0	894	2	3060	-2
30-08	586106	-15	858	11	3136	4
31-08	586066	0	824	-1	3210	1

Section 4.3 7-day average comparison

Fitted values and error for differential model. (7-day average)

Date	Sf	Se	If	Ie	Rf	Re
01-08	588340	0	573	0	1187	0
02-08	588273	-22	590	32	1237	-10
03-08	588204	3	607	32	1289	-35
04-08	588133	-14	625	51	1342	-37
05-08	588060	-10	643	53	1397	-43
06-08	587985	19	662	-41	1453	22
07-08	587908	-11	681	-12	1511	23
08-08	587828	-34	701	18	1571	16
09-08	587714	-5	743	-16	1643	21
10-08	587631	-20	768	17	1701	3
11-08	587545	-4	794	14	1761	-10
12-08	587456	-9	820	5	1823	5
13-08	587364	-25	847	44	1887	-17
14-08	587269	-7	875	18	1954	-9
15-08	587171	-32	904	25	2023	9
16-08	587054	30	932	1	2114	-31
17-08	586968	55	935	-43	2196	-11
18-08	586882	2	938	28	2279	-29
19-08	586796	9	941	73	2362	-81
20-08	586709	25	944	-16	2445	-7
21-08	586622	18	947	-39	2529	23
22-08	586535	-4	950	1	2613	5
23-08	586469	24	944	-68	2687	44
24-08	586407	61	937	-123	2756	62
25-08	586346	56	930	-111	2824	55
26-08	586285	91	923	-124	2892	33
27-08	586225	82	916	-82	2959	0
28-08	586165	87	909	-98	3026	11
29-08	586106	40	902	-6	3092	-34
30-08	586104	-13	860	9	3136	4
31-08	586064	2	826	-3	3210	1

Fitted values and error for stochastic model. (7-day average)

Date	<i>Sf</i>	<i>Se</i>	<i>If</i>	<i>Ie</i>	<i>Rf</i>	<i>Re</i>
01-08	588340	0	573	0	1187	0
02-08	588262	-11	601	21	1237	-10
03-08	588184	23	626	13	1290	-36
04-08	588106	13	649	27	1345	-40
05-08	588028	22	670	26	1402	-48
06-08	587950	54	690	-69	1460	15
07-08	587872	25	707	-38	1521	13
08-08	587794	0	723	-4	1583	4
09-08	587700	9	756	-29	1643	21
10-08	587607	4	790	-5	1703	1
11-08	587513	28	822	-14	1765	-14
12-08	587420	27	851	-26	1829	-1
13-08	587326	13	878	13	1896	-26
14-08	587233	29	902	-9	1965	-20
15-08	587139	0	925	4	2036	-4
16-08	587052	32	934	-1	2114	-31
17-08	586965	58	938	-46	2197	-12
18-08	586878	6	942	24	2280	-30
19-08	586791	14	945	69	2363	-82
20-08	586705	29	948	-20	2447	-9
21-08	586618	22	951	-43	2531	21
22-08	586531	0	954	-3	2615	3
23-08	586476	17	937	-61	2687	44
24-08	586421	47	923	-109	2756	62
25-08	586366	36	911	-92	2823	56
26-08	586311	65	899	-100	2890	35
27-08	586256	51	889	-55	2955	4
28-08	586201	51	879	-68	3020	17
29-08	586146	0	869	27	3085	-27
30-08	586106	-15	858	11	3136	4
31-08	586066	0	824	-1	3210	1

Section 0

Future predictions

Predicted values and difference. (30 iterations with one matrix)

Date	Sp	Sd	Ip	Id	Rp	Rd
01-09	586009	2	826	20	3266	-23
02-09	585951	0	828	-23	3321	23
03-09	585894	-9	830	-6	3376	15
04-09	585869	-21	832	-21	3431	10
05-09	585779	1	834	-16	3487	15
06-09	585722	32	836	-69	3542	37
07-09	585664	76	838	-113	3598	37
08-09	585607	74	839	-106	3654	32
09-09	585550	103	841	-119	3710	15
10-09	585492	131	842	-164	3766	33
11-09	585435	136	843	-154	3822	18
12-09	585378	143	845	-147	3878	3
13-09	585320	198	846	-207	3934	9
14-09	585263	226	847	-276	3991	49
15-09	585206	269	848	-314	4047	44
16-09	585148	307	848	-387	4103	81
17-09	585091	338	849	-433	4160	95
18-09	585034	375	850	-535	4217	159
19-09	584976	415	851	-622	4273	207
20-09	584919	458	851	-713	4330	255
21-09	584862	498	852	-734	4386	236
22-09	584804	537	852	-734	4443	198
23-09	584747	574	853	-735	4500	161
24-09	584690	621	853	-744	4557	123
25-09	584633	650	854	-735	4614	84
26-09	584575	694	854	-745	4671	51
27-09	584518	747	855	-783	4727	36
28-09	584461	803	855	-788	4784	-15
29-09	584404	833	855	-772	4841	-61
30-09	584346	877	855	-777	4898	-99

Predicted values and difference. (30 iteration, matrix update every 5th iteration)

Date	Sp	Sd	Ip	Id	Rp	Rd
01-09	586009	2	826	20	3266	-23
02-09	585951	0	828	-23	3321	23
03-09	585894	-9	830	-6	3376	15
04-09	585836	12	832	-21	3431	10
05-09	585779	1	834	-16	3487	15
06-09	585728	26	809	-42	3563	16
07-09	585675	65	800	-75	3624	11
08-09	585623	58	792	-59	3684	2
09-09	585571	82	785	-63	3744	-19
10-09	585519	104	778	-100	3803	-4
11-09	585586	-15	664	25	3850	-10
12-09	585548	-27	651	47	3900	-19
13-09	585511	7	640	-1	3950	-7
14-09	585473	16	629	-58	3998	42
15-09	585436	39	619	-85	4046	45
16-09	585446	9	517	-56	4137	47
17-09	585416	13	502	-86	4182	73
18-09	585387	22	488	-173	4225	151
19-09	585357	34	475	-246	4268	212
20-09	585328	49	463	-325	4309	276
21-09	585357	3	126	-8	4618	4
22-09	585337	4	116	2	4647	-6
23-09	585317	4	109	9	4674	-13
24-09	585296	15	103	6	4700	-20
25-09	585276	7	99	20	4724	-26
26-09	585265	4	108	1	4727	-5
27-09	585247	18	99	-27	4754	9
28-09	585229	35	93	-26	4778	-9
29-09	585211	26	88	-5	4801	-21
30-09	585193	30	85	-7	4822	-23

B. Python script

Section 3.1 Compartmental model based on a system of differential equations

Python script for calculating the fitted values and error terms of the differential model.

```
#Inputing the data for August
Date = ['01-08-20','02-08-20','03-08-20','04-08-20','05-08-20','06-08-20','07-08-20','08-08-20','09-08-20','10-08-20','11-08-20','12-08-20','13-08-20','14-08-20','15-08-20','16-08-20','17-08-20','18-08-20','19-08-20','20-08-20','21-08-20','22-08-20','23-08-20','24-08-20','25-08-20','26-08-20','27-08-20','28-08-20','29-08-20','30-08-20','31-08-20']
S = [588340,588251,588207,588119,588050,588004,587897,587794,587709,587611,587541,587447,587339,587262,587139,587084,587023,586884,586805,586734,586640,586531,586493,586468,586402,586376,586307,586252,586146,586091,586066]
I = [573,622,639,676,696,621,669,719,727,785,808,825,891,893,929,933,892,966,1014,928,908,951,876,814,819,799,834,811,896,869,823]
R = [1187,1227,1254,1305,1354,1475,1534,1587,1664,1704,1751,1828,1870,1945,2032,2083,2185,2250,2281,2438,2552,2618,2731,2818,2879,2925,2959,3037,3058,3140,3211]
# Importing the necessary packages
import matplotlib.pyplot as plt # To plot the results
# The function for performing the calculations for the fitted values
def Differential(t,a,b,S0,I0,R0):
    for i in range (1,t+1):
        dSdt = -a*S0*I0 # S'(0)
        dIdt = a*S0*I0 -b*I0 # I'(0)
        dRdt = b*I0 # R'(0)
        S0 = round(S0 + dSdt)
        I0 = round(I0 + dIdt)
        R0 = round(R0 + dRdt)
        Sf.append(S0)
        If.append(I0)
        Rf.append(R0)
# calculating parameters
A = []
B = []
for i in range (1,31):
    ai = (S[i-1]-S[i])/(S[i-1]*I[i-0])
    A.append(ai)
    bi = (R[i]-R[i-1])/I[i-1]
    B.append(bi)
# 30-day average
# Calculating The parameters
a = sum(A)/len(A)
b = sum(B)/len(B)
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
Sf.append(S0)
If.append(I0)
Rf.append(R0)
```

```

Differential(30,a,b,S0,I0,R0)
print("30 days")
print("a = "+str(a)+"\nb = "+str(b))
# Calculating the error terms
Se = [] # a list that will contain the error of S
Ie = [] # a list that will contain the error of I
Re = [] # a list that will contain the error of R
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.title (" S observed and S fitted.30-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)
plt.plot(Date,Rf)
plt.legend(["I observed","I fitted","R observed","R fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" I and R observed; I and R fitted.30-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)
plt.plot(Date,Re)
plt.legend(["S error","I error","R error"])
plt.xlabel("Date")
plt.ylabel("error")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title ("errors of S, I and R.30-day average")
plt.show()
# 14-day average
# calculating the parameters
# First 14 days
a_1 = sum(A[0:14])/(14)
b_1 = sum(B[0:14])/(14)
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R

```

```

# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
Sf.append(S0)
If.append(I0)
Rf.append(R0)
Differential(14,a_1,b_1,S0,I0,R0)
# Second 14 days
a_2 = sum(A[14:28])/(14)
b_2 = sum(B[14:28])/(14)
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[14]
I0 = I[14]
R0 = R[14]
Differential(14,a_2,b_2,S0,I0,R0)
# Last 2 days
a_3 = sum(A[28:30])/(2)
b_3 = sum(B[28:30])/(2)
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[28]
I0 = I[28]
R0 = R[28]
Differential(2,a_3,b_3,S0,I0,R0)
print(" 14 days")
print("a1 = "+str(a_1)+"\nb1 = "+str(b_1))
print("a2 = "+str(a_2)+"\nb2 = "+str(b_2))
print("a3 = "+str(a_3)+"\nb3 = "+str(b_3))
# Calculating the error terms
Se = [] # a list that will contain the error of S
Ie = [] # a list that will contain the error of I
Re = [] # a list that will contain the error of R
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.title (" S observed and S fitted. 14-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)
plt.plot(Date,Rf)
plt.legend(["I observed","I fitted","R observed","R fitted"])
plt.xlabel("Date")

```

```

plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" I and R observed; I and R fitted.14-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)
plt.plot(Date,Re)
plt.legend(["S error", "I error", "R error"])
plt.xlabel("Date")
plt.ylabel("error")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title ("errors of S, I and R.14-day average")
plt.show()
# 7-day average
# calculating the transition probabilities
# First 7 days
a_1 = sum(A[0:7])/(7)
b_1 = sum(B[0:7])/(7)
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
Sf.append(S0)
If.append(I0)
Rf.append(R0)
Differential(7,a_1,b_1,S0,I0,R0)
# Second 7 days
a_2 = sum(A[7:14])/(7)
b_2 = sum(B[7:14])/(7)

# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[7]
I0 = I[7]
R0 = R[7]
Differential(7,a_2,b_2,S0,I0,R0)
# Third 7 days
a_3 = sum(A[14:21])/(7)
b_3 = sum(B[14:21])/(7)
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[14]
I0 = I[14]
R0 = R[14]
Differential(7,a_3,b_3,S0,I0,R0)
# Fourth 7 days
a_4 = sum(A[21:28])/(7)
b_4 = sum(B[21:28])/(7)
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[21]
I0 = I[21]

```

```

R0 = R[21]
Differential(7,a_4,b_4,S0,I0,R0)
# Last 2 days
a_5 = sum(A[28:30])/(2)
b_5 = sum(B[28:30])/(2)
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[28]
I0 = I[28]
R0 = R[28]
Differential(2,a_5,b_5,S0,I0,R0)
print("7 days")
print("a1 = "+str(a_1)+"\nb1 = "+str(b_1))
print("a2 = "+str(a_2)+"\nb2 = "+str(b_2))
print("a3 = "+str(a_3)+"\nb3 = "+str(b_3))
print("a4 = "+str(a_4)+"\nb4 = "+str(b_4))
print("a5 = "+str(a_5)+"\nb5 = "+str(b_5))
# Calculating the error terms
Se = [] # a list that will contain the error of S
Ie = [] # a list that will contain the error of I
Re = [] # a list that will contain the error of R
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.title (" S observed and S fitted. 7-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)
plt.plot(Date,Rf)
plt.legend(["I observed","I fitted","R observed","R fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" I and R observed; I and R fitted.7-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)
plt.plot(Date,Re)
plt.legend(["S error","I error","R error"])
plt.xlabel("Date")
plt.ylabel("error")

```



```
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title ("errors of S, I and R.7-day average")
plt.show()
```

Section 3.2 Stochastic compartmental models

Python script for calculating the fitted values and error terms of the stochastic model.

```
#Inputing the data for August
Date = ['01-08-20','02-08-20','03-08-20','04-08-20','05-08-20','06-08-20','07-08-20','08-08-20','09-08-20','10-08-20','11-08-20','12-08-20','13-08-20','14-08-20','15-08-20','16-08-20','17-08-20','18-08-20','19-08-20','20-08-20','21-08-20','22-08-20','23-08-20','24-08-20','25-08-20','26-08-20','27-08-20','28-08-20','29-08-20','30-08-20','31-08-20']
S = [588340,588251,588207,588119,588050,588004,587897,587794,587709,587611,587541,587447,587339,587262,587139,587084,587023,586884,586805,586734,586640,586531,586493,586468,586402,586376,586307,586252,586146,586091,586066]
I = [573,622,639,676,696,621,669,719,727,785,808,825,891,893,929,933,892,966,1014,928,908,951,876,814,819,799,834,811,896,869,823]
R = [1187,1227,1254,1305,1354,1475,1534,1587,1664,1704,1751,1828,1870,1945,2032,2083,2185,2250,2281,2438,2552,2618,2731,2818,2879,2925,2959,3037,3058,3140,3211]
States = ["S", "I", "R"]
# Importing the necessary packages
import matplotlib.pyplot as plt # To plot the results
import numpy # To perform matrix calculations
# The function for performing the matrix calculations for the fitted values
def Matrix(P,t,S0,I0,R0):
    Pi = Id
    for i in range (1,t+1):
        Pi = numpy.array(Pi@P)
        St = round((S0*(Pi.item(0))) +(I0*(Pi.item(3))) +(R0*(Pi.item(6))))
        It = round((S0*(Pi.item(1))) +(I0*(Pi.item(4))) +(R0*(Pi.item(7))))
        Rt = round((S0*(Pi.item(2))) +(I0*(Pi.item(5))) +(R0*(Pi.item(8))))
        Sf.append(St)
        If.append(It)
        Rf.append(Rt)
# Identity matrix
Id = numpy.array([[1,0,0],[0,1,0],[0,0,1]])
# calculating transition proportions
U = []
V = []
for i in range (1,31):
    ui = (S[i-1]-S[i])/S[i-1]
    U.append(ui)
    vi = (R[i]-R[i-1])/I[i-1]
    V.append(vi)
# 30-day average
# Calculating Transition probability
u = sum(U)/len(U)
v = sum(V)/len(V)
u1 = 1 - u
v1 = 1-v
Spr = [round(u1,5),0,0]
Ipr = [round(u,5),round(v1,5),0]
Rpr = [0,round(v,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
```

```

print("Transition Probability matrix for 30-day average version")
for i,d in enumerate(data):
    line = '|'.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
# The Transition probability matrix
P = numpy.array([[u1, u, 0],[0, v1, v],[0, 0, 1]])
Sf.append(S0)
If.append(I0)
Rf.append(R0)
Matrix(P,30,S0,I0,R0)
# Calculating the error terms
Se = [] # a list that will contain the relative error of S in percentage
Ie = [] # a list that will contain the relative error of I in percentage
Re = [] # a list that will contain the relative error of R in percentage
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation=-45)
plt.title (" S observed and S fitted.30-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)
plt.plot(Date,Rf)
plt.legend(["I observed","I fitted","R observed","R fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation=-45)
plt.grid()
plt.title (" I and R observed; I and R fitted.30-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)

```

```

plt.plot(Date,Re)
plt.legend(["S error","I error","R error"])
plt.xlabel("Date")
plt.ylabel("Percentage (%)")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" Relative errors of S, I and R.30-day average")
plt.show()
# 14-day average
# calculating the transition probabilities
# First 14 days
u_1 = sum(U[0:14])/(14)
v_1 = sum(V[0:14])/(14)
u1_1 = 1-u_1
v1_1 = 1-v_1
Spr = [round(u1_1,5),0,0]
Ipr = [round(u_1,5),round(v1_1,5),0]
Rpr = [0,round(v_1,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 14-day average version part 1")
for i,d in enumerate(data):
    line = '|'.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
# The Transition probability matrix
P = numpy.array([[u1_1, u_1, 0],[ 0, v1_1, v_1],[ 0, 0, 1]])
Sf.append(S0)
If.append(I0)
Rf.append(R0)
Matrix(P,14,S0,I0,R0)
# Second 14 days
u_2 = sum(U[14:28])/(14)
v_2 = sum(V[14:28])/(14)
u1_2 = 1-u_2
v1_2 = 1-v_2
Spr = [round(u1_2,5),0,0]
Ipr = [round(u_2,5),round(v1_2,5),0]
Rpr = [0,round(v_2,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 14-day average version part 2")
for i,d in enumerate(data):
    line = '|'.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))

```

```

# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[14]
I0 = I[14]
R0 = R[14]
# The Transition probability matrix
P = numpy.array([[u1_2, u_2, 0],[ 0, v1_2, v_2],[ 0, 0, 1]])
Matrix(P,14,S0,I0,R0)
# Last 2 days
u_3 = sum(U[28:30])/(2)
v_3 = sum(V[28:30])/(2)
u1_3 = 1-u_3
v1_3 = 1-v_3
Spr = [round(u1_3,5),0,0]
Ipr = [round(u_3,5),round(v1_3,5),0]
Rpr = [0,round(v_3,5),1]
Names = ['S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 14-day average version part 3")
for i,d in enumerate(data):
    line = ''.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[28]
I0 = I[28]
R0 = R[28]
# The Transition probability matrix
P = numpy.array([[u1_3, u_3, 0],[ 0, v1_3, v_3],[ 0, 0, 1]])
Matrix(P,2,S0,I0,R0)
# Calculating the error terms
Se = [] # a list that will contain the relative error of S in percentage
Ie = [] # a list that will contain the relative error of I in percentage
Re = [] # a list that will contain the relative error of R in percentage
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.title (" S observed and S fitted. 14-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)
plt.plot(Date,Rf)

```

```

plt.legend(["I observed", "I fitted", "R observed", "R fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date, rotation='-45')
plt.grid()
plt.title (" I and R observed; I and R fitted.14-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)
plt.plot(Date,Re)
plt.legend(["S error", "I error", "R error"])
plt.xlabel("Date")
plt.ylabel("Percentage (%)")
plt.xticks(Date, rotation='-45')
plt.grid()
plt.title (" Relative errors of S, I and R.14-day average")
plt.show()
# 7-day average
# calculating the transition probabilities
# First 7 days
u_1 = sum(U[0:7])/(7)
v_1 = sum(V[0:7])/(7)
u1_1 = 1-u_1
v1_1 = 1-v_
Spr = [round(u1_1,5),0,0]
Ipr = [round(u_1,5),round(v1_1,5),0]
Rpr = [0,round(v_1,5),1]
Names = [' ', 'S', 'I', 'R']
data = [Names] + list(zip(States, Spr, Ipr, Rpr))
print("Transition Probability matrix for 7-day average version part 1")
for i,d in enumerate(data):
    line = '|'.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# Calculating the fitted values
Sf = [] # a list that will contain the fitted value of S
If = [] # a list that will contain the fitted value of I
Rf = [] # a list that will contain the fitted value of R
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[0]
I0 = I[0]
R0 = R[0]
# The Transition probability matrix
P = numpy.array([[u1_1, u_1, 0],[ 0, v1_1, v_1],[ 0, 0, 1]])
Sf.append(S0)
If.append(I0)
Rf.append(R0)
Matrix(P,7,S0,I0,R0)
# Second 7 days
u_2 = sum(U[7:14])/(7)
v_2 = sum(V[7:14])/(7)
u1_2 = 1-u_2
v1_2 = 1-v_2

```

```

Spr = [round(u1_2,5),0,0]
Ipr = [round(u_2,5),round(v1_2,5),0]
Rpr = [0,round(v_2,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 7-day average version part 2")
for i,d in enumerate(data):
    line = ''.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[7]
I0 = I[7]
R0 = R[7]
# The Transition probability matrix
P = numpy.array([ [u1_2, u_2, 0],[ 0, v1_2, v_2],[ 0, 0, 1] ])
Matrix(P,7,S0,I0,R0)
# Third 7 days
u_3 = sum(U[14:21])/(7)
v_3 = sum(V[14:21])/(7)
u1_3 = 1-u_3
v1_3 = 1-v_3
Spr = [round(u1_3,5),0,0]
Ipr = [round(u_3,5),round(v1_3,5),0]
Rpr = [0,round(v_3,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 7-day average version part 3")
for i,d in enumerate(data):
    line = ''.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('-' * len(line))
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[14]
I0 = I[14]
R0 = R[14]
# The Transition probability matrix
P = numpy.array([ [u1_3, u_3, 0],[ 0, v1_3, v_3],[ 0, 0, 1] ])
Matrix(P,7,S0,I0,R0)
# Fourth 7 days
u_4 = sum(U[21:28])/(7)
v_4 = sum(V[21:28])/(7)
u1_4 = 1-u_4
v1_4 = 1-v_4
Spr = [round(u1_4,5),0,0]
Ipr = [round(u_4,5),round(v1_4,5),0]
Rpr = [0,round(v_4,5),1]
Names = [' ','S','I','R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 7-day average version part 4")
for i,d in enumerate(data):
    line = ''.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:

```

```

    print('- ' * len(line))
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[21]
I0 = I[21]
R0 = R[21]
# The Transition probability matrix
P = numpy.array([[u1_4, u_4, 0],[ 0, v1_4, v_4],[ 0, 0, 1]])
Matrix(P,7,S0,I0,R0)
# Last 2 days
u_5 = sum(U[28:30])/(2)
v_5 = sum(V[28:30])/(2)
u1_5 = 1-u_5
v1_5 = 1-v_5
Spr = [round(u1_5,5),0,0]
Ipr = [round(u_5,5),round(v1_5,5),0]
Rpr = [0,round(v_5,5),1]
Names = [' ', 'S', 'I', 'R']
data = [Names] + list(zip(States,Spr,Ipr,Rpr))
print("Transition Probability matrix for 14-day average version part 5")
for i,d in enumerate(data):
    line = '|'.join(str(x).ljust(11) for x in d)
    print(line)
    if i == 0:
        print('- ' * len(line))
# The initial values for S, I and R at t = 0 (August 1st)
S0 = S[28]
I0 = I[28]
R0 = R[28]
# The Transition probability matrix
P = numpy.array([[u1_5, u_5, 0],[ 0, v1_5, v_5], [ 0, 0, 1]])
Matrix(P,2,S0,I0,R0)
# Calculating the error terms
Se = [] # a list that will contain the relative error of S in percentage
Ie = [] # a list that will contain the relative error of I in percentage
Re = [] # a list that will contain the relative error of R in percentage
for i in range (len(Date)):
    Sei = (S[i]-Sf[i])
    Se.append(Sei)
    Iei = (I[i]-If[i])
    Ie.append(Iei)
    Rei = (R[i]-Rf[i])
    Re.append(Rei)
# Plotting the fitted values and the error terms
plt.figure(figsize=(10,7))
plt.plot(Date,S)
plt.plot(Date,Sf)
plt.legend(["S observed","S fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.title (" S observed and S fitted. 7-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,I)
plt.plot(Date,If)
plt.plot(Date,R)

```

```
plt.plot(Date,Rf)
plt.legend(["I observed","I fitted","R observed","R fitted"])
plt.xlabel("Date")
plt.ylabel("People")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" I and R observed; I and R fitted.7-day average")
plt.show()
plt.figure(figsize=(10,7))
plt.plot(Date,Se)
plt.plot(Date,Ie)
plt.plot(Date,Re)
plt.legend(["S error","I error","R error"])
plt.xlabel("Date")
plt.ylabel("Percentage (%)")
plt.xticks(Date,rotation='-45')
plt.grid()
plt.title (" Relative errors of S, I and R.7-day average")
plt.show()
```


C. References

1. Aleman, D. M., Wibisono, T. G., & Schwartz, B. (2009). Accounting for individual behaviors in a pandemic disease spread model. *Winter Simulation Conference*, (pp. 1977-1985).
2. ALGEMEEN BUREAU voor de STATISTIEK. (2019). *Demographic Data 2015-2018*. Paramaribo. Retrieved from ABS Suriname.
3. Brauer, F., Castillo-Chavez, C., & Feng, Z. (2019). *Mathematical Models in Epidemiology*. New York: Springer Science+Business Media, LLC.
4. Directorate National Security . (2020, August, September). Retrieved from Covid-19: <https://covid-19.sr/>
5. *Hypothesis Test: Difference Between Means*. (2021, August). Opgehaald van Stat Trek: <https://stattrek.com/hypothesis-test/difference-in-means.aspx>
6. Kemeny, J. G., & Snell, J. L. (1960). *Finite Markov Chains* (2nd ed.). Princeton, New Jersey: Van Nostrand.
7. Kretzschmar, M. E., Rozhnova, G., & van Boven, M. (2020). *Effectiveness of isolation and contact tracing for containment and slowing down a COVID-19 epidemic: a modelling study*.
8. *Suriname: Coronavirus Pandemic Country Profile*. (2021, July 1). Retrieved from Our World in Data: <https://ourworldindata.org/coronavirus/country/suriname?country=~SUR>
9. Twumasi, C., Asiedu, L., & Nortey, E. N. (2019). Markov Chain Modeling of HIV, Tuberculosis, and Hepatitis B Transmission in Ghana . *Interdisciplinary Perspectives on Infectious Diseases*.
10. Winston, W. L. (2004). Markov Chains. In W. L. Winston, *Operations Research: Applications and Algorithms* (pp. 923-960). Toronto: Brooks/Cole.
11. World Health Organization. (2020). *Novel Coronavirus (2019-nCoV) situation report - 1*.
12. World Health Organization. (2020). *Novel Coronavirus (2019-nCoV) situation report - 22*.
13. World Health Organization. (2020). *Novel Coronavirus (2019-nCoV) situation report - 8*.
14. World Health Organization. (2021). *Weekly epidemiological update on COVID-19 - 13 July 2021*.