

UNIVERSITY OF WITWATERSRAND

SCHOOL OF MATHEMATICAL STATISTICS AND
ACTUARIAL SCIENCES

ORIGINAL PUBLISHED WORK SUBMITTED TO THE FACULTY
OF SCIENCE IN FULFILMENT OF THE DEGREE OF DOCTORATE
IN PHILOSOPHY

Stochastic Modelling of the Spread of Infectious Diseases

Author:

Raeesa DOCRAT

Supervisor:

Prof. Frank BEICHELT

June 27, 2022



WITS
UNIVERSITY

Contents

Abstract	v
Declaration	xi
List of Symbols	xiii
List of Abbreviations	xv
List of Figures	xix
List of Tables	xx
1 Introduction	1
1.1 Literature Review	2
1.2 Problem Statement	5
1.3 Aim and Objectives of the Study	5
1.3.1 Aim of the Study	5
1.3.2 Objectives	5
1.4 Infectious Disease Terminology	6
1.4.1 Exposure	7
1.4.2 Infection	8
1.5 Outline of Thesis	8
1.6 Presentations and Publications	9
2 Basic Models	11

2.1	Compartmental Models	11
2.1.1	Characterisations and Assumptions	12
2.1.2	The Basic Reproduction Number	13
2.2	The Deterministic SIR Model	13
2.2.1	Properties of the SIR Model	14
2.2.2	The Deterministic Stratified SIR Model	20
2.3	The Stochastic SIR Model	36
2.4	Conclusion	46
3	Justification for a spatial model for the spread of infectious diseases: Tuberculosis in South Africa as an example	47
3.1	Study Area and Methods	49
3.1.1	South Africa	49
3.1.2	Measuring Spatial Autocorrelation	49
3.1.3	Software	52
3.2	Results	53
3.2.1	Plot of the Proportion of TB incidence	53
3.2.2	Assessing the Spatial Autocorrelation	54
3.2.3	Hot and Cold Spot Analysis	56
3.3	Analysis	56
3.3.1	Spatial autocorrelation and the effects of the spatial weights matrix	56
3.3.2	Proportion of TB cases	57
3.4	Conclusion	58
4	Discrete-Time Stochastic Compartmental Models	61
4.1	Binomial Chain Models	63
4.1.1	Reed-Frost Model	63
4.1.2	The Greenwood Model	66

4.1.3	Tuckwell-Williams Model	66
4.1.4	Spatio-Stochastic Models	67
4.2	Simulation, Application and Analysis of Spatio-stochastic Model	76
4.2.1	Simulation Setup	76
4.2.2	Simulation Results	78
4.2.3	Application to Measles Data	82
4.3	Discussion	83
4.4	Conclusion	85
5	Conclusion	89
5.1	Review of thesis	89
5.2	Significance and contributions of this study	91
5.3	Limitations of the study	92
5.4	Further Work	93
	References	94
	Appendices	100
A	Mathematical Tools	101
A.1	Generating Functions	101
A.1.1	Ordinary Generating Functions	101
A.1.2	Bi-variate Generating Functions	102
A.1.3	Statistical Generating Functions	102
A.2	Differential Equations	104
A.2.1	First Order Differential Equations	104
A.2.2	Second Order Differential Equations	106
B	Stochastic Processes	107
B.1	Birth and Death Processes	107
B.1.1	Discrete-Time Birth and Death Processes	107

B.1.2	Continuous-Time Birth and Death Processes	111
C	Codes for the Spatio-Stochastic Models	119

Abstract

Background

The spread of infectious diseases is a world-wide problem that has a greater impact on low-income countries. Mathematical modelling is a useful tool to better understand these diseases and to plan prevention and interventions. These models can either be deterministic or stochastic. Generally, deterministic models are used as they are easier to understand and predict average behaviour. Stochastic models are probabilistic models which incorporate random fluctuations over time. Stochastic models allow the calculation of probabilities and likelihoods of possible progressions of an infectious disease. These models can be built to include many aspects and features of infectious diseases. One such inclusion is the spatial nature inherent in the spread of infectious diseases. The inclusion of a spatial component for the spread of infectious diseases is needed, especially in South Africa, where it is spatially heterogeneous, to accurately analyse and efficiently allocate resources to where it is most needed.

Methods

This thesis begins with a review of both deterministic and stochastic compartmental models to illustrate the uses of these models in the context of disease modelling. The spatial autocorrelation of the spread of TB in South Africa is then studied. The results, although intuitive, confirm the existence of spatial autocorrelation and hence the need for models which accounts for the spatial

heterogeneity. A spatio-stochastic model in discrete time and space is then developed using binomial chain models as a basis. The spatial and temporal differences are modelled in the number of contacts a person would encounter. The spatio-stochastic models are then further extended to account for different modelling scenarios: closed spatial populations, spatial populations with interactions, but no migrations and finally a spatial model with migrations and interactions from other spatial units. This developed model is then compared to existing models in a simulation study. Lastly, the developed model is applied to a measles data set.

Results

The simulation study illustrates the differences in the outcome of different cases of the spatio-stochastic model. The disease develops and progresses very differently, depending on the modelling case chosen. This highlights the importance of choosing the correct model. The simulation study demonstrated the differences between other binomial chain based models, namely the Reed-Frost and the Tuckwell Williams model. The inclusion of spatial heterogeneity produced outcomes that varied from the two non-spatial models. Lastly, the application of the model to a measles data set confirms that the inclusion of the spatial component allows a better fit of the models to the data.

Conclusion

The spatio-stochastic models are flexible and can be tailored to the dynamics of a particular disease's spatial spread. This aspect makes these models very useful in understanding and analysing the spatial spread of the infectious disease.

The developed spatio-stochastic model can be applied to a wide variety of cases. The fact that it is stochastic allows not only the ability to predict av-

erage behaviour, but also to calculate probabilities related to vital questions in planning and prevention. Questions related to the probability of outbreak in a particular area or which areas are likely to be hot spot areas can now be addressed. The modelling of the spread of infectious diseases using a spatial model, like the model developed in this thesis, is important to make informed decisions on how to stop and prevent further infection. The spatial models can provide information on sub-populations which can be used to allocate resources accordingly. It is also useful in identifying the effects of different interventions like vaccines. Due to globalisation and the level of connection in society, the spatio-stochastic modelling of infectious diseases is vital.

Acknowledgements

Alhumdullilah!

This thesis has been the central part of my life for the past 4 years and I have many people to thank for their support, guidance and insight over this period.

The first acknowledgement goes to my supervisor, Professor Frank Beichelt. Thank you for all your support and guidance over the past 5 years. I have been extremely lucky and grateful to have someone of your calibre and expertise as a supervisor. You have really played a big role in my academic career and without you, I would have not been in the position I am today.

I would like to thank my husband, Qiyaam for all of his support and patience. As a wife I am a handful, as a wife and a PhD-candidate I am a tornado of emotions!

I would like to acknowledge my family and friends, though many of them did not really understand the equations and models, they showered me with support, shoulders to cry on, editing drafts, on and kept me healthy and positive. I would like to thank in particular Rabia, who has helped a lot with editing and admin!

To my friends and colleagues at the school, especially Dr Justine Nasejje and Dr Raeesa Ganey, I would like to acknowledge your advice, support and problem solving any logistical problem.

I would like to acknowledge my ex-neighbour and mentor Professor Loyiso Nongxa for always being there for any problem, the constant support, the advice and motivation.

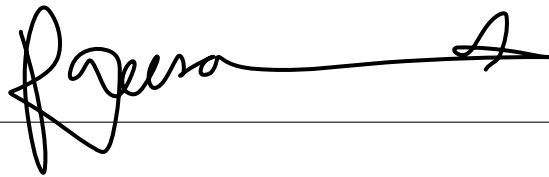
I would like to thank and acknowledge all those who helped when I reached out. I am grateful for all the help and guidance and hope that I too can contribute to others in the same way.

Finally, I would like to acknowledge the UCDP, for funding the courses and conferences that I have attended which enriched my knowledge and experience.

“If I have seen further it is by standing on the shoulders of Giants” - Isaac Newton

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Doctor in Philosophy at the University of Witwatersrand, Johannesburg. It has not been submitted before for any other degree or examination at any other University.

A handwritten signature in black ink, appearing to read 'Raeesa', followed by a long horizontal line that ends in a small loop.

Raeesa Manjoo-Docrat

27th day of June 2022 in Johannesburg

List of Symbols

$\mathbb{N}$	Set of all positive integers
$\mathbb{Z}$	Set of all integers
$\mathbb{Z}_+$	Set of all non-negative integers
$\mathbb{Q}$	Set of all rational numbers
$\mathbb{Q}'$	Set of all irrational numbers
$\mathbb{R}$	$(-\infty, \infty)$
$\mathbb{R}_+$	$[0, \infty)$
$*$	Convolution $g * h(t) = \int_0^\infty g(t-u)h(du)$
$\square$	Symbol after proof of a theorem
$f \circ g$	Composition of two functions f and g
$f^*(n)$	n^{th} Convolution of a function f
$f^{(n)}$	n^{th} derivative of function f
$o(x)$	Landau order symbol
X, Y, Z	Random variables
$\mu_x, E[X]$	Mean or expected value of random variable X
$\sigma_X^2, var(X)$	Variance of random variable X
$P_X(t)$	Probability generating function of a random variable X
$M_X(t)$	Moment generating function of a random variable X
$\phi_X(t)$	Characteristic function of a random variable X
$K_X(t)$	Cumulant generating function of a random variable X
$\{X(t), t \in \mathbf{T}\}$	Stochastic process with parameter space $\mathbf{T}$
$m(t)$	Trend function of a stochastic process, $t \in \mathbf{T}$

$p_{ij}, p_{ij}^{(n)}$	One, n -step transition probabilities of a homogeneous discrete time Markov chain
$p_{ij}(t)$	Transition probabilities of a continuous-time homogeneous Markov chain
π_i	Stationary state distribution of a homogenous Markov chain $\{\pi_i, i \in \mathbb{Z}\}$
λ_i	Birth rate
μ_i	Death rate
L	Lifetime

List of Abbreviations

CDC	Centers for Disease Control and Prevention
CGF	Cumulant Generating Function
CHF	Characteristic Generating Function
E	Exposed
EVD	Ebola Virus Disease
HIV	Human Immunodeficiency Virus
I	Infective
MGF	Moment Generating Function
NICD	National Institute for Communicable Diseases
OGF	Ordinary Generating Function
PGF	Probability Generating Function
R	Recovered
S	Susceptible
SARS	Severe Acute Respiratory Syndrome
SEIR	Model with states Susceptible-Exposed-Infective-Recovered
SIR	Model with states Susceptible-Infective-Recovered
SIRS	Model with states Susceptible-Infective-Recovered-Susceptible
TB	Tuberculosis
WHO	World Health Organisation

List of Figures

1.1	<i>Elements in the Chain of Infection.</i>	7
1.2	<i>Outcomes of Exposure.</i>	7
1.3	<i>Stages of Infection and Disease.</i>	8
2.1	<i>Diagrammatic representation of the SIR model.</i>	11
2.2	<i>Classification of SIR Models.</i>	12
2.3	<i>Simulation of a Deterministic SIR model with $\beta = 0.7$, $\gamma = 0.2$ and $N = 1000$.</i>	14
2.4	<i>Plot of $R(\infty)$ vs R_0.</i>	16
2.5	<i>Illustration of flattening the curve. Source: CDC</i>	17
2.6	<i>Diagrammatic representation of the SIRD model.</i>	18
2.7	<i>Illustration of the effect of social distancing.</i>	19
2.8	<i>Sensitivity analysis with regard to ψ.</i>	20
2.9	<i>South African population per age.</i>	22
2.10	<i>Steps in the modelling process.</i>	23
2.11	<i>Model structure of SEIR model.</i>	24
2.12	<i>Matrix plot of the 4 age groups.</i>	26
2.13	<i>Timeline of Lockdown Levels.</i>	28
2.14	<i>Plot of the actual versus observed infections of the calibrated model.</i>	31
2.15	<i>Timeline of South Africa's Vaccination Roll-out.</i>	32

2.16	<i>Observed and predicted total infections: a) observed data, b) base model, c) 10% less virulent, d) 20% less virulent, e) 10% more virulent, f) 20% more virulent.</i>	35
2.17	<i>Observed and predicted total infections: a) observed data, b) 90% efficacy, c) 80% efficacy, d) 70% efficacy, e) 60% efficacy, f) 50% efficacy.</i>	35
2.18	<i>Observed and predicted total infections with 30% reduction in virulence and 63% efficacy of the vaccination.</i>	36
2.19	<i>Illustration of the transition between states.</i>	42
3.1	<i>Plot of South Africa Districts and Provinces.</i>	49
3.2	<i>Plot of TB incidence in South Africa.</i>	53
3.3	<i>Plot of the local Moran's I statistic using queen adjacency. . . .</i>	55
3.4	<i>Plot of the local Moran's I statistic using 8 nearest neighbours. .</i>	55
3.5	<i>Plot of the local Moran's I statistic using threshold distance of 248 km.</i>	56
3.6	<i>Plot of hot and cold spots using different weight inputs. The white indicates no significant relationship, the red is a hot spot, which is a clustering of high values, the dark blue is a cold spot, which is clustering of low values and the light blue is a clustering of both high and low values.</i>	57
4.1	<i>Simulations of a Reed-Frost model with $p = 0.001$, $I_0 = 10$ and $S_0 = 2000$.</i>	65
4.2	<i>Simulations of a Reed-Frost model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.</i>	65
4.3	<i>Final size of a Reed-Frost model with $p = 0.001$, $I_0 = 10$ and $S_0 = 2000$.</i>	65

4.4	<i>Final size of a Reed-Frost model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.</i>	65
4.5	<i>Simulations of a Greenwood model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.</i>	66
4.6	<i>Final size of a Greenwood model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.</i>	66
4.7	<i>Transition graph Spatio-Stochastic Model of an individual.</i>	69
4.8	<i>Transition graph Spatio-Stochastic SEIR Model of an individual.</i>	72
4.9	<i>A diagrammatic representation of the population: (a) Reed-Frost and Tuckwell-Williams populations, (b) Spatio-Stochastic Case I population, (c) Spatio-Stochastic Case II population, (d) Spatio-Stochastic Case III population.</i>	78
4.10	<i>Final size of 15 000 simulations of Reed-Frost model varying the probability of transmission and infection: (a) $p = 0.005$, (b) $p = 0.0075$, (c) $p = 0.01$.</i>	78
4.11	<i>Final size of Reed-Frost model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.</i>	79
4.12	<i>Final size of Tuckwell-Williams model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.</i>	79
4.13	<i>Final size of Spatio-stochastic I model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.</i>	80
4.14	<i>Final size of Spatio-stochastic model I over 15 000 simulations in : (a) subpopulation 1, (b) subpopulation. 2</i>	81

4.15	<i>Final size of Spatio-stochastic II model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.</i>	81
4.16	<i>Final size of Spatio-stochastic III model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.</i>	82
4.17	<i>95% Confidence intervals fit for Models 1,2,3 and 4 over 1000 simulations.</i>	83
4.18	<i>Comparison of the final size of the Spatio-stochastic Models.</i>	84
4.19	<i>Violin plot of the final size simulations for each of the models.</i>	84
B.1	<i>Transition graph of a birth and death process with finite state space on condition $r_i = 0$</i>	108

List of Tables

2.1	Case Fatality Rates per age group.	22
2.2	The age-groups used in the age-stratified compartmental model.	23
2.3	Description of the different compartments used in the SEIR model.	24
2.4	Model Parameters.	26
2.5	The age-groups used in the age-stratified compartmental model.	27
2.6	The lockdown levels and scaling.	28
2.7	The dates of the waves in South Africa.	29
2.8	The lockdown levels and scaling after calibration.	30
2.9	The RMSE of different scenarios.	31
2.10	Model Parameters of Vaccination Model	34
2.11	Model Parameters of Combined Model	42

3.1	Summary of spatial weights	50
3.2	Results of Moran's I test for the different weights.	54
4.1	Parameters of the Reed-Frost Simulations.	76
4.2	Parameters of the Tuckwell-Williams Simulations.	77
4.3	Parameters of the Spatio-stochastic Model Simulations.	77
4.4	Summary of Final Size per Model.	85
A.1	Combinatorial construction of operations used in counting . . .	102

Chapter 1

Introduction

An infectious disease is one that is caused by an infectious or pathogenic agent developing in or on an infected host. Viruses, bacteria, parasites and fungi are pathogenic agents which can cause infectious diseases. Infectious agents can be spread from an infected host either directly or indirectly. Direct methods of infection include physical contact, sneezing, and coughing. Indirect infections can occur through contaminated objects or by vectors such as mosquitoes or ticks.

An important aspect of health policies should be to control and to reduce the incidence and distribution of infectious diseases especially during an epidemic outbreak. In low-income countries, infectious diseases are more common (World Health Organization (2018)). Low-income countries generally lack sound health care, infrastructure and finances for the successful control of infectious diseases. The World Health Organization (2018) indicates that infectious diseases are amongst the top 10 contributors to deaths in low-income countries.

To minimise the occurrence of infections and reduce the number of future deaths, intervention and prevention health policies need to be implemented. Resources are limited and therefore they should be used efficiently. The spread and extent of infectious diseases are often modelled using mathematical models. They are quick and low-cost methods for comparing various strategies to maximise the effect of the different strategies, based on the limited data available (Maude et al. (2012)).

Mathematical modelling of infectious diseases started in the eighteenth century. Since then, many scientists have adapted and improved the models. A model describing the spread of an infectious disease can either be deterministic or stochastic. *Deterministic models* are typically based on a system of differential equations that describe the modelled aspects. Deterministic mod-

els predict the average development of infectious diseases. *Stochastic models or probabilistic models* incorporate random fluctuations over time. Stochastic models account for random changes in epidemics that may arise over time.

The compartmental model also known as the Susceptible, Infected and Recovered (SIR) model is commonly used to model and analyse the spread of infectious diseases since the base model can be adapted to account for different properties of infectious diseases. The stochastic compartmental model can be used to calculate the likelihood of an epidemic, the stationary distribution of the number of people infected (if it exists) and the probability distribution of the final size of an epidemic. The compartmental model retains the assumption of population-wide homogeneous mixing. This assumption is not realistic and can affect the analysis of the models. A compartmental model that accounts for this, such as a model with the spatial component is much better suited for modelling the situation in a country such as South Africa. South Africa's population varies greatly by location and a model which includes this spatial effect will be more useful in the efficient allocation of resources.

1.1 Literature Review

Infectious diseases can be transferred directly or indirectly from an infected host to a susceptible host (Barreto et al. (2006)). Until the twentieth century, infectious diseases were the most serious health issue in the world (Barreto et al. (2006)). Since then a focus of attention on infectious diseases by various organisations and institutions has resulted in many beneficial changes in the spread and control of infectious diseases (Nelson and Williams (2013)). In 1979, smallpox was the first infectious disease to be eradicated and in 2006, the worldwide incidence of polio was reduced by 99% (Russell (2011)). Improved sanitation, childhood immunisations and antibiotics lead to the reduction in the spread of many infectious diseases (Reingold (2000)).

Human immunodeficiency virus (HIV), ebola virus disease (EVD), severe acute respiratory syndrome (SARS), tuberculosis (TB), malaria, and, most recently, COVID-19 are all emerging infectious diseases that we face today. Ongoing efforts have been made to control these modern diseases. Even though a lot of progress has been made, challenges such as antibiotic resistant pathogens, adherence failure and poor health policies affect the control of these diseases (Nelson and Williams (2013)). Six factors explaining the emerging diseases have been identified. Human demographics and behaviour, economic growth and land use, international travel and trade, microbial adaptation, and changes and breakdowns in public health interventions are all factors to consider (McMichael (2004)).

Infectious diseases are a burden on global economics and public health. Developing countries, however, seem to be worst affected. In these countries, infectious diseases and nutritional deficiencies dominate the causes of morbid-

ity (Boutayeb (2006)). Many infectious diseases such as malaria, HIV, and TB belong to the top ten causes of death in low-income countries (World Health Organization (2018)) and South Africa is no exception. According to the Centers of Disease Control and Prevention (CDC); TB is the biggest contributor to deaths in South Africa. Other leading causes of death in South Africa include HIV, influenza and pneumonia. The country inherited a fractured health system which was further burdened with HIV/AIDS (Coovadia et al. (2009)). HIV/AIDS was greatly underestimated and the lack of proper policies led to a failure of the healthcare system (Coovadia et al. (2009)).

Mathematical modelling is a useful method for simulating and mirroring real-life situations. Modelling the spread of infectious diseases is a key application of mathematics. It is a cost-efficient and rapid way of comparing strategies to reduce the spread of infectious diseases (Maude et al. (2012), Silal et al. (2014)). There are several types of mathematical models utilised in predicting the spread of infectious disease, both within the host and amongst susceptible individuals (Gutierrez et al. (2015)). The models can be broadly categorised as one of three methodologies; statistical-based models, mathematical state-space models and machine-learning models (Siettos and Russo (2013)). Epidemic surveillance is carried out using statistical approaches such as spatial models, regression techniques, and time series. Mathematical models are often used to forecast the progression of infectious diseases. With the advance of computers and technology, machine-learning methods make use of the wealth of online and digital data to forecast the evolution of epidemics (Siettos and Russo (2013)). In this thesis, the focus is on the mathematical state-space models; specifically, one of the most widely used models, the compartmental or SIRS model. This model has a number of individuals in each state of the disease. It includes the susceptible population (S), the infected population (I), and the recovered population (R). People can move between the compartments S, I and R at certain rates. This basic model can be modified in a variety of ways to account for a wide variety of diseases (Silal et al. (2016)).

The first compartmental model was developed by David Bernoulli in the late eighteenth century to model the gain in life expectancy after implementing control measures against smallpox (Allen et al. (2008)). The deterministic compartmental model was formalised by the seminal work of McKendrick and Kermack in 1927. In their thesis they derive integro-differential equations to describe the spread of epidemics such as the plague and cholera in a closed population. Frost(1886-1966) developed a modified mathematical model of the one proposed by Soper. The Reed-Frost model is one of the simplest discrete time stochastic models which describes the evolution of an infection over generations. Bartlett developed the continuous time stochastic SIRS model (Siettos and Russo (2013)). Bartlett pointed out that the deterministic models are insufficient. He proposed that mastering difficulties that arise due to the complexity of stochastic models should be aided by Monte-Carlo methods. Bailey et al. (1975) provides detailed work on the discrete and continuous SIRS

models, along with the estimation of parameters.

Since then, mathematical modelling has evolved from the simple SIRS model to complex compartmental models (Jing et al. (2018), Meng et al. (2016), Jacquez and Simon (1993)). Compartments are added to include features of infectious diseases, such as an exposed compartment for diseases which have a significant latent period. The SEIR model is the name given to this model. The base SIRS model can also be adjusted by incorporating demographic properties of the data such as births and deaths. Some diseases such as influenza display seasonality. The SIRS model can be tweaked to account for seasonality in such diseases by making the transition rate a sinusoidal function of time. Another important factor affecting the susceptibility and infectibility of an individual is the age-structure. Many deterministic and stochastic compartmental models have been developed to include the effects of age. Compartmental models have also been developed to include interventions and preventative measures such as vaccinations. Many modifications have been applied to these models giving rise to a selection of more sophisticated and complex models.

Even though these models are versatile and include various key aspects of the disease, many of them retain the assumption of population-wide parameters, which in turn limits the analysis of these models (Chen et al. (2014)). To make the models a more realistic representation, a version which includes the spatial component is more appropriate. The spatial interaction plays a crucial role in the modelling of the spread of an infectious disease. In the modelling of the spread of an infectious disease, taking into account spatial interaction is crucial. The heterogeneity caused by the segregation of populations is one of the main sources of heterogeneity in epidemic modelling (Meakin and Keeling (2018)). Models that include interactions between different sub-populations are known as a meta-population models (Mishra et al. (2010)) and it is one of the ways that spatial dependencies can be included. Another model which accounts for the spatial properties describes the spread of infectious disease in spatially-structured populations by reaction diffusion equations (Allen et al. (2008)).

Although the compartmental model has both deterministic and stochastic formulations, this thesis will focus on the stochastic models. With the use of regression models, time-series, and generalised linear functions, there are several ways to stochastically model the spread of infectious diseases. Stochastic models, though more complex than the deterministic counterparts, are more realistic and asymptotically tend to the deterministic models (Dietz (1967), Mollison (1995), Daley and Gani (2001)). Some other desirable aspects that can be obtained from stochastic models is the ability to calculate the probability of a disease outbreak, the stationary probability distribution of the number of infected (if it exists), the final size distribution and the expected duration of an epidemic (Allen (2010)). These properties are useful for better planning preventative measures and the efficient allocation of resources. Using

stochastic processes is another method which one can apply to include stochasticity. Branching processes, birth and death processes, competition processes and finite Markov chains are some of the stochastic processes which are useful in modelling the spread of infectious diseases (Dietz (1967), Allen (2010)). Stochastic compartmental models are based on these stochastic processes.

1.2 Problem Statement

Stochastic models such as birth and death processes and branching processes have been applied in epidemiology to better understand the spread of diseases, and how to plan preventative measures. We can distribute health services effectively and reasonably efficiently if we have a good understanding of the nature of an infectious disease. Mathematical models are important for modelling the spread of infectious diseases because they capture the important features or mechanisms and allow simulation of real-world scenarios. Additional features of diseases can be included by adapting and enhancing the stochastic compartmental model. The spatial heterogeneity of infectious disease transmission is one such feature. This is an important factor to consider when modelling the spread of infectious diseases. A spatio-stochastic model would be required to better understand the spread of infectious diseases. This model can be applied to data to study the progression of diseases both spatially and temporally.

1.3 Aim and Objectives of the Study

1.3.1 Aim of the Study

The main aim is to establish and derive a spatio-stochastic model of infectious disease transmission.

1.3.2 Objectives

The objectives of this research are to:

- **Review and apply current compartmental models to the spread of disease**

This involves reviewing and exploring compartmental models to view the potential for improvement and enhancement of these models. The versatility of compartmental models are also demonstrated with the use of applications.

- **Justify the use of a spatial model for the spread of infectious diseases**

The spatial autocorrelation of an infectious disease in South Africa is studied. The results confirm the existence of spatial autocorrelation and hence the need for a model which accounts for the heterogeneity.

-
- **Develop a spatio-stochastic model of the spread of infectious diseases**

The current stochastic discrete models used for the spread of infectious diseases are reviewed and then further developed to derive a spatio-stochastic model. This model is further extended to account for the different modelling scenarios.

- **Compare developed spatio-stochastic models with existing models**

Simulations of the developed model will be compared to the simulations of existing models of the spread of infectious diseases.

- **Apply the spatio-stochastic model on available datasets**

The spatio-stochastic model will then be applied to existing/available datasets. These applications will be used to illustrate the uses of the developed spatio-stochastic model.

1.4 Infectious Disease Terminology

This section defines terms and concepts related to infectious disease epidemiology, which will be referred to throughout this thesis.

As defined early in Chapter 1, a disease caused by an infectious agent is known as an infectious disease. The spread of infectious disease is described by a chain of six elements that are needed for infection of an individual (Krämer et al. (2010)). Figure 1.1 illustrates the chain of infection as proposed by Krämer et al. (2010):

Defining the elements of the chain of infection:

- **Infectious Agent**- this agent is the pathogen causing infection, it can be bacteria, fungi, virus or parasite.
- **Susceptible Host**- an individual member of the population who is at risk of becoming infected by the infectious agent.
- **Portal of Entry**- the route through which the infectious agent enters the body of the host.
- **Mode of Transmission**- this is the movement of an infectious disease.
- **Portal of Exit**- the path taken by the infectious agent to leave the host.
- **Reservoir**- population of organisms in which the infectious agent naturally lives and reproduces.

Other factors such as the vectors and the host-populations with vector borne diseases are also relevant but not included in the chain of infection by Krämer et al. (2010).

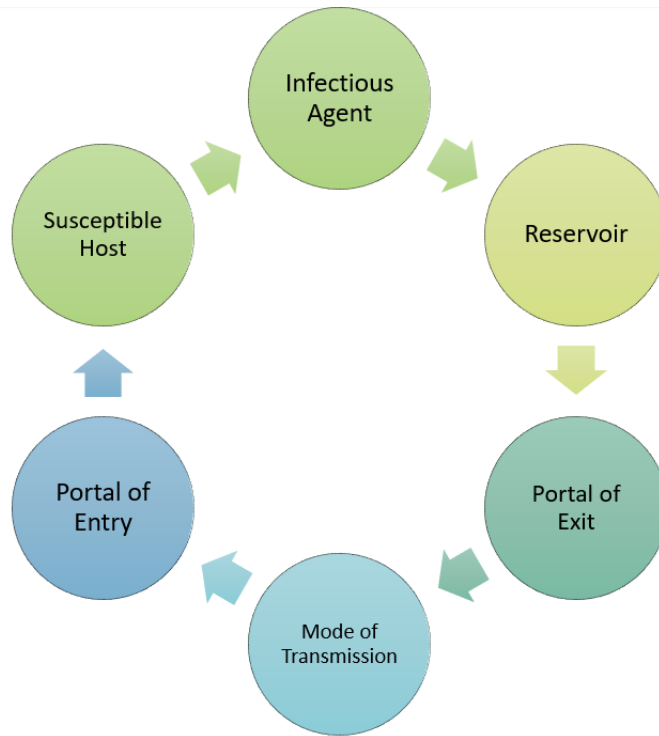


Figure 1.1: *Elements in the Chain of Infection.*

1.4.1 Exposure

When an individual has been in contact with an infectious agent, it is known as exposure. Exposure to infection agents can, but does not always result in disease.

Figure 1.2, adapted from Krämer et al. (2010), describes the possible outcomes of exposure to an infectious disease.

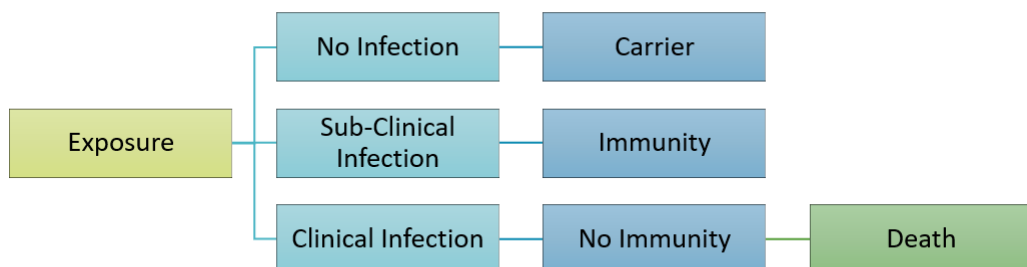


Figure 1.2: *Outcomes of Exposure.*

1.4.2 Infection

Infection is the time moment when a pathogen enters a host and starts reproducing itself there.

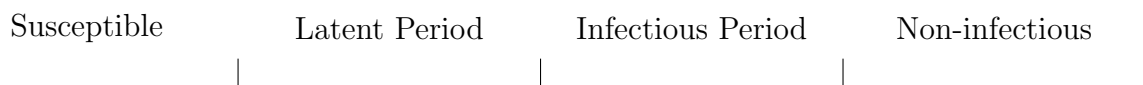
Stages of Infection

The following periods are different stages of infection:

- The **latent period** is the time interval between infection and the start of the infectious phase.
- The **incubation period** is the time between infection and the appearance of clinical symptoms.
- The **infectious period** is the period during which an infected person may spread the disease to a susceptible person.

Figure 1.3 below illustrates the stages of infection and disease.

Infectiousness



Disease

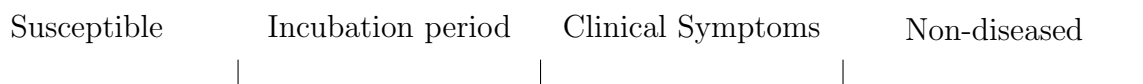


Figure 1.3: *Stages of Infection and Disease.*

1.5 Outline of Thesis

Chapter 2 outlines basic models that are currently used to model the spread of infectious diseases. This includes the SIR or compartmental model which is used in both the deterministic and stochastic form to model a variety of infectious diseases. The models are then extended and applied to model different aspects of the COVID-19 pandemic.

Chapter 3 investigates the spatial autocorrelation that exist in South Africa with the use of TB data. This in turn justifies the development of a model which includes a spatial component.

Chapter 4 begins with a brief introduction into binomial chain models, such as

the Reed-Frost and Greenwood models which are typically as a discrete-time stochastic process to model the spread of infectious diseases. The Tuckwell and Williams formation is then defined. This model is used as the basis of the derivation of the spatial model with its variations. This chapter is the theoretical foundation of the developed model. The spatio-stochastic model is derived under different scenarios and further extended to include the latent or exposed state. The developed models are then compared to the previous model with aid of a simulation study. Finally the spatio-stochastic model is applied to a measles dataset to illustrate the uses of the model.

Chapter 5 concludes the work done in this thesis, summarising the concepts and highlighting the novel work. The limitations and future work is also discussed in this chapter.

The appendix defines and outlines the mathematical tools that are needed to simplify and solve the stochastic models developed. The two main mathematical tools used throughout this thesis are generating functions, to simplify complicated expression and differential equations, which appear throughout the derivations. The appendix also introduces the stochastic models typically used in the modelling the transmission of infectious diseases. Birth and death processes are defined along with properties that are later needed to simplify models for infectious diseases.

1.6 Presentations and Publications

Sections from this thesis has been presented as listed:

- **Conference Presentation-** “*A Spatio-stochastic Model for the Spread of Infectious Diseases* ”, Mathematics Conference and Workshop for Emerging Researchers, Drakensberg, (January, 2022).
- **Article Submission-** “*Justification for the Spatial Modelling of the Spread of Infectious Diseases*”, Journal of Infectious Diseases (May, 2021).
- **Article Publication-** “*A Spatio-stochastic Model for the Spread of Infectious Diseases* ”, Journal of Theoretical Biology (January, 2022).
- **Online Workshop-** “*Infectious Disease Online Workshop*”, (September, 2021). <https://sites.google.com/view/idmworkshop/home>
- **Poster Presentation-** “*Modelling the Spread of Infectious Diseases using Spatio-stochastic Model*”, 30th International Biometric Conference (IBC, 2020).
- **Seminar Presentation-** “*Modelling the Spread of Tuberculosis in South Africa using a Spatio-stochastic Model*”, University of the Witwatersrand(August, 2019).
- **Poster Presentation-** “*Modelling the Spread of Tuberculosis in South Africa*”, Spatial Statistics Conference, Sitges, Spain (July, 2019)

Chapter 2

Basic Models

This chapter outlines basic models that are currently used to model the spread of infectious diseases. This includes the SIR or compartmental model which is used in both the deterministic and stochastic form to model a variety of infectious diseases. The models are then extended and applied to model different aspects of the COVID-19 pandemic and HIV/AIDS.

2.1 Compartmental Models

In Epidemiology, compartmental models are a type of model that is frequently used to model the spread of infectious diseases. The compartmental model is versatile and can be adapted to include a wide range of aspects of infectious diseases by creating additional compartments or by adjusting the transition rates or probabilities. This model is built under the assumption that at any given time point an individual can be classified as susceptible, infected and infectious and recovered or immune.

Let S , I and R represent the number of susceptible, infectious and recovered individuals in a population. Transitions can be made to and from these compartments depending on the dynamics of a specific disease. One such model is the SIR model first developed by Kermack and McKendrick (1927). In this model only susceptible individuals can get infected and, once infected for a certain time period they will recover and develop permanent immunity to the infectious disease. Figure 2.1 is the diagrammatic representation of the SIR compartmental model.



Figure 2.1: *Diagrammatic representation of the SIR model.*

If all individuals, once recovered immediately become susceptible to the disease, then a SIS model would be used. If there was a latent period, which infections only develop after, then a SEIR model would be more appropriate, with a new compartment representing the number of individual who have a latent infection or are exposed.

In this thesis many models are reviewed and studied. Figure 2.2 is a representation of the different models represented and covered in this thesis. The SIR model is a type of compartmental model. One should note that deterministic models can have continuous and discrete formulations as well.

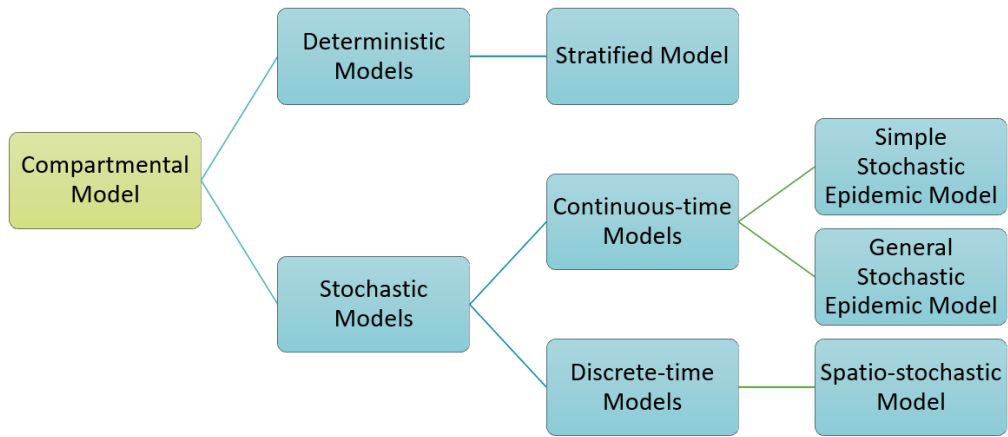


Figure 2.2: *Classification of SIR Models.*

2.1.1 Characterisations and Assumptions

The following characterisations and assumptions are made for the SIR-model

- The population size is fixed, $S + I + R = N$, where N is constant over time. The population is closed, there are no entries or departures into and from the population.
- Susceptible individuals are infected and after getting infected for some time, the individual recovers and becomes immune.
- There is homogeneous mixing of individuals which means each individual is equally likely of contacting another individual.
- Demographic and spatial features are ignored.
- No individual has initial immunity.

-
- Incubation is instantaneous.
 - Individuals within a compartment are identical in behaviour and characteristics, ie the population is homogeneous.
 - Duration of infectiousness of an individual is equal to the duration of the disease.

2.1.2 The Basic Reproduction Number

The basic reproduction number(R_0) is defined as the number of successfully infected contacts during the period of infectiousness in an entirely susceptible population from one infected individual.

R_0 is a measure of the severity of an epidemic. If $R_0 < 1$, then one individual infects less than one other person, and the endemic will die. If $R_0 > 1$, then one individual infects more than one person and the endemic increases.

2.2 The Deterministic SIR Model

The deterministic SIR model, also known as the general epidemic model, describes a disease where individuals can recover and in doing so develop permanent recovery. The number of susceptible individuals can decrease when an individual becomes infected. The number of infected individuals at a particular time can increase if a susceptible becomes infected or it can decrease if an infected person recovers. Diseases such as chicken pox and mumps would follow this model.

Let $S(t)$, $I(t)$ and $R(t)$ denote the mean number of susceptibles, infected and recovered individuals in the population at time t . Let N be the total number of individuals in the population. Since the population is closed, we assume that $S(t) + I(t) + R(t) = N$.

The general epidemic models is thus defined by the following set of differential equations:

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N} \quad (2.1)$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t) \quad (2.2)$$

$$\frac{dR(t)}{dt} = \gamma I(t), \quad (2.3)$$

with initial conditions: $S(0) = s_0$, $I(0) = i_0$ and $R(0) = 0$. and β is the transmission rate and γ , the recovery rate. These equations were first proposed by Kermack and McKendrick (1927)

The term $\beta S(t)I(t)$ is based on the fact that susceptibles must be in contact with infected individuals to become infected. This along with the assumption of a homogeneous population and homogeneous mixing implies that infections will occur at a rate proportional to the product of the number of susceptibles and infected at that time. This is also known as the *mass-action* law which was first formulated in the early nineteenth century by William Hamer(1862-1932).

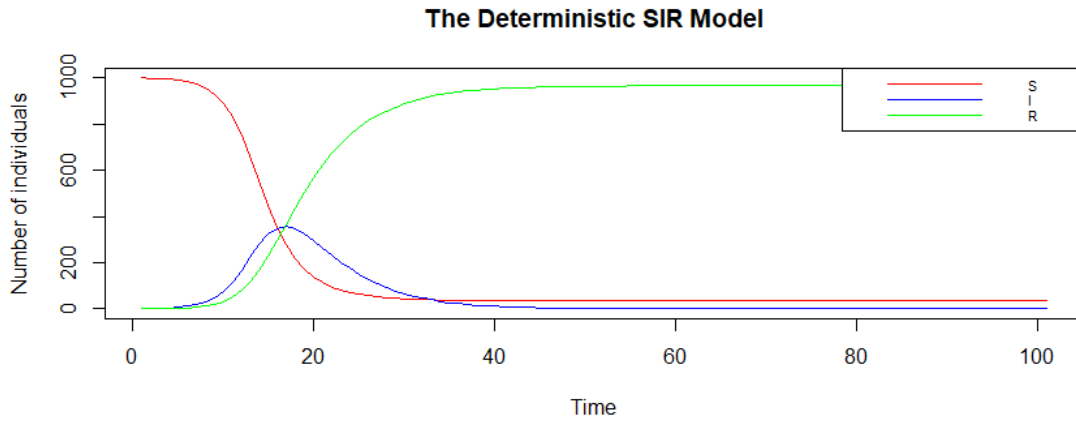


Figure 2.3: *Simulation of a Deterministic SIR model with $\beta = 0.7$, $\gamma = 0.2$ and $N = 1000$.*

Figure 2.3 illustrates the progression of a deterministic model with those particular parameters. Initially, almost all of the population is susceptible with at least one individual being infected. The infected individual spreads the disease to the susceptible individuals and over time the number of susceptible decreases and eventually the number of recovered individuals increases. The rate and severity of the epidemic will depend on the values of the parameters chosen.

2.2.1 Properties of the SIR Model

The Basic Reproduction Number

To derive the basic reproduction number for the deterministic SIR model, we will consider when a disease is epidemic. This occurs when the number of people infected increases:

$$\frac{dI(t)}{dt} > 0 \quad (2.4)$$

Substituting equation 2.2,

$$\frac{\beta S(t)I(t)}{N} - \gamma I(t) > 0 \quad (2.5)$$

$$\frac{\beta S(t)I(t)}{N\gamma} > I(t). \quad (2.6)$$

Initially, everyone except the initially infected is susceptible, ie $S(t) \cong N$, thus

$$\frac{\beta}{\gamma} > 1 \quad (2.7)$$

Thus, the reproduction number, $R_0 = \frac{\beta}{\gamma}$.

This implies that if $\beta > \gamma$, an epidemic will occur and if $\beta < \gamma$, a minor outbreak will occur.

The Final State Distribution

Once an infectious disease has run its course, the final state is of interest to gauge the severity and extent of an infectious disease. The final state can be derived by dividing equation (2.1) by (2.3) that gives:

$$\frac{dS(t)}{dR(t)} = -R_0 S(t). \quad (2.8)$$

Integrating the differential equation with initial values $S(0) = s_0$ and $R(0) = 0$, yield:

$$S(t) = s_0 e^{-R_0 R(t)}. \quad (2.9)$$

Taking the limit as $t \rightarrow \infty$, and replacing $S(\infty) = N - R(\infty)$ results in:

$$N - R(\infty) = s_0 e^{-R_0 R(\infty)}. \quad (2.10)$$

The final size of the recovered individual is thus related to the basic reproduction number.

Figure 2.4 was generated assuming a population of 1000 individuals where one individual was initially infected. It shows the relationship of the basic reproduction number with the final size of the immune population, $R(\infty)$.

The COVID-19 Pandemic and Flattening the Curve: A Glance at South Africa

In this study, the current COVID-19 pandemic, that changed the way that many people lived is studied. The outbreak was first reported in Wuhan

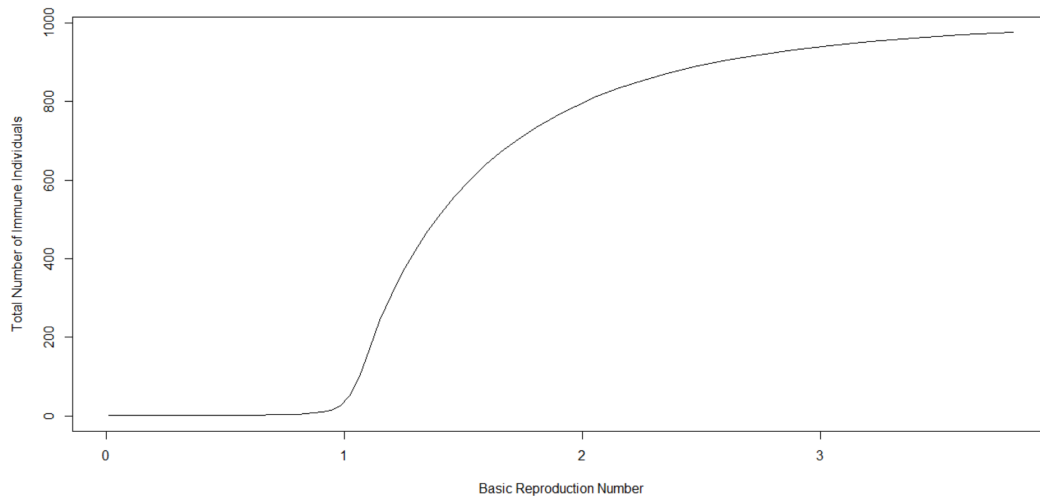


Figure 2.4: Plot of $R(\infty)$ vs R_0 .

Province, China towards the end of December 2019. Over the following 100 days, the disease had spread across the world and some first world countries became epicenters of the disease.

The COVID-19 is a type of corona disease said to have originated from the consumption of bats. It is a respiratory disease which is spread through droplets (sneezing and coughing) and through infected or contaminated objects. Initially people were cautioned to use safe practises such as the regular washing of hands, not touching your face, coughing or sneezing into the crease of your elbow, and remaining at home if you were feeling sick or exposed. This progressed to the encouragement of social-distancing. This involves avoiding social gatherings and interactions unless absolutely necessary, limiting the number of people at social events and maintaining a 1.5m distance between people in public spaces. However, South Africa had many imported cases and the number of infections were increasing rapidly.

The growth infectious diseases are usually exponential and the South African government has become increasingly concerned with the effect on the health care system. The disease death rate is around 1% to 5 % and the elderly and people with compromised immune systems are at most risk. Around 15% of all cases are severe enough to require medical attention. Many of these patient require ventilators to assist with breathing. If the spread of the disease continues exponentially, the healthcare systems will be overwhelmed and will not be able to accommodate the increasing amount of COVID-19 patients or the usual patients. This is how the term ‘flattening the curve’ has gained attraction. This is an epidemiological term referring to slowing the rate of infection to allow for the health infrastructure to keep up.

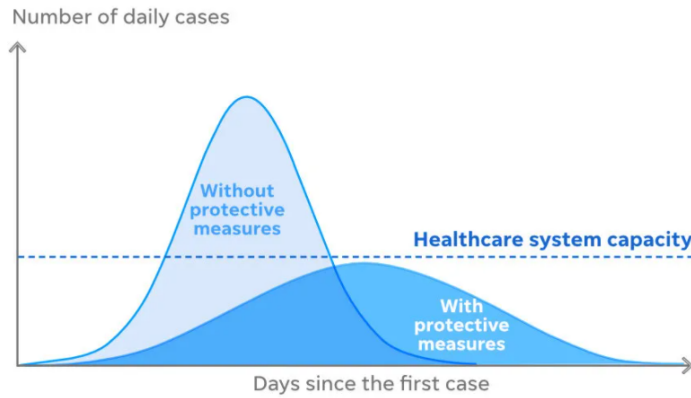


Figure 2.5: *Illustration of flattening the curve. Source: CDC*

An infectious disease's infection rate is high at first and then gradually increases until it reaches a peak. It then decreases and dies off. From Figure 2.5, it can be seen that the “flattening of the curve” is essentially decreasing the initial rate of infection and delaying the peak so that health care systems can keep up and also prepare for the increase in hospitalisations. The government implemented a 21 day lockdown, which was then extended by an additional 15 days. This indeed had the intended effect of flattening the curve in South Africa allowing for more time for preparation for the pandemic. Although flattening the curve extends the time until the infection dies out, it creates time which allows better management of the disease.

A simple deterministic model has been developed from the model proposed by Kermack and McKendrick (1927) to better understand this concept of flattening the curve and the effect on the progression and spread of the COVID-19.

The Model

The model that is being used to analyse the effect of social distancing and different lock-down strategies on the spread of the disease is a simple SIRD model. Since this is a novel disease with many uncertainties regarding the dynamics, a basic model is used to emphasise the known aspects of the disease. The simple SIRD model will be further extended to include other aspects of COVID-19, like the incubation period. Figure 2.6 is a representation of the transition graph of the SIRD model with the addition of a parameter, ψ , to account for social distancing.

The COVID-19 is a novel disease which behaves differently in different countries and for different variants and a lot is yet to be discovered.

The following assumptions have been made:

1. The population is closed. There are no births, deaths and there is no migration in and out of the population.

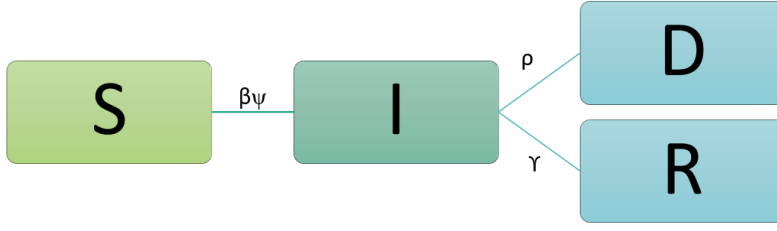


Figure 2.6: *Diagrammatic representation of the SIRD model.*

2. There are four states, susceptible (S), infected (I), recovered (R) and Dead (D).
3. There is homogenous mixing of individuals.
4. Once individuals recover, they infer permanent immunity.
5. The median infectious period is approximately 14 days (Gomes (2020)).
6. The mode value of the estimated R_0 was taken as 2.7 (Silal et al. (2020)).

The system of differential equations of the model are:

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)\psi}{N} \quad (2.11)$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)\psi}{N} - \rho I(t) \quad (2.12)$$

$$\frac{dR(t)}{dt} = \gamma I(t) \quad (2.13)$$

$$\frac{dD(t)}{dt} = \rho I(t), \quad (2.14)$$

with initial conditions: $S(0) = s_0$, $I(0) = i_0$ and $R(0) = 0$. and β is the transmission rate and γ , the recovery rate, ρ is the disease death rate and ψ is a social-distancing parameter introduced to monitor the effect of social-distancing and other flattening curve measures. $0 < \psi < 1$ where $\psi = 0$ when no interventions are taken or are successful and $\psi = 1$ when the interventions completely stop the transmission of the disease.

Figure 2.7 is an illustration of the impact that changing the social distancing parameter had on the progression of the diseases. It can be seen that when no interventions are taken the disease progresses quickly, reaching a peak early on. As you increase the social distancing parameter, the length of the disease increases, however, the total number of recovered and deaths do not change until the parameter is quite big, ie when all interventions are adhered to.

Sensitivity Analysis of ψ

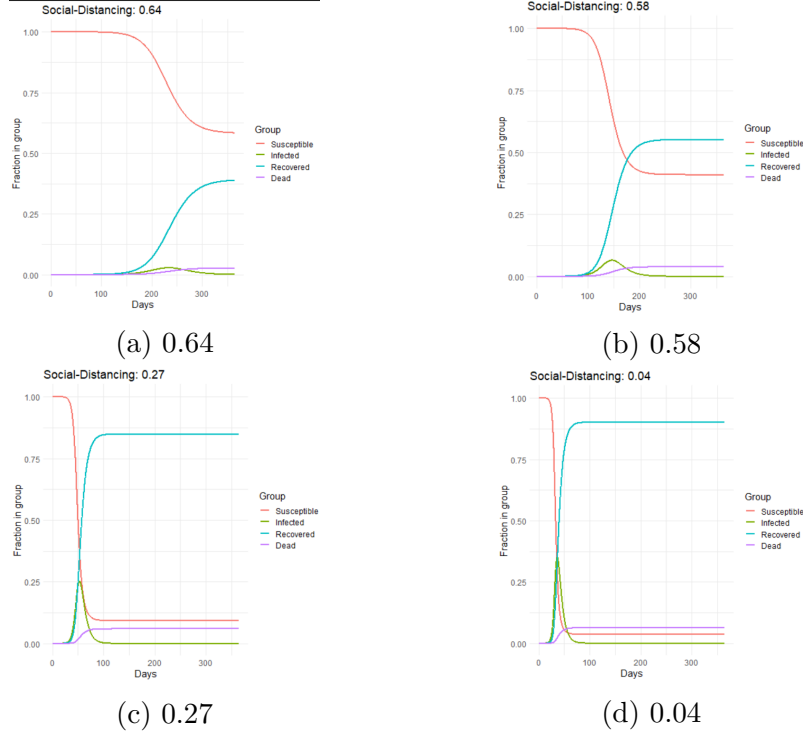


Figure 2.7: *Illustration of the effect of social distancing.*

A single parameter sensitivity analysis was conducted on ψ , the social distancing parameter and the proportion of dead individuals. 10 000 Simulations were run, varying ψ from a standard uniform distribution.

From Figure 2.8, it can be seen that the proportion of dead individuals is sensitive to the choice of ψ . This is an indication that the amount of social-distancing can impact the number of deaths. It can also be observed that when $\psi > 70\%$, the proportion of deaths is zero. This implies, that the social distancing only at most requires a 70% reduction infections and more stringent policies than that will not result in fewer deaths.

The Basic Reproduction Number of the COVID-19 Pandemic

An epidemic occurs when the number of infections increases over time, ie

$$\frac{dI(t)}{dt} > 0 \quad \forall t > 0. \quad (2.15)$$

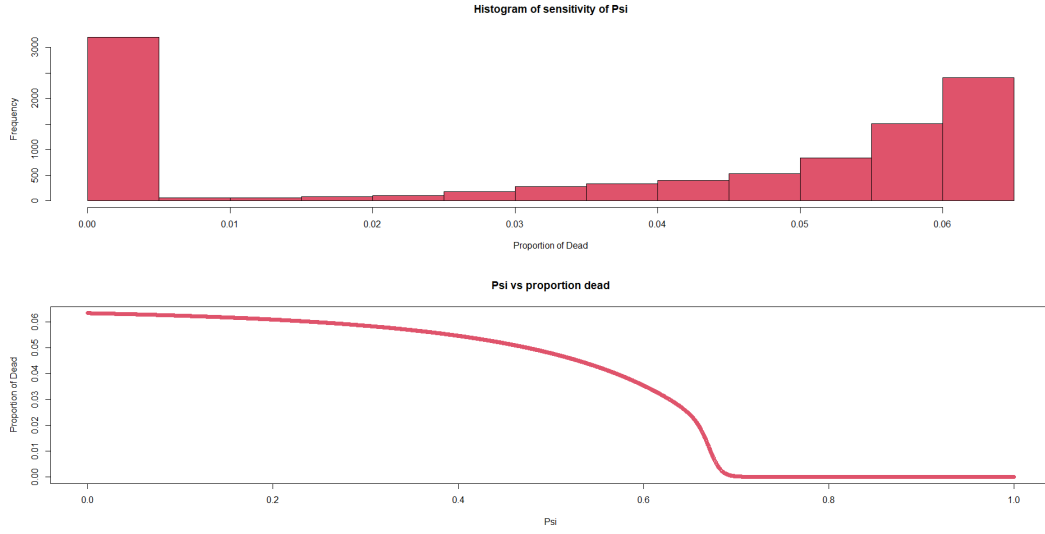


Figure 2.8: *Sensitivity analysis with regard to ψ .*

Thus, R_0 of this model can be calculated as follows:

$$\frac{\beta S(t)I(t)\psi}{N} - \gamma I(t) - \rho I(t) > 0 \quad (2.16)$$

$$\frac{\beta S(t)I(t)\psi}{N} > (\gamma + \rho)I(t) \quad (2.17)$$

$$\frac{\beta\psi}{\gamma + \rho} > 1, \quad (2.18)$$

since $S(t) \cong N$, when the outbreak just begins.

Therefore, $R_0 = \frac{\beta\psi}{\gamma + \rho}$. This social distancing parameter affects the R_0 and by increasing the parameter, we can decrease the R_0 , so that an epidemic can be avoided.

2.2.2 The Deterministic Stratified SIR Model

A fundamental assumption of the general epidemic model is that the population within each compartment is homogeneous. This assumption does not hold true for the majority of infectious diseases and therefore this deterministic model is not the most appropriate.

One solution is proposed in Daley and Gani (2001) is to stratify the population. The stratification can be done according to age, gender, economic class or even spatially.

Assume the population is broken into m sub-populations. Then $S_j(t)$, $I_j(t)$ and $R_j(t)$ represent the number of susceptible, infected and immune individuals in the j^{th} sub-group at time t , where $j = 1, 2, \dots, m$ and $t \geq 0$.

Then the differential equations for the j^{th} sub-group as derived by Daley and Gani (2001):

$$\frac{dS_j(t)}{dt} = -S_j(t) \left(\frac{\beta_{1j}I_1(t)}{N_1} + \frac{\beta_{2j}I_2(t)}{N_2} + \dots + \frac{\beta_{mj}I_m(t)}{N_m} \right) \quad (2.19)$$

$$\frac{dI_j(t)}{dt} = S_j(t) \left(\frac{\beta_{1j}I_1(t)}{N_1} + \frac{\beta_{2j}I_2(t)}{N_2} + \dots + \frac{\beta_{mj}I_m(t)}{N_m} \right) - \gamma_j I_j(t) \quad (2.20)$$

$$\frac{dR_j(t)}{dt} = \gamma_j I_j(t), \quad (2.21)$$

with initial conditions: $S_j(0) = s_{j0}$, $I_j(0) = i_{j0}$ and $R_j(0) = 0$, for $j = 1, 2, \dots, m$.

The contact rate β_{ij} is the pairwise infection rate between a susceptible in the i^{th} sub-group and an infected in the j^{th} sub-group. The recovery rate, γ_i is the rate at which an individual in the j^{th} sub-group becomes immune.

In subsequent section, the compartmental model will be stratified with respect to age to further look at the impact of age on the model. Even though the stratified model accounts for heterogeneity, deterministic models are not always suitable. They are not suitable in cases when there is uncertainty in the final state, such as in small groups.

Deterministic models also fail to provide standard errors on parameters that have been obtained. Deterministic models provide good estimates of average behaviour of populations but for aspects such as disease extinction, stochastic models are better suited.

The COVID-19 Pandemic and Stratification: Investigating the effect of age in South Africa

Background

The COVID-19 pandemic developed differently throughout the world. The spread of the disease and the number of deaths differed in most of the countries. This could be due to a variety of factors such as the density of the population, the systems in place, living conditions and how the outbreak was managed. One of the earliest observations about the behaviour of the disease was that it is more fatal to the elderly and those with compromised immune systems. The death rate amongst the elderly were significantly higher when compared to other individuals. This could be a reason why ageing population such as many European countries were worse hit than younger populations like those found in African countries (Stein (2020)). The following table summaries the case fatality rate as given by Stein (2020).

In this section, the deterministic stratified compartmental model is used to analyse the spread of COVID-19 amongst the different age groups. South

Table 2.1: Case Fatality Rates per age group.

Age Group	Case Fatality Rate
10-19	0.2%
20-29	0.2%
30-39	0.2%
40-49	0.4%
50-59	1.3%
60-69	3.6%
70-79	8%
80 $\geq$	14.8%

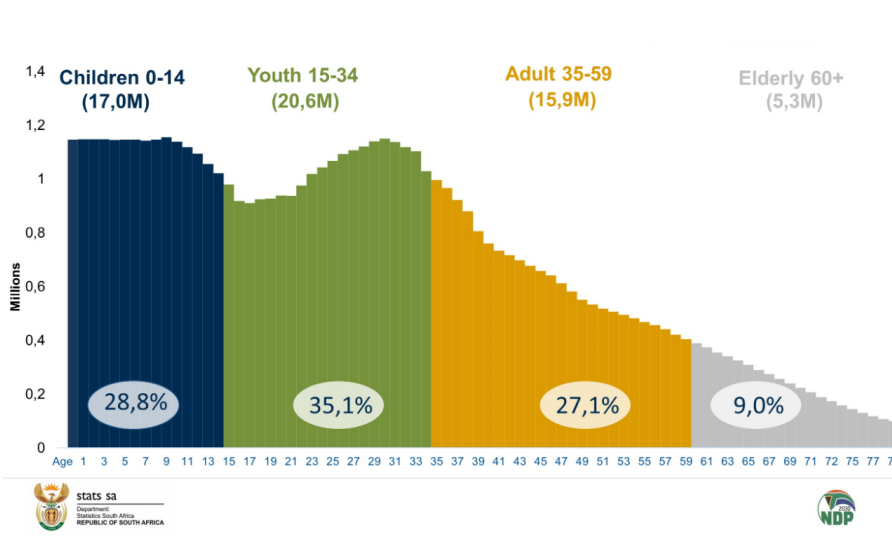


Figure 2.9: *South African population per age.*

Africa, like many other African countries has a relatively young population. The age distribution of South Africa according to Stats SA is graphed in Figure 2.9. Almost two thirds of the population is under 35 years old where the case fatality is around 0.2%. The age-stratified will be used to model the spread of the disease during the first two waves. The third and fourth wave age-stratified model will be adjusted by adding a compartment and transitions to account for the vaccinations. The vaccination roll-out in South Africa was implemented according to age.

The age-stratified model will be split into 4 groups as given by Table 2.2: The ages were stratified in this manner to look at the impact on the spread and later vaccination roll out on the South African population. The first group holds a large proportion of the population with the majority containing individuals attending school or university. The last group is the most vulnerable to the “so-called” COVID-19 fatalities.

Table 2.2: The age-groups used in the age-stratified compartmental model.

Age Group Number	Range(years)	Number of individuals
1	0-19	21 994 779
2	20-44	24 106 854
3	45-59	8 144 767
4	60+	4 640 496

Methodology

In this section we build an age-stratified model and use it to further develop a vaccination model for the spread of COVID-19 in South Africa. The first step is to build the model. The model that will be used is not the simple model defined in Section 2.2, but will have many other compartments to account for the various aspects of the disease. The model is built using the model that is used by the National Institute of Communicable Diseases (NICD)(Silal et al. (2020)) as a basis. This model has been altered and stratified by age. The model will then be calibrated using the case data from the first wave. The case data from the second wave will then be used to test the model. The COVID-19 vaccination was not available until after the second wave. The model is thus updated to include the vaccination. The model is then recalibrated with the third wave's data.

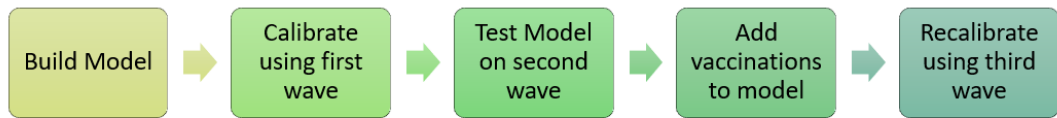


Figure 2.10: *Steps in the modelling process.*

The following assumptions are made in the stratified model:

1. All individuals within a group are identical and there is homogeneous mixing within groups.
2. Asymptomatic individuals are less infectious than those with symptomatic infections.
3. Individuals which are hospitalised in a general ward or in ICU cannot spread the infection as they are hospitalised and separated from the population.
4. Only individuals hospitalised die as they have conditions severe enough to be fatal.

-
5. The death rate for each group is constant within a group but varies between groups.

Table 2.3: Description of the different compartments used in the SEIR model.

Compartment	Description
S	Number of susceptible individuals
E	Number of exposed individuals
A	Number of individuals with an asymptomatic infection
I	Number of individuals with a mild infection
GW	Number of infected individuals admitted into a general ward
ICU	Number of infected individuals in ICU
R	Number of recovered individuals
D	Number of dead individuals

Age-stratified SEIR model

The model dynamics are represented by the transition graph in Figure 2.11:

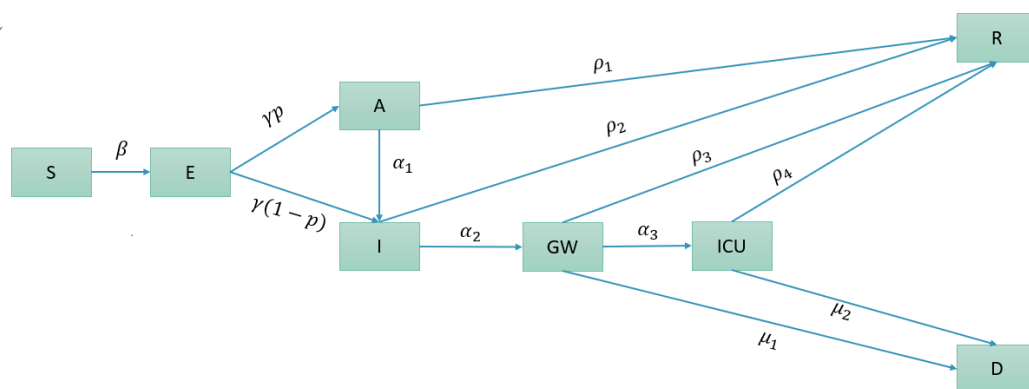


Figure 2.11: *Model structure of SEIR model.*

The system of differential equations of the model are given by:

$$\frac{dS_i(t)}{dt} = -S_i(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_j(t) + GW_j(t) + r_a A_j(t))}{N_j} \right] \quad (2.22)$$

$$\frac{dE_i(t)}{dt} = S_i(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_j(t) + GW_j(t) + r_a A_j(t))}{N_j} \right] - \gamma E_i(t) \quad (2.23)$$

$$\frac{dA_i(t)}{dt} = p\gamma E_i(t) - \rho_1 A_i(t) - \alpha_1 A_i(t) \quad (2.24)$$

$$\frac{dI_i(t)}{dt} = (1 - p)\gamma E_i(t) + \alpha_1 A_i(t) - \alpha_2 I_i(t) - \rho_2 I_i(t) \quad (2.25)$$

$$\frac{dGW_i(t)}{dt} = \alpha_2 I_i(t) - \alpha_3 GW_i(t) - \rho_3 GW_i(t) - \mu_{1i} GW_i(t) \quad (2.26)$$

$$\frac{dICU_i(t)}{dt} = \alpha_3 GW_i(t) - \rho_4 ICU_i(t) - \mu_{2i} ICU_i(t) \quad (2.27)$$

$$\frac{dR_i(t)}{dt} = \rho_1 A_i(t) + \rho_2 I_i(t) + \rho_3 GW_i(t) + \rho_4 ICU_i(t) \quad (2.28)$$

$$\frac{dD_i(t)}{dt} = \mu_{1i} GW_i(t) + \mu_{2i} ICU_i(t), \quad (2.29)$$

where $i = 1, 2, 3, 4$, represents each age class.

Parametrisation

The model was parametrised using both literature and available data (see Table 2.4).

1. Age-dependent transmission rates and the contact matrices

An important set of parameters are the age-dependent transmission rates (β_{ij}). This was obtained by using the following relationship:

$$\beta_{ij} = \frac{R_0 c_{ij}^*}{D}, \quad (2.30)$$

where c_{ij}^* is a scaling factor derived from the contact matrix and D is the mean duration of the infectious period.

The contact matrix was obtained from Prem et al. (2021). In their research, the authors estimate the contact matrices for 152 countries including South Africa. An illustration of the contact matrix is given by Figure 2.12.

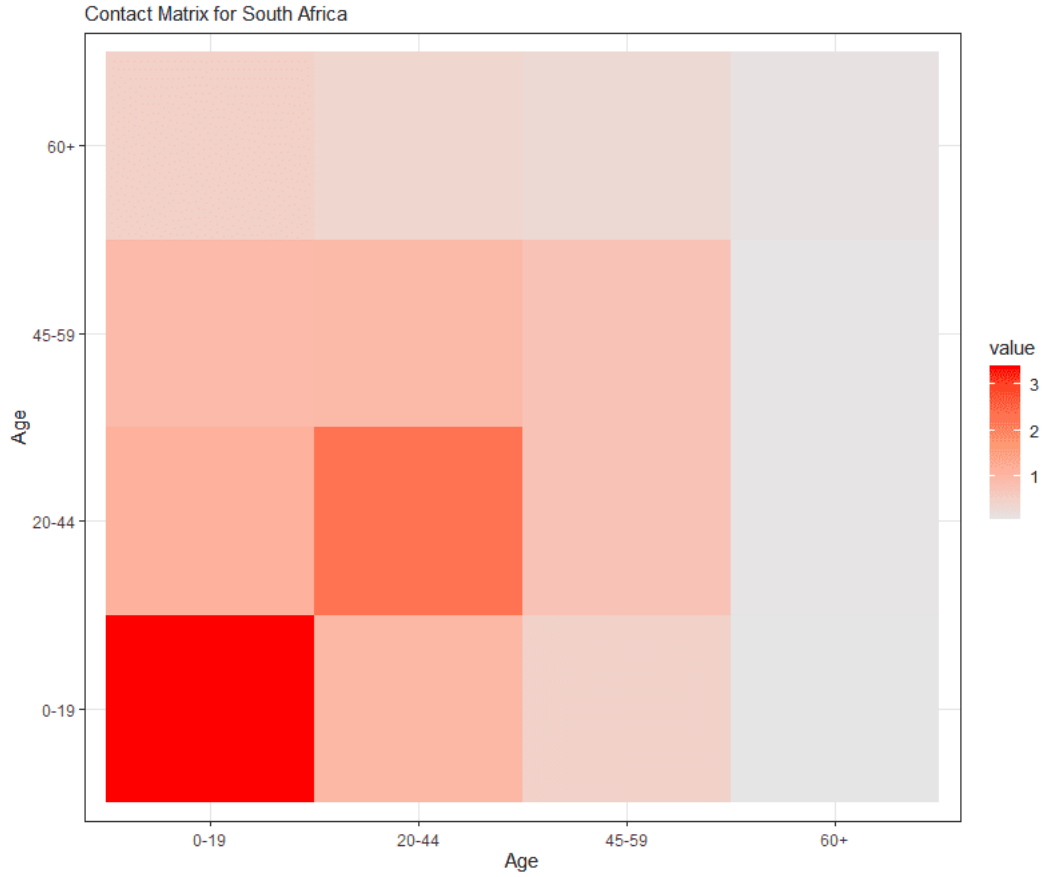


Figure 2.12: *Matrix plot of the 4 age groups.*

2. Key Parameters from Literature

Table 2.4: Model Parameters.

Parameter	Description	Estimate	Reference
r_a	Relative infectivity of asymptomatic cases	0.75	Silal et al. (2020) Fabris-Rotelli et al. (2021)
$\frac{1}{\gamma}$	Number of days in incubation	$(2)^{-1}$	Fabris-Rotelli et al. (2021)
p	Proportion of asymptomatic infections	0.75	Silal et al. (2020), Fabris-Rotelli et al. (2021)
$\frac{1}{\rho_1}$	Number of days with asymptomatic infection	$(7)^{-1}$	Silal et al. (2020), Fabris-Rotelli et al. (2021)
$\frac{1}{\rho_2}$	Number of days with symptomatic infection	$(5)^{-1}$	Silal et al. (2020)
$\frac{1}{\rho_3}$	Number of days with infection in general ward infection	$(12)^{-1}$	Silal et al. (2020)
$\frac{1}{\rho_4}$	Number of days with asymptomatic infection	$(16)^{-1}$	Silal et al. (2020)
$\frac{1}{\alpha_1}$	Number of days until symptoms develop	$(2)^{-1}$	Silal et al. (2020)
$\frac{1}{\alpha_2}$	Number of days until hospitalisation	$(5)^{-1}$	Silal et al. (2020)
$\frac{1}{\alpha_3}$	Number of days until ICU	$(4)^{-1}$	Silal et al. (2020)

For the age-stratified model, the key parameters in Table 2.4 are assumed to be the same across all ages. This assumption could be relaxed if there is data available to obtain more up to date estimates.

3. Age-dependent Death Rates

The death rate was obtained by using the case fatality rate for South Africa for COVID-19 as per the Johns Hopkins Coronavirus Resource Center. This was then broken up into age-groups according to the report by Pillay-van Wyk et al. (2020), which has been adjusted for the age-groups and summarised in Table 2.5.

Table 2.5: The age-groups used in the age-stratified compartmental model.

Age Group Number	Range(years)	Proportion of Deaths
1	0-19	0.004
2	20-44	0.1385
3	45-59	0.3035
4	60+	0.554

The death rate for the hospitalised individuals is assumed to be 0.015 and for the individuals in ICU 0.0225 (Ngonghala et al. (2020)). The age dependent death rates were obtained by the product of the proportion of deaths in each category and the corresponding death rates.

Model Initialising and Specifics

1. Lock-down Levels

During the COVID-19 pandemic, the South African government has released different levels of lockdown of the country to curb the spread of COVID-19 cases particularly around the peaks to prevent strain on the health systems.

The R_0 was adjusted by a scaling factor to account for the change in the behaviour of the public. The details are given below:

Table 2.6: The lockdown levels and scaling.

Lockdown Level	Scaling Factor
5	0.7
4	0.75
3	0.8
2	0.85
1	0.9

The same scaling was used for levels and their corresponding adjusted levels.

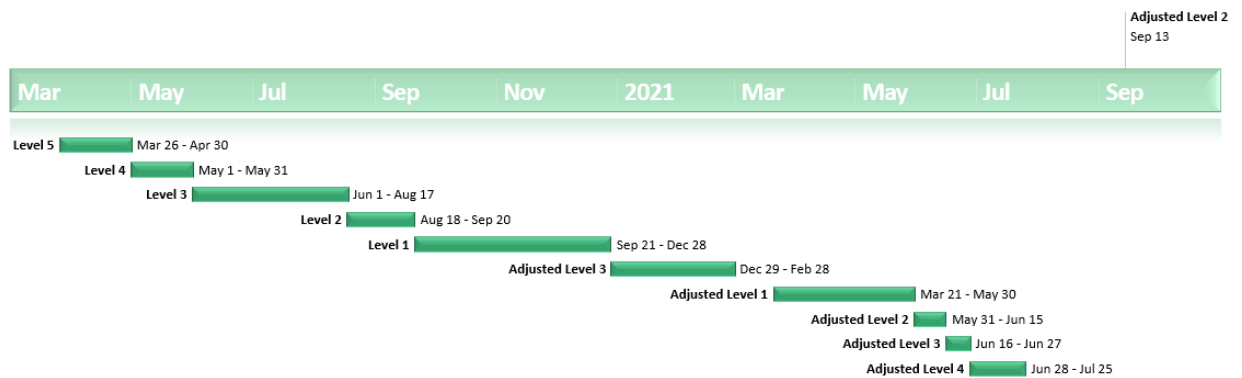


Figure 2.13: *Timeline of Lockdown Levels.*

2. Waves Time frames

Table 2.7: The dates of the waves in South Africa.

Wave Number	Time period
1	13 June 2020- 29 August 2020
2	26 December 2020-13 February 2021
3	10 June 2021-28 August 2021

3. Under-counting and Seeding

When modelling infectious diseases, especially in the context of a global pandemic caused by novel disease, we need to account for the fact that the information that is recorded about the infectious diseases is not the actual observations. The number of cases tested positive is subject to the number of tests done. Since the entire population is not tested, the cases that are reported under-represent the actual cases of infectious diseases. This is made worse by the presence of asymptomatic individuals. Typically, people test only when they have symptoms, and once again the positive cases will be biased towards symptomatic people as they would predominantly be taking tests. This needs to be accounted for in the model. The proportion under-counting is assumed to be a factor of 4.8 (Fabris-Rotelli et al. (2021)) for the purposes of the model. It is also assumed that the number of asymptomatic individuals is 4 times the number of symptomatic infections. This will be included in the seeding and initialising of the model as well as when comparing it to the observed output.

Another area where the data is not representative of the actual scenario is with the recorded deaths. According to Dorrington et al. (2021), the actual deaths from COVID-19 are between 2.5 and 3 times bigger than the recorded official deaths. This estimate is obtained by looking at the excess deaths when comparing the same period from pre-COVID times. This too, is taken into account when fitting and calibrating the model.

4. Variants of COVID-19

Like all viruses, SARS-CoV-2 may change over time. This mutation can have a negligible impact on the properties of the virus. However, some mutations may alter the properties of the virus, this can change how easily the virus spread, the severity of infection, the effectiveness of vaccines and other therapeutic medicines or diagnostic tools (CDC). The World Health Organisation has been working with many field experts, institutions, and researchers to monitor and assess the evolution of the SARS-CoV-2 since January 2020. New variants which pose an increased risk to global public health are now being categorized. According to the WHO, two of the main categories are Variants of Interest (VOI) and

Table 2.8: The lockdown levels and scaling after calibration.

Lockdown Level	Scaling Factor
5	0.7
4	0.8
3	0.83
2	0.88
1	0.92

Variants of Concern (VOC). A Variant of Concern has an increase in transmissibility, or an increase in virulence, or decreases the effectiveness of public health and social measures (CDC). Among the VOCs are the Alpha, Beta, Gamma, Delta, and Omicron variants. Variants of Interest have genetic changes that are predicted or known to affect virus characteristics and identified to cause significant transmission or multiple COVID-19 CLUSTERS in multiple countries. Some VOI include the Lamba and Mu variants.

To account for the variability introduced by variants, the model will be run over different scenarios, each with varying levels of infectivity both greater and less than the base model.

Model Calibration using First Wave

The model was calibrated using the first wave infections data. The reduction in the impact of lockdown levels, proportion of asymptomatic infections and relative infectivity of asymptomatic infections were adjusted until the model represented the observed data. The new values of the lock-down levels that were used in the following models are as follows:

Model Testing using Second Wave

The recalibrated model was then used to test the goodness of fit for different case scenarios. The root mean square error (RMSE) was used to compare the fit of different cases. Since COVID-19 is a novel disease with several mutations, when performing the goodness of fit, the calculations are run over different levels adjusting the baseline for possible differences between the variants. We consider cases where the variant is less and more virulent than the initial variant. Table 2.9 summarises the RMSE for the different scenarios.

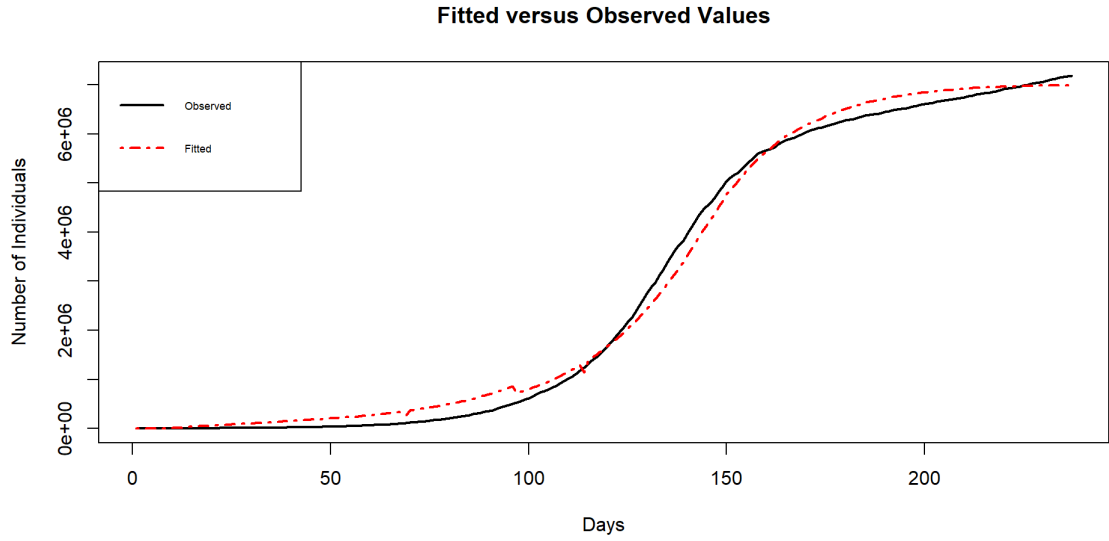


Figure 2.14: *Plot of the actual versus observed infections of the calibrated model.*

Table 2.9: The RMSE of different scenarios.

Case Adjustments	Root Mean Square Error
No lockdown adjustment	369 675.4
10% less virulent	1 347 004
20% less virulent	1 436 863
50% less virulent	1 485 781
10% more virulent	247 478.60
20% more virulent	1750 548
50% more virulent	12 741 041
Unadjusted	1 070 153

The scenarios associated with the best fit among the options were the ones assuming the same virulence, but no effect from lock-down measures and the option considering a variant which is 10% more virulent than the base model. The RMSE increases quite substantially for the other cases and this makes sense since the growth is exponential and the differences from the observed data would quickly diverge.

Addition of Vaccinations

We now add in the vaccinated individuals. Typically, vaccinations are added in as an additional compartment to form a V-SEIR model, however, this formulation is used with vaccinations that are not *leaky*. The term *leaky* refers to the fact that unlike many other vaccines that prevent infection, the COVID-19 vaccine *reduces* infections (Shinde et al. (2021)). Furthermore, it has been

found that the immunity from these vaccinations wane over time (Chemaitelly et al. (2021)). The main benefit of the vaccine is not that it prevents infection, but that it prevents severe and critical cases. These aspects are taken into consideration when building the model and the details are given below.

To include vaccinated individuals, the model has been further stratified between the vaccinated and unvaccinated population with different parameters for infection between them. This formulation is more suitable for the COVID-19 vaccination.

1. Model Assumptions

To simplify the model, the following assumptions in addition to the previous assumptions will be made:

- All individuals within a strata are identical and there is homogenous mixing within groups and amongst subgroups.
- Asymptomatic and vaccinated individuals are less infectious than those who have not been vaccinated or who are symptomatic.
- At least one vaccination dose needs to be given so that people are considered vaccinated.
- Vaccinations reduce the death rate in hospitalised and ICU patients.
- The vaccination immunity is constant throughout a wave and there is no loss of immunity during a 6-month modelling period.

2. Vaccination Strategy

The vaccine strategy in South Africa was rolled out in many phases as indicated by Figure 2.15. The first to be vaccinated were front-line health workers, followed by teachers, inmates and then different age groups starting from the older individuals. In the model, an initial rate per age-group was estimated from the data during February, this mainly applied to health care workers across all ages, this rate was assumed to be the same until the age-specific vaccinations and then those rates were increased accordingly. For the third wave, the elderly over 60 years were predominately vaccinated.

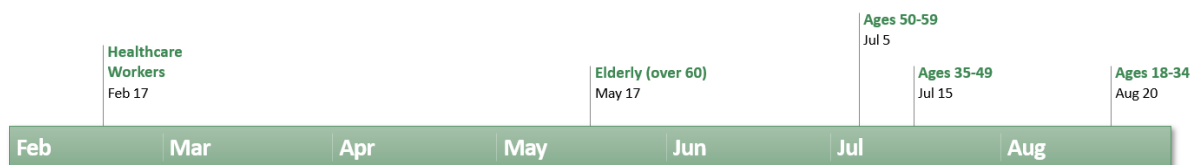


Figure 2.15: *Timeline of South Africa's Vaccination Roll-out.*

3. Model Equations and Parameters

The system of differential equations for the unvaccinated sub-populations, denoted by $\bar{V}$, are given by:

$$\frac{dS_{\bar{V}i}(t)}{dt} = -S_{\bar{V}i}(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_{\bar{V}j}(t) + GW_{\bar{V}j}(t) + r_a A_{\bar{V}j}(t))}{N_{\bar{V}j}} + \sum_{j=1}^4 \frac{\beta_{ij}r_v(I_{Vj}(t) + GW_{Vj}(t) + r_a A_{Vj}(t))}{N_{Vj}} \right] - v_i S_{\bar{V}i}(t) \quad (2.31)$$

$$\frac{dE_{\bar{V}i}(t)}{dt} = S_{\bar{V}i}(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_{\bar{V}j}(t) + GW_{\bar{V}j}(t) + r_a A_{\bar{V}j}(t))}{N_{\bar{V}j}} + \sum_{j=1}^4 \frac{\beta_{ij}r_v(I_{Vj}(t) + GW_{Vj}(t) + r_a A_{Vj}(t))}{N_{Vj}} \right] - \gamma E_{\bar{V}i}(t) \quad (2.32)$$

$$\frac{dA_{\bar{V}i}(t)}{dt} = p\gamma E_{\bar{V}i}(t) - \rho_1 A_{\bar{V}i}(t) - \alpha_1 A_{\bar{V}i}(t) \quad (2.33)$$

$$\frac{dI_{\bar{V}i}(t)}{dt} = (1-p)\gamma E_{\bar{V}i}(t) + \alpha_1 A_{\bar{V}i}(t) - \alpha_2 I_{\bar{V}i}(t) - \rho_2 I_{\bar{V}i}(t) \quad (2.34)$$

$$\frac{dGW_{\bar{V}i}(t)}{dt} = \alpha_2 I_{\bar{V}i}(t) - \alpha_3 GW_{\bar{V}i}(t) - \rho_3 GW_{\bar{V}i}(t) - \mu_{\bar{V}1i} GW_{\bar{V}i}(t) \quad (2.35)$$

$$\frac{dICU_{\bar{V}i}(t)}{dt} = \alpha_3 GW_{\bar{V}i}(t) - \rho_4 ICU_{\bar{V}i}(t) - \mu_{\bar{V}2i} ICU_{\bar{V}i}(t) \quad (2.36)$$

$$\frac{dR_{\bar{V}i}(t)}{dt} = \rho_1 A_{\bar{V}i}(t) + \rho_2 I_{\bar{V}i}(t) + \rho_3 GW_{\bar{V}i}(t) + \rho_4 ICU_{\bar{V}i}(t) \quad (2.37)$$

$$\frac{dD_{\bar{V}i}(t)}{dt} = \mu_{\bar{V}1i} GW_{\bar{V}i}(t) + \mu_{\bar{V}2i} ICU_{\bar{V}i}(t), \quad (2.38)$$

where $i = 1, 2, 3, 4$, represents each age class.

The system of differential equations of the vaccinated population, denoted V , are given by:

$$\frac{dS_{Vi}(t)}{dt} = -S_{Vi}(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_{\bar{V}j}(t) + GW_{\bar{V}j}(t) + r_a A_{\bar{V}j}(t))}{N_{\bar{V}j}} + \sum_{j=1}^4 \frac{\beta_{ij}r_v(I_{Vj}(t) + GW_{Vj}(t) + r_a A_{Vj}(t))}{N_{Vj}} \right] + v_i S_{\bar{V}i}(t) \quad (2.39)$$

$$\frac{dE_{Vi}(t)}{dt} = S_{Vi}(t) \left[\sum_{j=1}^4 \frac{\beta_{ij}(I_{\bar{V}j}(t) + GW_{\bar{V}j}(t) + r_a A_{\bar{V}j}(t))}{N_{\bar{V}j}} + \sum_{j=1}^4 \frac{\beta_{ij}r_v(I_{Vj}(t) + GW_{Vj}(t) + r_a A_{Vj}(t))}{N_{Vj}} \right] - \gamma E_{Vi}(t) \quad (2.40)$$

$$\frac{dA_{Vi}(t)}{dt} = p\gamma E_{Vi}(t) - \rho_1 A_{Vi}(t) - \alpha_1 A_{Vi}(t) \quad (2.41)$$

$$\frac{dI_{Vi}(t)}{dt} = (1-p)\gamma E_{Vi}(t) + \alpha_1 A_{Vi}(t) - \alpha_2 I_{Vi}(t) - \rho_2 I_{Vi}(t) \quad (2.42)$$

$$\frac{dGW_{Vi}(t)}{dt} = \alpha_2 I_{Vi}(t) - \alpha_3 GW_{Vi}(t) - \rho_3 GW_{Vi}(t) - \mu_{V1i} GW_{Vi}(t) \quad (2.43)$$

$$\frac{dICU_{Vi}(t)}{dt} = \alpha_3 GW_{Vi}(t) - \rho_4 ICU_{Vi}(t) - \mu_{V2i} ICU_{Vi}(t) \quad (2.44)$$

$$\frac{dR_{Vi}(t)}{dt} = \rho_1 A_{Vi}(t) + \rho_2 I_{Vi}(t) + \rho_3 GW_{Vi}(t) + \rho_4 ICU_{Vi}(t) \quad (2.45)$$

$$\frac{dD_{Vi}(t)}{dt} = \mu_{V1i} GW_{Vi}(t) + \mu_{V2i} ICU_{Vi}(t), \quad (2.46)$$

where $i = 1, 2, 3, 4$, represents each age class.

For the stratified vaccinated model, most of the parameters are assumed to be the same between both strata, in particular, parameters related to the periods of infectiousness, hospitalisation and ICU. The parameters which are adjusted because of the vaccination are the relative infectivity of vaccinated individuals, the age specific vaccination and death rates. The values for these parameters are summarised in Table 2.10 below:

The relative infectivity of vaccinated individuals depend on various factors such as the variant, type of vaccination, number of doses and time since vaccination to mention a few (Shinde et al. (2021), Sadoff et al. (2021), Chemaitelly et al. (2021)). The model is run over different values between the ranges in Table 2.10.

Table 2.10: Model Parameters of Vaccination Model

Parameter	Description	Estimate	Reference
r_v	Relative infectivity of vaccinated individuals	0.2 – 0.5	Shinde et al. (2021)
v_1	Vaccination rate in age group 1	1437/14	Estimated
v_2	Vaccination rate in age group 2	2902/14	Estimated
v_3	Vaccination rate in age group 3	1559/14	Estimated
v_4	Vaccination rate in age group 4	627/14	Estimated
μ_{V1i}	Death rate of vaccinated people in GW and age group i	0.1 – 0.3	Sadoff et al. (2021)
μ_{V2i}	Death rate of vaccinated people in ICU and age group i	0.1 – 0.3	Sadoff et al. (2021)

South Africa’s vaccination roll-out gave priority to health care workers, teachers, inmates and different age groups. The initial vaccination rate was estimated per age group with vaccination data from the month of February. According to the CDC, the period until the vaccination is effective is 14 days, so the rate of vaccinated per day was further multiplied by the time until the vaccination is effective before an individual enters the vaccinated compartments. This rate was assumed to be the same and increased per age group once those vaccinations had started.

Results and Discussion

The vaccination model was run over different assumptions about the level of efficacy of the vaccination against infection and the level of virulence of the variant(s) present at the time.

Figure 2.16 is an illustration of the observed values against the different levels of virulence of the variant. It was run under assumptions that the variant is 10%, 20% more and less virulent against the base model. The case that provided the closest results was the assumption of 20% less virulent.

Figure 2.17 is an illustration of the different cases of assumed efficacy of the vaccination. The levels considered were 90%, 80%, 70%, 60% and 50%. The case of 80% efficacy against infection was the closest fit to the observed data. Figure 2.18 was obtained by varying both the level of efficacy of the vaccination and the virulence of the variant at the time. Initially, the model underestimates the number of infections, but it then begins to converge.

The scenario runs indicate the level of variability just introduced by the variant and the efficacy of the vaccination which need to be carefully considered when building and running the model.

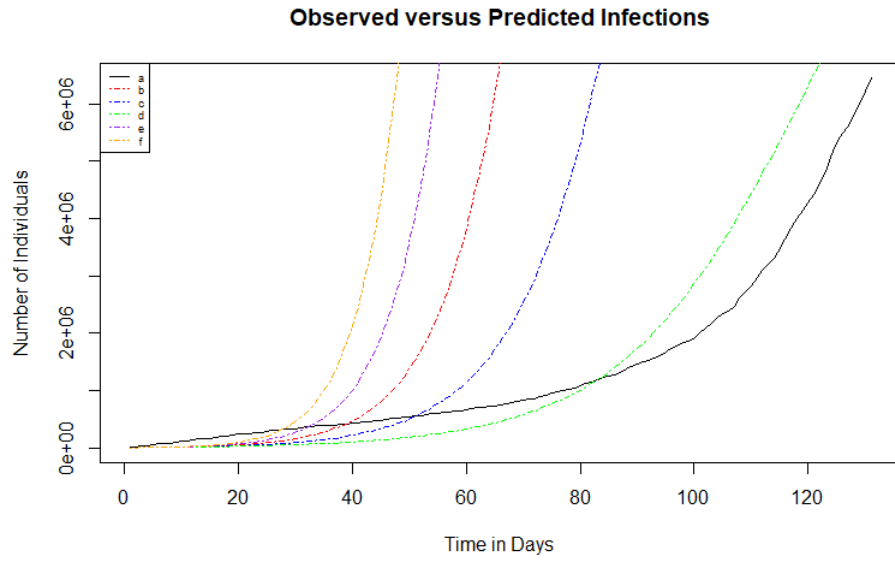


Figure 2.16: *Observed and predicted total infections: a) observed data, b) base model, c) 10% less virulent, d) 20% less virulent, e) 10% more virulent, f) 20% more virulent.*

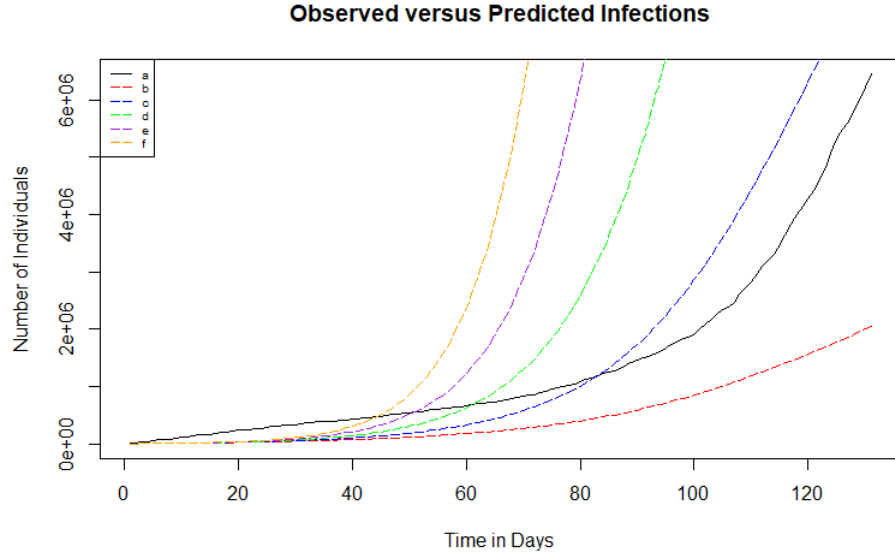


Figure 2.17: *Observed and predicted total infections: a) observed data, b) 90% efficacy, c) 80% efficacy, d) 70% efficacy, e) 60% efficacy, f) 50% efficacy.*

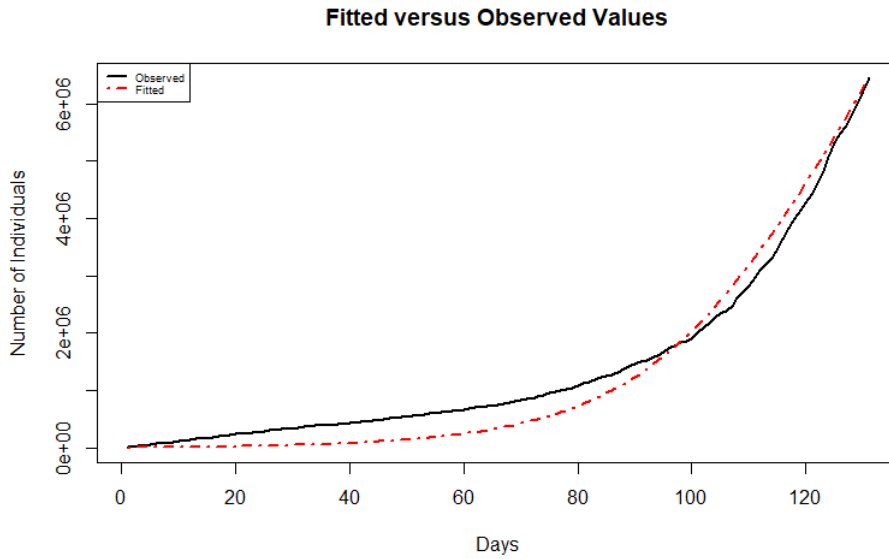


Figure 2.18: *Observed and predicted total infections with 30% reduction in virulence and 63% efficacy of the vaccination.*

This example on the building of the COVID-19 model shows the many elements that need to be considered when building a model for the spread of novel infectious diseases. An age-stratified model was chosen based on the fact that COVID-19 has non-uniform attack rates on different age-groups, this is further motivated by the fact that the vaccination roll-out in South Africa was implemented according to age. Though the model calibrated with the latest data produces a good fit of the observed data, for prediction this can produce diverging results due to the constant changing of the virus and circumstances. To overcome this, the model should be fit over shorter periods and over varying scenarios to observe the possible range of outcomes. The introduction of vaccinations introduced many additional uncertainties regarding the length of immunity, the efficacy levels and the effectiveness against different variants. The model can be improved by considering relaxing some of the assumptions regarding these aspects to create a more comprehensive model accounting for these features.

2.3 The Stochastic SIR Model

Though the deterministic model is useful for the modelling of average behaviour of large populations, the stochastic model is more useful in smaller networks and for assessing the variability of estimates and predictions. The stochastic compartmental model is analogous to the deterministic models with three states: susceptible (S), Infected (I) and recovered (R).

Consider a stochastic SI model, also known as the simple stochastic epidemic model, with a total population of N individuals. In this model, a susceptible individual can become infected when in contact with an infected individual. An infected individual would remain infected and there is no removal by death or recovery. For each $t \geq 0$, Let

- N represents the total number of individuals at any time point and assume that it is constant.
- $S(t)$ represents the random number of susceptible individuals at time t ,
- $I(t)$ represents the random number of infected individuals at time t ,

Assume that $(S, I)(0) = (s_0, i_0)$, where i_0 is the initial number of infected individuals and the population is closed.

$$S(t) + I(t) = N \quad \text{for } t \geq 0 \quad (2.47)$$

Assume that epochs in which individuals change their state are randomly determined in continuous time. It is assumed that $\{(S, I)(t), t \geq 0\}$ is a homogeneous Markov chain in continuous time with state space the non negative integers satisfying equation 2.31.

Since the system is closed, the Markov chain is univariate. It can be written as $\{S(t), t \geq 0\}$ which is a pure death process. The following derivations for the simple stochastic epidemic are obtained from Beichelt (2018).

The transition probabilities during the time interval $(t, t + \Delta t)$ are defined as:

$$P\{S(t + \Delta t) = i - 1 | S(t) = i\} = \beta i(N - i)\Delta t + o(\Delta t) \quad (2.48)$$

$$P\{S(t + \Delta t) = i | S(t) = i\} = 1 - \beta i(N - i)\Delta t + o(\Delta t), \quad (2.49)$$

with $S(t) = 0$ being an absorbing state and $\beta > 0$ is the pairwise infection parameter between the infected and susceptible individuals.

The forward Kolmogorov equations for the state probabilities, $p_i(t) = P(S(t) = i | s(0) = s_0)$, $(0 < i < N - i_0)$ are :

$$\frac{dp_N(t)}{dt} = -\beta(N - i_0)(i_0)p_N(t) \quad (2.50)$$

$$\frac{dp_i(t)}{dt} = -\beta i(N - i)p_i(t) + \beta(i + 1)(N - i - 1)p_{i+1}(t), \quad 1 \leq i \leq N - 1 \quad (2.51)$$

$$\frac{dp_0(t)}{dt} = \beta(N - 1)p_1(t). \quad (2.52)$$

Equations 2.34, 2.35 and 2.36 are very complex and the explicit results may not be obtained easily. To simplify the approach, Laplace transforms are used. Laplace transform $\hat{p}_i(\theta) = \int_0^\infty e^{-\theta t} p_i(t) dt$, $Re(\theta) > 0$, $0 \leq i \leq N$ are applied to equations 2.34, 2.35 and 2.36 giving:

$$\int_0^\infty e^{-\theta t} \frac{dp_N(t)}{dt} dt = \theta \hat{p}_N(\theta) - 1 = -\beta(N - i_0)(i_0) \hat{p}_N(\theta) \quad (2.53)$$

$$\int_0^\infty e^{-\theta t} \frac{dp_i(t)}{dt} dt = \theta \hat{p}_i(\theta) = -\beta i(N - i) \hat{p}_i(\theta) + \beta(i + 1)(N - i - 1) \hat{p}_{i+1}(\theta), \quad 1 \leq i \leq N - 1 \quad (2.54)$$

$$\int_0^\infty e^{-\theta t} \frac{dp_0(t)}{dt} dt = \theta \hat{p}_0(\theta) = \beta(N - 1) \hat{p}_1(\theta). \quad (2.55)$$

Simplifying equations 2.37, 2.38 and 2.39 further, we derive:

$$\hat{p}_N(\theta) = \frac{1}{\theta + \beta((N - i_0)(i_0))} \quad (2.56)$$

$$\hat{p}_i(\theta) = \frac{\beta(i + 1)(N - i - 1)}{\theta + \beta i(N - i)} \hat{p}_{i+1}(\theta) \quad (2.57)$$

$$\hat{p}_0(\theta) = \frac{\beta(N - 1)}{\theta} \hat{p}_1(\theta). \quad (2.58)$$

Solving equations 2.40, 2.41 and 2.42 recursively, we obtain

$$\hat{p}_i(\theta) = \prod_{j=1}^{N-i_0} \frac{\beta(j + 1)(N - j - 1)}{\theta + \beta j(N - j)}, \quad 0 \leq j \leq N - 1. \quad (2.59)$$

An alternate method of obtaining probabilities without Laplace transforms is with the use of generating functions.

Define the generating function for $\{S(t), t \geq 0\}$ as

$$\varphi(z, t) = E[z^{S(t)}] = \sum_{j=0}^{N-i_0} P(S(t) = j | S(0)) z^j, \quad (2.60)$$

where $0 \leq S(t) \leq N - i_0$.

Since $S(t)$ is a pure death process, the results in the appendix can be used to simplify the result.

The General Epidemic Model

The general epidemic model or stochastic SIR model is based on more realistic assumptions. In this model, a susceptible individual can become infected when in contact with an infected individual and the infected individuals can recover from the disease after a recovery period. For each $t \geq 0$, Let

- N represents the total number of individuals at any time point and assume that it is constant,

-
- $S(t)$ represents the random number of susceptible individuals at time t ,
 - $I(t)$ represents the random number of infected individuals at time t ,
 - $R(t)$ represents the random number of recovered individuals at time t ,

Assume that $(S, I, R)(0) = (s_0, i_0, 0)$, where i_0 is the initial number of infected individuals and

$$S(t) + I(t) + R(t) = N \quad \text{for } t \geq 0 \quad (2.61)$$

We assume that $\{(S, I)(t) | t \geq 0\}$ is a bivariate Markov process.

The derivations that follow are from Daley and Gani (2001).

Let i and j , $0 \leq i + j \leq N$, $0 \leq i \leq s_0$, represent the number of susceptible and infected individuals respectively. Assume there is homogeneous mixing during the time interval $(t, t + \Delta t)$ such that infections occur at rate $\beta ij \Delta t$ and removals at rate $\gamma j \Delta t$. The conditional transition probabilities are denoted as:

$$P\{(S, I)(t + \Delta t) = (i - 1, j + 1) | (S, I)(t) = (i, j)\} = \beta ij \Delta t + o(\Delta t) \quad (2.62)$$

$$P\{(S, I)(t + \Delta t) = (i, j - 1) | (S, I)(t) = (i, j)\} = \gamma j \Delta t + o(\Delta t) \quad (2.63)$$

$$P\{(S, I)(t + \Delta t) = (i, j) | (S, I)(t) = (i, j)\} = 1 - \beta ij \Delta t - \gamma j \Delta t + o(\Delta t), \quad (2.64)$$

where β is the pairwise infection parameter, which is assumed to be constant and γ is the recovery rate.

The state probabilities are denoted as

$$p_{ij}(t) = P\{(S, I)(t) = (i, j) | (S, I)(0) = (s_0, i_0)\} \quad (2.65)$$

Therefore the Kolmogorov forward equations are:

$$\frac{dp_{s_0, i_0}(t)}{dt} = -i_0(\beta s_0 + \gamma)p_{s_0, i_0}, \quad (2.66)$$

$$\frac{dp_{i, j}(t)}{dt} = \beta(i + 1)(j - 1)p_{i+1, j-1} - j(\beta i + \gamma)p_{i, j} + \gamma(j + 1)p_{i, j+1}, \quad 0 \leq i + j \leq N, 0 \leq i \leq s_0. \quad (2.67)$$

With initial conditions:

$$p_{s_0, i_0}(0) = 1 \quad \text{and} \quad p_{i, j}(0) = 0 \quad \text{otherwise} \quad (2.68)$$

Define the PGF of the bivariate Markov Process, $\{(\mathbf{S}, \mathbf{I})(t)\}$, by $\varphi(z, w, t)$, where

$$\begin{aligned}\varphi(z, w, t) &= \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} p_{i,j}(t) z^i w^j \\ &= E[z^{\mathbf{S}(t)} w^{\mathbf{I}(t)} | (\mathbf{S}, \mathbf{I})(0) = (s_0, i_0)].\end{aligned}\quad (2.69)$$

The partial derivatives of the PGF are

$$\frac{\partial \varphi(z, w, t)}{\partial z} = \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} i p_{i,j}(t) z^{i-1} w^j \quad (2.70)$$

$$\frac{\partial \varphi(z, w, t)}{\partial w} = \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} j p_{i,j}(t) z^i w^{j-1}. \quad (2.71)$$

Multiplying equations 2.50 and 2.51 by z^i and w^j and summing over i and j gives

$$\begin{aligned}\frac{\partial \varphi(z, w, t)}{\partial t} &= \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} \beta(i+1)(j-1) p_{i+1,j-1} z^i w^j - \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} j(\beta i + \gamma) p_{i,j} z^i w^j \\ &\quad + \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} \gamma(j+1) p_{i,j+1} z^i w^j.\end{aligned}\quad (2.72)$$

Substituting equations 2.54 and 2.55 and simplifying yields

$$\frac{\partial \varphi(z, w, t)}{\partial t} = w(w\beta - z) \frac{\partial^2 \varphi}{\partial z \partial w} + \gamma(1 - w) \frac{\partial \varphi}{\partial w}. \quad (2.73)$$

This can be solved numerically or with the use of Laplace transforms.

Relationship between the stochastic and deterministic SIR model

Assume that $\{(\mathbf{S}, \mathbf{I})(t) | t \geq 0\}$ is a Markov process with transition probabilities as given by equations 2.39 and 2.40.

Let $\Delta \mathbf{S} = \mathbf{S}(t + \Delta t) - \mathbf{S}(t)$ and $\Delta \mathbf{I} = \mathbf{I}(t + \Delta t) - \mathbf{I}(t)$ denote the increments of the number of susceptible and infective individuals respectively over a period of Δt .

Taking the expected value of the increments, we get

$$E[\Delta \mathbf{S}] = -\beta \mathbf{S}(t) \mathbf{I}(t) \Delta t \quad (2.74)$$

$$E[\Delta \mathbf{I}] = (\beta \mathbf{S}(t) \mathbf{I}(t) - \gamma \mathbf{I}(t)) \Delta t. \quad (2.75)$$

Therefore the increments can be written as :

$$\Delta \mathbf{S} = -\beta \mathbf{S}(t) \mathbf{I}(t) \Delta t + \Delta Z_1 \quad (2.76)$$

$$\Delta \mathbf{I} = (\beta \mathbf{S}(t) \mathbf{I}(t) - \gamma \mathbf{I}(t)) \Delta t - \Delta Z_1 + \Delta Z_2, \quad (2.77)$$

where ΔZ_1 and ΔZ_2 are conditionally centered Poisson increments with mean 0 and conditional variance $\beta S(t)I(t)\Delta t$ and $-\gamma I(t)\Delta t$. Dropping Z_1 and Z_2 terms and letting Δt tend to zero, we get:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) \quad (2.78)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t), \quad (2.79)$$

$$(2.80)$$

which corresponds to the ordinary differential equations of the deterministic SIR model.

Mathematical Modelling HIV Transmission using Generating Functions

HIV is an infectious virus that attacks the immune system, making infected people vulnerable to other diseases. If left untreated, it develops to the last stage of HIV also known as AIDS.

HIV is spread by contact with bodily fluids. AIDS was first identified in South Africa in 1983 (Karim and Karim (2010)) and since then it has been one of the biggest contributors to mortality. According to Organization et al. (2016), 70% of people living with HIV/AIDS are in middle-income countries. The main goal of the Global Health Sector Strategy on HIV/AIDS is to see the end of the AIDS epidemic as a public health threat by 2030.

In this example, the application of generating functions are used for the simplification of the analysis of stochastic SIR models is studied in context of HIV transmission.

A basic model of HIV transmission will be developed into a stochastic SIRS model. The reason that we use a stochastic model is stated from Waema and Olowofeso (2015):

- Biological variables are subject to random fluctuations, making them stochastic by definition.
- Models that are stochastic have more detail than models that are deterministic.
- The average of stochastic models yield deterministic models.

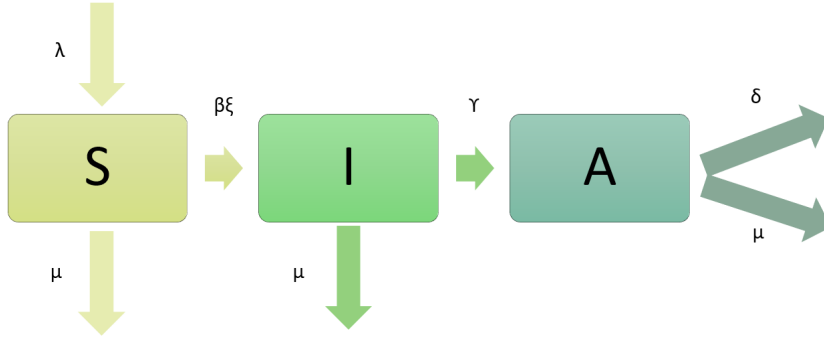


Figure 2.19: *Illustration of the transition between states.*

Let $S(t)$, $I(t)$, $A(t)$ represent the number of susceptible, HIV infected and AIDS infected individuals at time t respectively. Let $N(t) = S(t) + I(t) + A(t)$ represent the total number of individuals at time t . The parameters and notation are given in Table 2.11:

Table 2.11: Model Parameters of Combined Model

Parameter	Description
μ	Natural death rate unrelated to HIV/AIDS
λ	Birth rate
δ	Disease related death rate
β	Rate of sexual contacts
ξ	Probability of infection
γ	Progression to AIDS rates

The model is defined by the following transition probabilities:

$$P\{(S, I)(t + \Delta t) = (i - 1, j) | (S, I)(t) = (i, j)\} = \mu i \Delta t + o(\Delta t) \quad (2.81)$$

$$P\{(S, I)(t + \Delta t) = (i + 1, j) | (S, I)(t) = (i, j)\} = \lambda i \Delta t + o(\Delta t) \quad (2.82)$$

$$P\{(S, I)(t + \Delta t) = (i - 1, j + 1) | (S, I)(t) = (i, j)\} = \beta \xi i \Delta t + o(\Delta t) \quad (2.83)$$

$$P\{(S, I)(t + \Delta t) = (i, j - 1) | (S, I)(t) = (i, j)\} = \gamma j \Delta t + \mu j \Delta t + o(\Delta t) \quad (2.84)$$

$$P\{(S, I)(t + \Delta t) = (i, j) | (S, I)(t) = (i, j)\} = 1 - \lambda i \Delta t - \beta \xi i \Delta t - \gamma j \Delta t - \mu j \Delta t - \mu i \Delta t - o(\Delta t). \quad (2.86)$$

Define the state probabilities as:

$$p_{ij}(t) = P\{(S, I)(t) = (i, j) | (S, I)(0) = (s_0, i_0)\} \quad (2.87)$$

Therefore the Kolmogorov forward equations are:

$$\frac{dp_{s_0, i_0}(t)}{dt} = -\lambda s_0 - \beta \xi s_0 - \mu i_0 \quad (2.88)$$

$$\begin{aligned} \frac{dp_{i,j}(t)}{dt} = & \lambda(i-1)p_{i-1,j} + \beta \xi(i+1)p_{i+1,j-1} + \mu(i+1)p_{i+1,j} + \gamma(j+1)p_{i,j+1} \\ & + \mu(j+1)p_{i,j+1} - (\lambda i + \beta \xi i + \gamma j + \mu j + \mu i)p_{i,j}. \end{aligned} \quad (2.89)$$

with initial conditions:

$$p_{s_0, i_0}(0) = 1 \quad \text{and} \quad p_{i,j}(0) = 0 \quad \text{otherwise} \quad (2.90)$$

Solution using PGF method

Denote the PGF of the bivariate Markov Process, $\{(\mathbf{S}, \mathbf{l})(t)\}$, as $\varphi(z, w, t)$, where

$$\begin{aligned} \varphi(z, w, t) &= \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} p_{i,j}(t) z^i w^j \\ &= E[z^{\mathbf{S}(t)} w^{\mathbf{l}(t)} | (\mathbf{S}, \mathbf{l})(0) = (s_0, i_0)]. \end{aligned} \quad (2.91)$$

The partial derivatives of the PGF is given by:

$$\frac{\partial \varphi(z, w, t)}{\partial z} = \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} i p_{i,j}(t) z^{i-1} w^j \quad (2.92)$$

$$\frac{\partial \varphi(z, w, t)}{\partial w} = \sum_{0 \leq i \leq s_0} \sum_{0 \leq j \leq i_0 - i} j p_{i,j}(t) z^i w^{j-1}. \quad (2.93)$$

Multiplying equations 2.72 and 2.73 by z^i and w^j and summing over i and j gives

$$\begin{aligned} \frac{\partial \varphi(z, w, t)}{\partial t} = & \sum \sum \lambda(i-1)p_{i-1,j} z^i w^j + \sum \sum \beta \xi(i+1)p_{i+1,j-1} z^i w^j \\ & + \sum \sum \gamma(j+1)p_{i,j+1} z^i w^j + \sum \sum \mu(i+1)p_{i+1,j} z^i w^j \\ & + \sum \sum \mu(j+1)p_{i,j+1} z^i w^j - \sum \sum (\lambda i + \beta \xi i + \gamma j + \mu j + \mu i)p_{i,j} z^i w^j. \end{aligned}$$

Substituting equations 2.76 and 2.77 and simplifying yields

$$\frac{\partial \varphi(z, w, t)}{\partial t} = \lambda z^2 \frac{\partial \varphi}{\partial z} + w \beta \xi \frac{\partial \varphi}{\partial z} + \mu \frac{\partial \varphi}{\partial z} + (\gamma + \mu) \frac{\partial \varphi}{\partial w} - z(\lambda + \beta \xi) \frac{\partial \varphi}{\partial z} - (\gamma + \mu) w \frac{\partial \varphi}{\partial w}. \quad (2.94)$$

Further simplification gives

$$\frac{\partial \varphi(z, w, t)}{\partial t} = ((\lambda z - \mu)(z - 1) + \beta \xi(w - z)) \frac{\partial \varphi}{\partial z} + ((\gamma + \mu)(1 - w)) \frac{\partial \varphi}{\partial w}. \quad (2.95)$$

Let $((\lambda z - \mu)(z - 1) + \beta\xi(w - z)) = f(w, z)$ and let $(\gamma + \mu)(1 - w) = g(w)$. Then equation 2.79 can be written as

$$f(w, z) \frac{\partial \varphi}{\partial z} + g(w) \frac{\partial \varphi}{\partial w} - \frac{\partial \varphi}{\partial t} = 0. \quad (2.96)$$

This is a quasi-linear first order partial differential equation. The method of characteristic functions will be used to solve further.

We define the characteristic equations as

$$\frac{d\varphi}{ds} = 0 \quad (2.97)$$

$$\frac{dt}{ds} = -2 \quad (2.98)$$

$$\frac{dw}{ds} = g(w) \quad (2.99)$$

$$\frac{dz}{ds} = f(w, z). \quad (2.100)$$

Solving the system, from (2.80) we get

$$d\varphi = 0 \quad (2.101)$$

$$\therefore \varphi = c_1. \quad (2.102)$$

From (2.81),

$$\frac{dt}{ds} = -1 \quad (2.103)$$

$$dt = -ds \quad (2.104)$$

$$t = s + c_2. \quad (2.105)$$

From (2.82)

$$\frac{dw}{ds} = g(w) \quad (2.106)$$

$$\frac{dw}{g(w)} = ds, \quad (2.107)$$

which gives

$$\int \frac{dw}{g(w)} = s + c_3. \quad (2.108)$$

Evaluating the integral gives:

$$\frac{\ln |1 - w|}{\gamma + \mu} = s + c_3. \quad (2.109)$$

Making w the subject, we get:

$$\begin{aligned}\frac{\ln |1-w|}{\gamma+\mu} &= s + c_3 \\ \ln |1-w| &= (\gamma+\mu)(s+c_3) \\ w &= 1 - e^{(\gamma+\mu)(s+c_3)}.\end{aligned}\tag{2.110}$$

From (2.83),

$$\frac{dz}{ds} = f(w, z) \tag{2.111}$$

$$\frac{dz}{f(w, z)} = ds, \tag{2.112}$$

which gives

$$\int \frac{dz}{f(w, z)} = s + c_4. \tag{2.113}$$

Substituting $f(w, z)$:

$$\int \frac{dz}{((\lambda z - \mu)(z - 1) + \beta \xi(w - z))} = s + c_4.$$

Further substituting (2.84) , we get

$$t + \int \frac{dz}{((\lambda z - \mu)(z - 1) + \beta \xi(1 - e^{(\gamma+\mu)(s+c_3)} - z))} = c_4. \tag{2.114}$$

Further simplification gives,

$$\int \frac{dz}{((\lambda z - \mu)(z - 1) + \beta \xi(1 - e^{(\gamma+\mu)(s+c_3)} - z))} = \int \frac{dz}{\lambda z^2 - z(\lambda + \mu + \beta \xi) + \mu + \beta \xi(1 - e^{(\gamma+\mu)(s+c_3)})} \tag{2.115}$$

$$= \int \frac{dz}{az^2 + bz + c}, \tag{2.116}$$

where $a = \lambda$, $b = -(\lambda + \mu + \beta \xi)$ and $c = \mu + \beta \xi(1 - e^{(\gamma+\mu)(s+c_3)})$, which has solution

$$\int \frac{dz}{az^2 + bz + c} = \begin{cases} \frac{2}{\sqrt{\Delta}} \arctan \frac{2ax+b}{-\sqrt{\Delta}} + c, & \Delta < 0 \\ \frac{1}{\sqrt{\Delta}} \ln \left| \frac{2ax+b-\sqrt{\Delta}}{2ax+b+\sqrt{\Delta}} \right|, & \Delta > 0, \end{cases} \tag{2.117}$$

where $\Delta = b^2 - 4ac$.

The general solution of the PDE in explicit form is

$$\varphi(z, w, t) = F(c_2, c_3, c_4). \quad (2.118)$$

where F is any differentiable function between three variables.

The solution of the above PDE, does not yield a useful result, however, the moments can be found numerically once $F()$ has been defined based on initial conditions.

2.4 Conclusion

This chapter defined and demonstrated various deterministic and stochastic compartmental models which are used in the modelling of infectious diseases.

Under the deterministic compartmental models, the basic SIR model was defined along with the assumptions of the model. This model was illustrated with a simplified COVID-19 example high-lighting the impact of social-distancing. The stratified model was then defined and demonstrated with the building of a COVID-19 model that accounts for vaccinations. As a novel infectious disease, there is much uncertainty that lies in model estimates due to the unknown nature and of the disease and the different mutations. To account for these uncertainties, it is recommended to conduct a sensitivity analysis to understand the impact of changing parameters, to run different case scenarios so that you can account for some of the uncertainty and to predict over shorter intervals and to frequently update the model as new information arises.

The stochastic versions of the compartmental model was then defined and demonstrated. The stochastic compartmental model can be defined as Markov chains leading to a wealth of information from the theory of stochastic processes. The stochastic compartmental model is used to model HIV transmission to illustrate the usefulness of the underlying stochastic theory to solve the stochastic differential equations with the use of generating functions.

This chapter emphasises the multidisciplinary nature of the modelling of infectious diseases and the wide range of techniques that can be used just within the space of one type of model. Deterministic models are often used due to the fact that they are easier to understand and implement, however, to account for the lack of randomness, models will be need to be run over several cases to observe different outcomes. A stochastic model allows the random fluctuation to be incorporated into the model, however these models are more difficult to run and solve and can become computationally expensive.

In this chapter, a vital piece missing in the modelling of infectious disease, is the spatial component. The next chapter will justify the need for a spatial component in the modelling of infectious disease.

Chapter 3

Justification for a spatial model for the spread of infectious diseases: Tuberculosis in South Africa as an example

This chapter investigates the spatial autocorrelation that exists in South Africa with the use of TB data. This in turn justifies the development of a model which includes the spatial component.

SIR models are based on the assumption of homogeneous mixing, which in turn implies population-wide parameters. The use of these parameters limit the analysis of the model output (Chen et al. (2014)). John Snow, the father of epidemiology, demonstrated the need for spatial modelling with the spatial analysis of the cholera cases in the 1800 London epidemic. The spatial modelling of infectious diseases in South Africa is of vital importance due to the spatial distribution of people and facilities. This section justifies the need by analysing the TB infections in SA.

Tuberculosis (TB) is an airborne infectious disease caused by *mycobacterium tuberculosis* and it is spread by sneezing, coughing or speaking. The disease has claimed over one billion lives over the past two centuries (Castillo-Chavez and Song (2004)). Tuberculosis is one of the biggest contributors to death around the world with 80% of the deaths stemming from only 23 countries in East Asia and Africa (Organization (2018)). South Africa, China and India alone account for 40% of the world-wide cases (Houben et al. (2016)). The post 2015-TB strategy aims to achieve a 50% reduction in world-wide incidence and 75% decrease in mortality by 2025 which they hope to further reduce to 95% and 90% respectively by 2035.

When an individual is infected by TB there is a 90% chance that their immune system will be able to effectively contain it. In such cases, an individual develops latent TB. If an individual's immune system fails to contain the in-

fection, an active TB infection develops. An individual with latent TB can live their entire life without it ever developing into an active TB infection. Latent TB infections are more likely to develop into an active infection if the immune system is weakened or compromised. Malnutrition, co-infection and age are some of the factors which weaken the immune system leaving the individual with a higher risk of contracting active TB (Castillo-Chavez and Song (2004)). Treatment for active TB consists of multiple drugs administered over 9 months. Failure to adhere to the drug therapy or incorrect treatment leads to the evolution of the bacteria to drug resistant strains. Expenses are high for treatment especially in developing countries such as South Africa. To achieve the post 2015-TB strategy policies need to be appropriately designed.

A study of 11 mathematical models by Houben et al. (2016) showed that interventions significantly impact the decrease in incidence and mortality. In South Africa, Houben et al. (2016) showed that continuous treatment for those receiving anti-retrovirals, screening at primary healthcare clinics and improving the linkage to care and treatment access are significant factors for South Africa in order to address to achieve its goals. The study demonstrated the importance of country-context and preventative measures to reach global targets. It also highlighted the fact that South Africa's TB program is overwhelmed by the effects of HIV/AIDS. However, the modelling failed to focus on high risk groups such as people living in informal settlements. It also failed to get a reasonable estimate of mixing between the general and high-risk groups. This motivates the need for a spatial model which includes the heterogeneity from the spatial distribution of people. Due to South Africa's apartheid system, the areas in South Africa are spatially divided. There are affluent areas with sufficient resources, sparsely populated by individuals with a high standard of living and there are poorer areas which are densely populated with high levels of unemployment, poverty and HIV/AIDS rates.

Understanding the drivers of spatial heterogeneity in the spread of TB is key to develop targeted health responses that are effective in the control of TB (Shaweno et al. (2018), Dowdy et al. (2012), Shrestha et al. (2021)). Some of the drivers of the spatial heterogeneity include the choice of spatial resolution, the type of the model and new infections rates versus reactivation rates (Shaweno et al. (2018)). Recent transmission drives TB in high-burden settings whereas reactivation drives the spread in low-endemic settings (Shaweno et al. (2018), Shrestha et al. (2021)). Associations between TB and demographic and socio-economic variables have been found. These risk factors tend to also cluster in hyper-endemic hotspots (Shrestha et al. (2021)). Dowdy et al. (2012) has found that hotspots drive the spread of TB and concluded that targeting hotspots are more effective than targeting the general population in preventing the spread of TB. In this chapter the spatial autocorrelation of the TB infection rates are assessed to justify the need for a spatial component.

3.1 Study Area and Methods

3.1.1 South Africa

The number of incidence cases of drug-sensitive TB infections is studied during 2017 for each of the 52 districts of the 9 provinces in which South Africa is partitioned is studied.

Stretching latitudinally from 22° S to 35° S and longitudinally from 17° E to 33° E, South Africa's surface area covers $1\,219\,602\text{ km}^2$ (Figure 3.1).

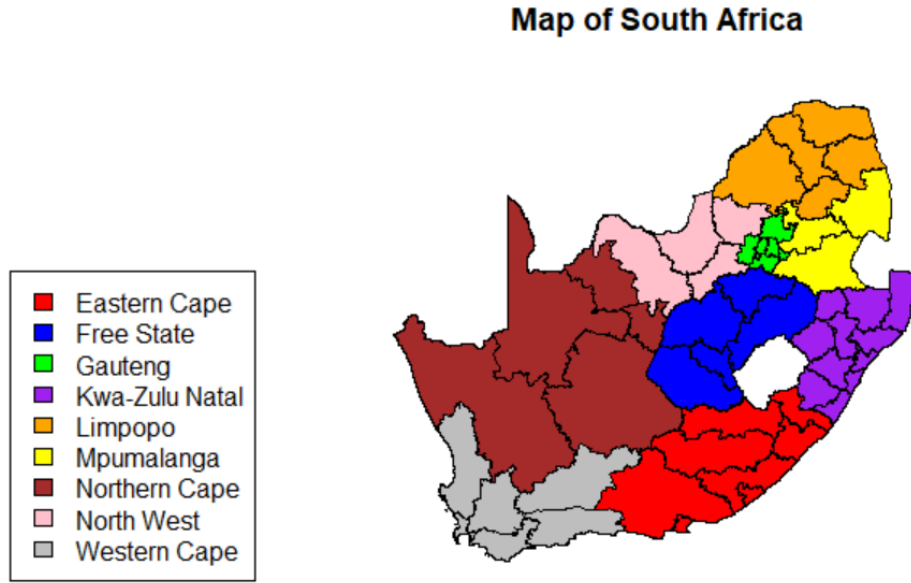


Figure 3.1: *Plot of South Africa Districts and Provinces.*

3.1.2 Measuring Spatial Autocorrelation

Let $Z(s)$ be an attribute observed at location s . The spatial autocorrelation between s_i and s_j is the degree of similarity between them and is measured as the standardised covariance between the locations, $Z(s_i)$ and $Z(s_j)$. Positive autocorrelation implies that the feature is positively associated with the surrounding areas and a negative autocorrelation indicates that the feature is negatively associated with the values of surrounding areas.

Spatial Weight Matrix and Neighbourhood Structure

Spatial autocorrelation depends on the neighbourhood structure of the data, and the measurement of spatial relationships must be defined.

Let n be the number of spatial units. The spatial weight matrix, W , is a $n \times n$ positive symmetric matrix with element w_{ij} at location i and j . The values of the weights, w_{ij} , for each pair of locations are assigned by some pre-set rules

Table 3.1: Summary of spatial weights

Name of weight	w_{ij}
Inverse distance	$\frac{1}{d_{ij}}$
Inverse square distance	$\frac{1}{d_{ij}^2}$
Threshold distance at D	$\mathbb{1}_{0 \leq d_{ij} \leq D}(d_{ij})$
K -nearest neighbour	$\mathbb{1}_{\text{if } j \text{ one of the } k \text{ nearest neighbours}(s_i)}$
Queen adjacency	$\mathbb{1}_{\text{if } i \text{ and } j \text{ share a common vertex or boundary}}$
Rook adjacency	$\mathbb{1}_{\text{if } i \text{ and } j \text{ share a common boundary}}$
Bishop adjacency	$\mathbb{1}_{\text{if } i \text{ and } j \text{ share a common vertex}}$

which defines the spatial relation among locations. By convention $w_{ii} = 0$ for all $i = 1, 2, \dots, n$, with

$$W = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{nn} \end{bmatrix}. \quad (3.1)$$

There are two main ways to measure spatial relationships of regions: distance and contiguity (state of bordering or being in contact). The spatial weight is often chosen based on the data and the way that the neighbours interact. Background knowledge on how the weights decay is used to choose appropriate weights (Lee and Ogburn (2017)). The weight matrix is used to calculate the extent of the autocorrelation, and thus, would affect the result. Lee and Ogburn (2017) observed that the weight matrix affects the power of the test for autocorrelation, but not the validity. Table 3.1 summarises some of the weights that can be used, where $\mathbb{1}$ is the indicator function.

In this thesis spatial autocorrelation will be assessed by Moran's I using a variety of spatial weight matrices.

Moran's I

A common measure used to assess the spatial autocorrelation of a continuous attribute Z , is the Moran's I (Moran (1950)). It calculates the ratio between the product of a variable and its spatial lag with the product of the variable adjusted for the spatial weights used.

$$I = \frac{n}{(n-1)S^2w_{..}} \sum_i \sum_j w_{ij} (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z}), \quad (3.2)$$

where

$$\begin{aligned}\bar{Z} &= \frac{\sum_{i=1}^n Z_i}{n}, \quad \text{and} \\ S^2 &= \frac{1}{n-1} \sum_i^n (Z(s_i) - \bar{Z})^2, \\ w_{..} &= \sum_i \sum_j w_{ij}.\end{aligned}$$

The value usually ranges from -1 , (indicating perfect dispersion) to 1 (indicating perfect correlation). A value near 0 indicates no spatial pattern.

The hypothesis for Moran's I test is:

$$\begin{aligned}H_0: & \text{The spatial pattern is random} \left(E[I] = \frac{-1}{n-1} \right) \\ H_1: & \text{The spatial pattern is correlated} \left(E[I] \neq \frac{-1}{n-1} \right)\end{aligned}$$

The test statistic can be calculated as

$$z(I) = \frac{I - E[I]}{\sqrt{Var(I)}}, \quad (3.3)$$

where $E[I] = \frac{-1}{n-1}$ and $Var(I) = E[I^2] - E[I]^2$.

The exact mean and variance of Moran's I under the assumption of independence of the error terms is derived by Getis and Ord (1992). Sen (1976) found the standardised Moran's I to be asymptotically normal. This asymptotic distribution holds for when there are at least 50 observations.

Global G

An area where high values cluster is known as a hot spot, and a cold spot when low values cluster. The above measures of autocorrelation are not able to indicate the type of correlation. For that reason, we use the Global G (Ord and Getis (1995)). The Global G is able to distinguish between hot and cold spots:

- G is large if high values cluster (hot spot)
- G is low if low values cluster (cold spot)

The observed G-value is calculated as

$$G_i = \frac{\sum_{j=1}^n w_{ij} Z(s_j)}{\sum_{j=1}^n Z(s_j)}, \quad (3.4)$$

where $Z(s_j)$ is the magnitude of variable Z at incident location j over all n and w_{ij} is the spatial weight defining the relationship between s_i and S_j .

The expected value is

$$E[G_i] = \frac{\sum_{j=1}^n w_{ij}}{n}. \quad (3.5)$$

and the variance is

$$var(G_i) = \frac{\sum_{j=1}^n w_{ij}(n - \sum_{j=1}^n w_{ij})}{n - 1} \quad (3.6)$$

G is asymptotically normally distributed (Getis and Ord (1992)) when the underlying distribution of variable Z is normally distributed or when the number of observations become large enough.

Local Indicators of Spatial Autocorrelation(LISA)

Local measures calculate the correlation for each area. Anselin (1995) defined a statistic to be a LISA if:

1. the degree of spatial autocorrelation is measured.
2. it can be calculated for each observation in the data
3. it is proportional to a global measure of spatial autocorrelation.

Local Moran's I

A local Moran's I satisfies the above conditions and can be constructed as one of the n components which compromise the global test.

$$I(s_i) = \frac{n}{(n-1)S^2} (Z(s_i) - \bar{Z}) \sum_j w_{ij} (Z(s_j) - \bar{Z}) \quad (3.7)$$

3.1.3 Software

The calculations, analysis and plots were done using the free-source statistical program, R (R Core Team (2018)). The following packages were used:

- `sp` (Pebesma and Bivand (2005), Bivand et al. (2013))
- `spdep` (Bivand and Wong (2018))
- `tmap` (Tennekes (2018))
- `rgdal` (Bivand et al. (2018))
- `rgeos` (Bivand and Rundel (2018))

3.2 Results

3.2.1 Plot of the Proportion of TB incidence

The number of TB cases per district were used to calculate the proportion of TB per district. This was done by dividing the number of TB cases in each district by the population in each district. The proportion of TB per district was then standardised by subtracting the mean proportion by the standard deviation.

Figure 3.2 is a plot produced using the standardised proportion of TB cases reported. This shows the distribution of TB across the districts in South Africa. There appears to be spatial autocorrelation which will be verified by the Moran's I.

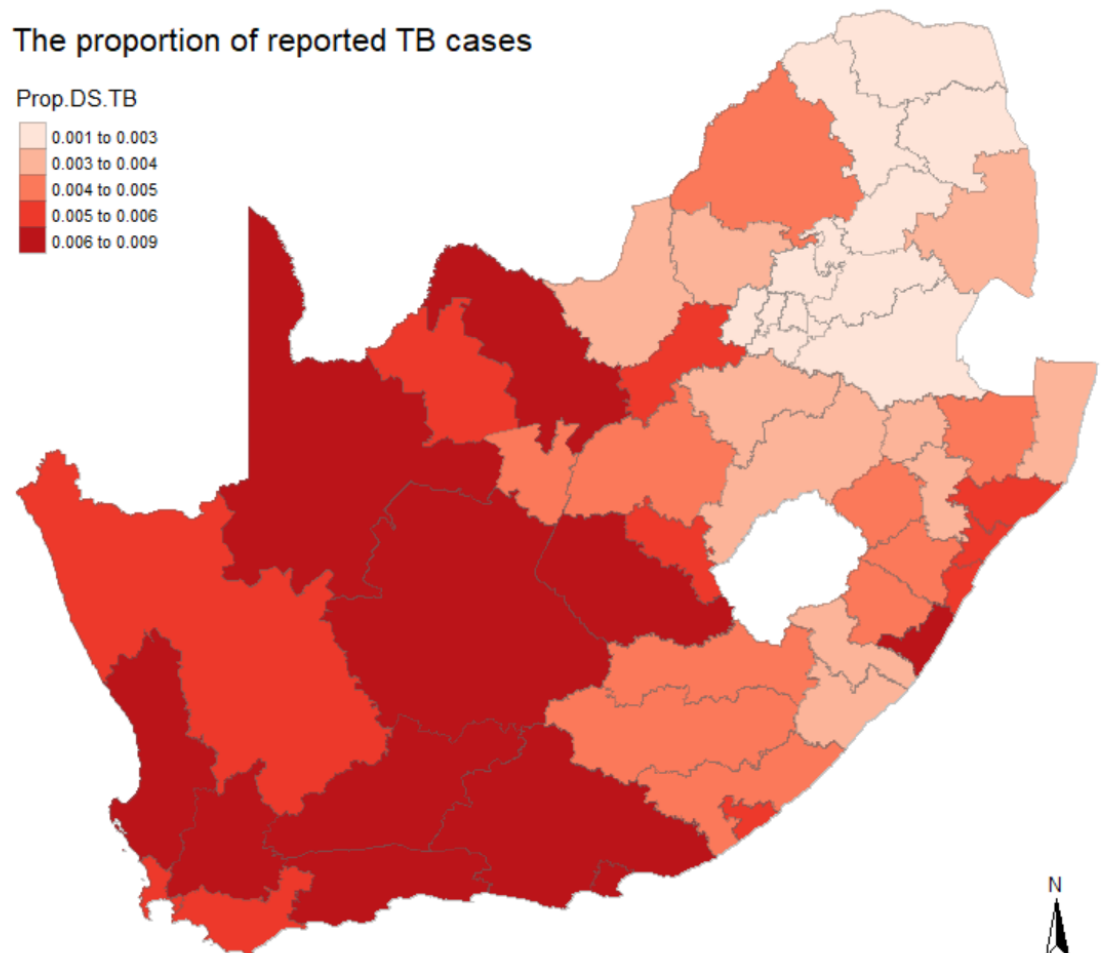


Figure 3.2: *Plot of TB incidence in South Africa.*

3.2.2 Assessing the Spatial Autocorrelation

The spatial autocorrelation was assessed using the Moran's I. Table 3.2 is a summary of the results obtained from the analysis. Moran's I was calculated using a selection of weight matrices as inputs.

Table 3.2: Results of Moran's I test for the different weights.

Weight	Moran's I	Expectation	Variance	P-value
Queen	0.6994	-0.0196	0.0086	$8.96e - 15$
Rook	0.6994	-0.0196	0.0086	$8.96e - 15$
1- nearest neighbour	0.8600	-0.0196	0.0274	$1.069e - 07$
2- nearest neighbours	0.7372	-0.0196	0.0148	$5.227e - 10$
3- nearest neighbours	0.7696	-0.0196	0.0105	$1.297e - 14$
4- nearest neighbours	0.7301	-0.0196	0.0077	$< 2.2e - 16$
5- nearest neighbours	0.7093	-0.0196	0.0061	$< 2.2e - 16$
6- nearest neighbours	0.6950	-0.0196	0.0050	$< 2.2e - 16$
7- nearest neighbours	0.6540	-0.0196	0.0042	$< 2.2e - 16$
8- nearest neighbours	0.6564	-0.0196	0.0035	$< 2.2e - 16$
9- nearest neighbours	0.6422	-0.0196	0.0031	$< 2.2e - 16$
10- nearest neighbours	0.6163	-0.0196	0.0027	$< 2.2e - 16$
248 km threshold distance	0.6043	-0.0196	0.0066	$1.379e - 14$
256 km threshold distance	0.6074	-0.0196	0.0063	$3.374e - 15$
264 km threshold distance	0.5906	-0.0196	0.0061	$6.071e - 15$
272 km threshold distance	0.6026	-0.0196	0.0057	$< 2.2e - 16$
280 km threshold distance	0.5942	-0.0196	0.0053	$< 2.2e - 16$
288 km threshold distance	0.5921	-0.0196	0.0052	$< 2.2e - 16$
300 km threshold distance	0.5793	-0.0196	0.005	$< 2.2e - 16$

From Table 3.2, it can be seen that even though the spatial relationship was defined in different ways, all input weights indicate statistically significant spatial autocorrelation.

The observed Moran's I ranges between 0.5793 – 0.8600.

The Local Moran's I was plotted by selecting some of the weight matrices calculated in Table 3.2. The queen adjacency, 8 nearest neighbours and a weights with a threshold distance of 248 km was chosen to assess the local autocorrelation and the hot and cold spots. The Local Moran's I plots (Figures 3.3, 3.4 and 3.5) also indicate varying results based on the weights chosen as an input.

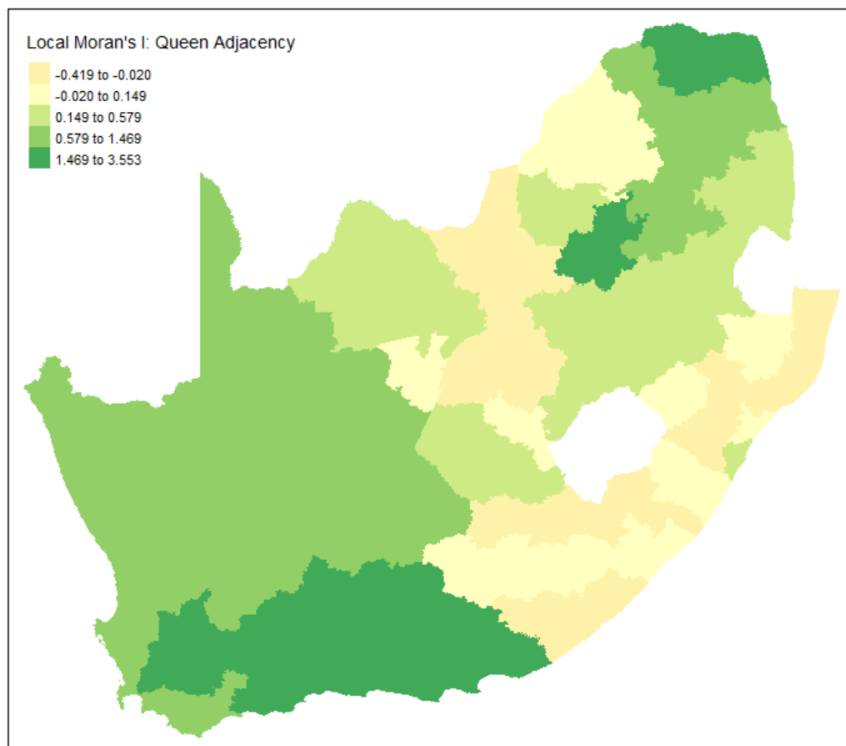


Figure 3.3: *Plot of the local Moran's I statistic using queen adjacency.*

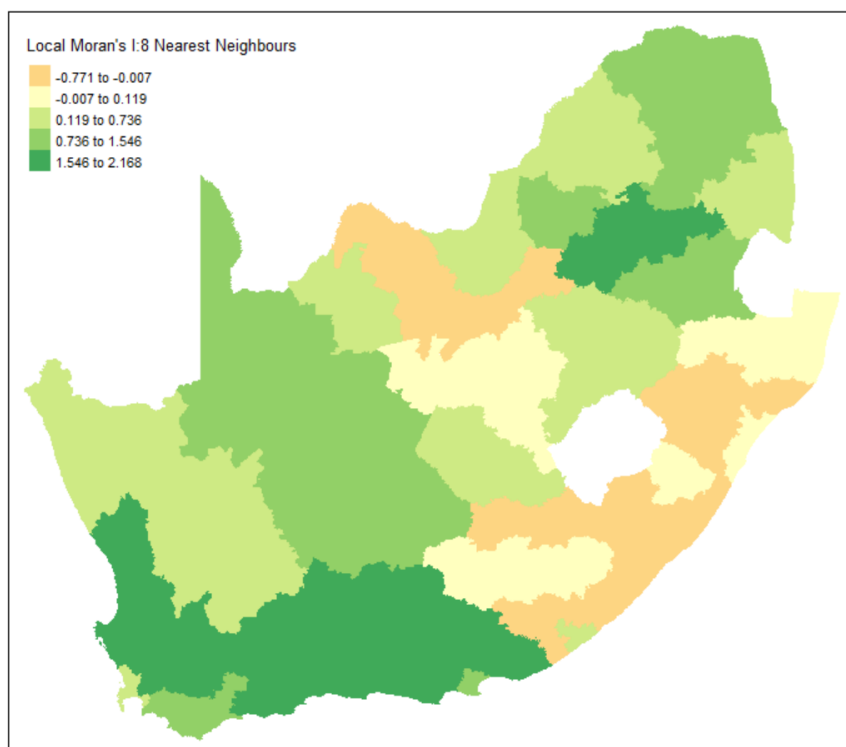


Figure 3.4: *Plot of the local Moran's I statistic using 8 nearest neighbours.*

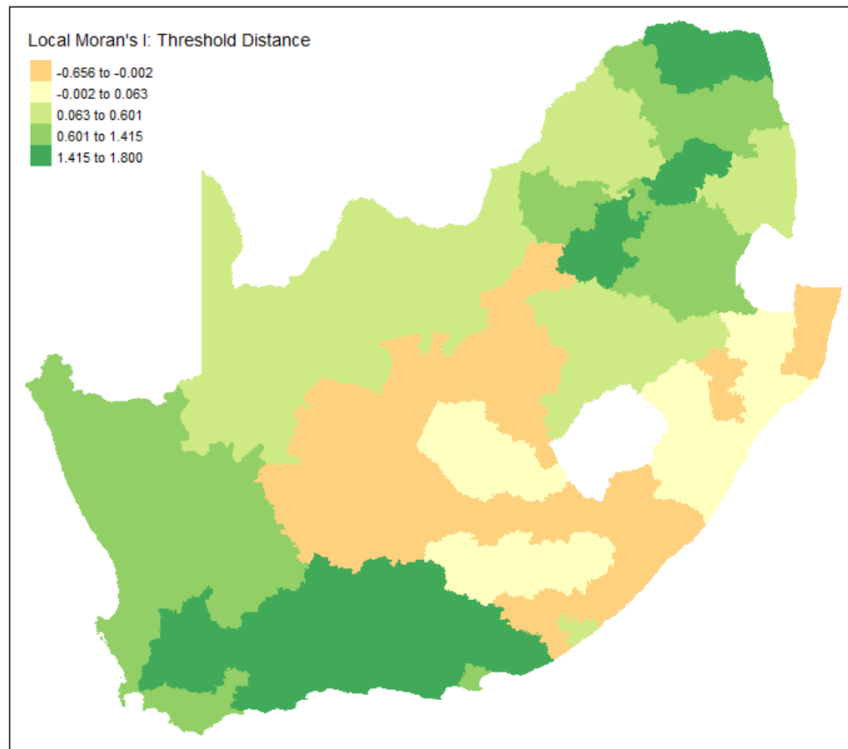


Figure 3.5: *Plot of the local Moran's I statistic using threshold distance of 248 km.*

3.2.3 Hot and Cold Spot Analysis

Figure 3.6 indicated the hot and cold spots for three different weight inputs. Once again, this has been done on a queen adjacency, 8 nearest neighbours and a threshold distance of 248 m. Each weight input results in different hot and cold spots. All plots in Figure 3.6 contain a hot spot at the southern part of South Africa.

3.3 Analysis

3.3.1 Spatial autocorrelation and the effects of the spatial weights matrix

On further examination of the global Moran's I with a variety of weights, it can be seen that the choice of weight has an impact on the value of the Moran's I statistic and ultimately the interpretation of the spatial autocorrelation. All of the Moran's I statistics, as given in Table 3.2 are positive, indicating a positive spatial autocorrelation of the proportion of TB infections. The range of the values (0.5793–0.8600) and the strength of the spatial relationship varies from moderate to strong. Though the Moran's I differs, the corresponding p-values all indicate a significant spatial relationship between the areas.

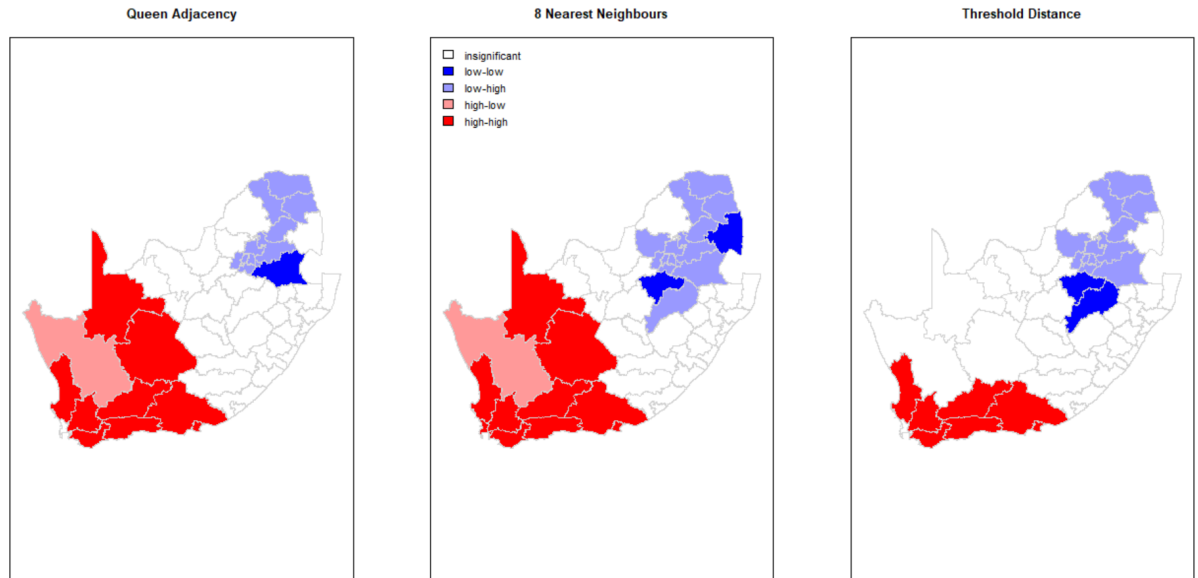


Figure 3.6: *Plot of hot and cold spots using different weight inputs. The white indicates no significant relationship, the red is a hot spot, which is a clustering of high values, the dark blue is a cold spot, which is clustering of low values and the light blue is a clustering of both high and low values.*

The local Moran's I, as plotted by Figures 3.3, 3.4, and 3.5 also indicate varying results based on the type of weight chosen. When a weight matrix defined by queen adjacency is used (Figure 3.3), the south western and north eastern tip of the map display positive spatial autocorrelation whilst the other regions display negative spatial autocorrelation. However, the negative spatial autocorrelation is not as strong as the positive correlation.

When the 8 nearest neighbour is used to define the spatial weight matrix, the resulting plot (Figure 3.4) indicate more positively correlated regions with very few negatively related regions. The plot produced using the threshold distance (Figure 3.5) is very similar to the 8 nearest neighbour plot (Figure 3.4), just with one or two more regions with negative spatial autocorrelation.

Once again, the choice of the spatial matrix affects the interpretation of the results.

The hot and cold spot analysis (Figure 3.6), echo the sentiment of the results of the Moran's I. The hot and cold spots, differ based on the choice of weights, however they all contain a hot spot at the southern tip of South Africa. The cold spots are different from each weight definition.

3.3.2 Proportion of TB cases

The plot of the proportion of TB cases reported per district (Figure 3.2) reveals high levels of infection in the southern and western areas of South Africa.

The northern region of South Africa displays lower levels of infection.

The Moran's I statistics calculated for the range of weights all indicate a significant spatial autocorrelation. The Moran's I indicates a positive relationship between spatial units. This makes intuitive sense, since people with the disease can only affect people around them. This justifies the need for a model which takes into consideration the spatial structure of the proportion of TB infected individuals.

Though the spatial correlation is evident, regardless of the weight, it has become clear that the interpretation of the spatial clustering is highly dependent on the type of weights. This is an aspect which needs to be dealt with extensively when building a spatio-stochastic model for the spread of infectious diseases.

3.4 Conclusion

Many of the areas in the Eastern Cape are rural and sparsely populated. The high proportion of TB cases can be due to the lack of resources in those areas or due to the quality of data recorded. Population estimates of sparsely populated rural areas might not be as accurate as their urban counterparts. A model which accounts for the bias in the data should be considered to more accurately estimate the autocorrelation which exists in a region.

Another factor to consider is the scale which the TB cases were reported. The district level might be too big and a smaller level which reports per ward should be considered to evaluate spatial autocorrelation at a smaller scale. Smaller scales are favourable as they can be aggregated to derive information about a bigger region. A downfall of using a smaller scale is that it is more costly with respect to computing. The spatial weight matrix for an area with split into N regions has dimensions $N \times N$. The smaller the division of the study area, the more computationally intensive the analysis will be.

Another aspect which this study could be improved is with the addition of mobility or connectivity between individuals. This type of data will give insight on the spreading patterns of TB. Other data sources that can be included is data on the socio-economic circumstances of individuals to see if there is a relationship between the hot and cold spots.

The last factor which might affect results is amount of land which is habitable. Not all land in South Africa is habitable due to the terrain. A plot which overlays the geographical make-up such as the elevation and terrain can also provide useful insight to the observed patterns.

This chapter has clearly indicated the need to account for spatial autocorrelation in a model of the spread of infectious diseases. The following chapters

investigate the most effective way to incorporate the spatial component into a compartmental model (or SIR model). There are various methods currently used to account for spatial structure such as network models, patch models and distance transmission models. The following chapter looks at adding a stochastic spatial component so that the model can better represent the spread of the infectious diseases.

Chapter 4

Discrete-Time Stochastic Compartmental Models

This chapter begins with a brief introduction into binomial chain models, such as the Reed-Frost and Greenwood models which are typically discrete-time stochastic process to model the spread of infectious diseases. The Tuckwell and Williams formation is then defined. This model is used as the basis of the derivation of the spatial model with its variations. This chapter is the theoretical foundation of the developed spatio-stochastic model.

In this thesis, the focus is on the mathematical state-space models; specifically, one of the most widely used models, the compartmental or SIRS model. This model has a number of individuals in each phase of the disease. It includes the susceptible population (S), the infected population (I), and the recovered population (R). People move between the compartments S, I and R at certain rates. This basic model can be modified in a variety of ways to account for a wide variety of diseases (Silal et al. (2016)). This chapter is based on the paper Manjoo-Docrat (2022).

Even though these models are versatile and include various key aspects of the disease, many of them retain the assumption of population-wide parameters, which in turn limits the analysis of these models (Chen et al. (2014)). To make the models a more realistic representation, a version which includes the spatial component is more appropriate. The heterogeneity caused by the segregation of populations is one of the main sources of heterogeneity in epidemic modelling (Meakin and Keeling (2018)). The spatial interaction plays a crucial role in the modelling of the spread of an infectious disease. As seen from chapter 3, it is important to consider the spatial relationship defined between spatial areas as it affects the models and analysis. There are many ways to include spatial heterogeneity in an infectious disease transmission model such as gravity models, patch models, distance transmission models, multigroup models as well as network models (Riley (2007)). Eggo et al. (2011) uses a gravity model to analyse the spatial variation and spread of influenza in 1918. Another model which accounts for the spatial properties describes the spread of infec-

tious disease in spatially-structured populations by reaction diffusion equations (Allen et al. (2008)). Wu et al. (2017) make use of a reaction-diffusion model which had different random and local dispersion rates. The use of both the gravity model and reaction-diffusion models provided additional insight on the spread of diseases spatially. Models that include interactions between different sub-populations are known as a meta-population models (Mishra et al. (2010)) and it is one of the ways that spatial dependencies can be included. The meta-population model used by Ming et al. (2016), Keeling and Gilligan (2000), Angulo et al. (2013) and Bisin and Moro (2020) include the spatial heterogeneity by creating sub-populations. These models provide prediction accuracy with additional insights which can be useful to authorities for planning and prevention. Ming et al. (2016) use a stochastic model to study the spread of heterogeneous populations from spatio-temporal data. The parameter associated with disease spread is replaced with a single parameter that can be used for the analysis of the spatial heterogeneity. Tuckwell and Toubiana (2007) develop a spatial model with parameters describing the within-host dynamics which are randomly generated throughout the population. The spatial location is generated in a rectangle and interaction between two individuals follow a Poisson distribution with a mean that is inversely related to the distance between the two individuals.

Although the compartmental model has both deterministic and stochastic formulations, the focus of this paper will be on the stochastic models. With the use of regression models, time-series, and generalised linear functions, there are several ways to stochastically model the spread of infectious diseases. Stochastic models, though more complex than the deterministic counterparts, are more realistic and asymptotically tend to the deterministic models (Dietz (1967), Mollison (1995), Daley and Gani (2001)). Some other desirable aspects that can be obtained from stochastic models is the ability to calculate the probability of a disease outbreak, the stationary probability distribution of the number of infected, the final size distribution and the expected duration of an epidemic (Allen (2010)). These properties are useful for better planning of preventative measures and the efficient allocation of resources. Using stochastic processes is another method which one can apply to include stochasticity. Branching processes, birth and death processes, competition processes and finite Markov chains are some of the stochastic processes which are useful in modelling the spread of infectious diseases (Dietz (1967), Allen (2010)). Stochastic compartmental models are based on these stochastic processes. Compartmental stochastic models are often used and the binomial chain models form the basis of such models.

The model proposed in this section is an extension of the stochastic SIR model developed in Tuckwell and Williams (2007). The model has been developed and extended to include a spatial component to account for the spatial heterogeneity that exists in infectious disease transmission. The spatial interaction is included in the number of contacts that an individual encounters which is

an intuitive method of including the spatial heterogeneity when comparing to other models. The model is demonstrated with three cases, one with only local transmission among spatial units, a model with interaction between spatial units, but no migration and finally a model with both migration and interaction between spatial units. This model allows for flexibility when considering different spatial scenarios which result from behavioural changes in the population which is a key component when building models for the analysis of the spread of infectious diseases. The behavioural aspects such as lockdowns and social distancing are important to include in the models diseases transmission (Toxvaerd (2020), Bisin and Moro (2020), Engle et al. (2020)). The spatio-stochastic model is compared with other discrete-time and space models and scenarios with the use of a simulation study which demonstrates the uses and range of the model.

A variety of discrete time and space models to analyse the spread of infectious diseases have been developed over the years. Discrete time models are appropriate when the infectious period is short relative to the longer latent period. New infections are seen to occur in generations which are separated by the latent period as a discrete time unit.

4.1 Binomial Chain Models

One of the earliest discrete time stochastic model was developed by Reed and Frost in 1928. The model was a binomial chain model with a probability of infection which was assumed to be geometric. This model is reliable for analysing close contact transmission.

In 1931, Greenwood developed another binomial chain model similar to the Reed-Frost model but with a constant probability of infection. These models are examples of Markov models since they possess the ‘memoryless’ property.

Extensions and adaptations of some of the simplest models have been developed leading to a wide-array of options for modelling the spread of infectious diseases. The reformulation and definitions for the binomial chain models are based on the monograph Daley and Gani (2001).

4.1.1 Reed-Frost Model

The Reed- Frost Model was first developed in 1928. Reed and Frost developed chain binomial models of epidemic spread for the use of teaching a course in biostatistics and epidemiology at Johns Hopkins University. This model is an example of a Markov model since the transition probabilities in a generation only depend on the previous generation.

Let X_t represent the number of susceptible individuals at time t and let Y_t represent the number of infective individuals at time t , where $t = 0, 1, 2, \dots$. Let p denote the probability of contact between any infective and any susceptible individual at time t . Let β be the probability of a contact resulting in infection.

Let α denote the probability that there is no infection which is then calculated as follows:

$$\begin{aligned}\alpha &= P(\text{no infection given no contact}) + P(\text{no infection given contact}) \\ \alpha &= 1 - p + p(1 - \beta)\end{aligned}$$

A susceptible individual at time t , will become infected at time $t + 1$, if the individual was infected by at least one of the infected. The probability of remaining susceptible when there are Y_t infected individuals is thus α^{Y_t} . This probability is independent for each of the X_t individuals. Thus the distribution of $X_t \sim \text{bin}(x_{t-1}, \alpha^{y_t})$. The Y_t infected individuals behave independently and the conditional probabilities are defined as follows:

$$\begin{aligned}P(Y_{t+1} = y_{t+1} | X_0 = x_0, Y_0 = y_0, \dots, X_t = x_t, Y_t = y_t) &= P(Y_{t+1} = y_{t+1} | X_t = x_t, Y_t = y_t) \\ &= \binom{x_t}{y_{t+1}} (1 - \alpha^{y_t})^{y_{t+1}} (\alpha^{y_t})^{x_t - y_{t+1}},\end{aligned}\tag{4.1}$$

where $X_{t+1} = X_t - Y_{t+1}$.

Assume the following initial conditions:

$$X_0 = n \quad \text{and} \quad Y_0 = m.\tag{4.2}$$

The complete probability of the chain $y_1, y_2, \dots, y_{t+1} = 0$ is obtained by conditioning sequentially and using the Markov property of the chain.

$$\begin{aligned}P(Y_1 = y_1, \dots, Y_k = y_k, y_{t+1} = 0 | X_0 = n, Y_0 = m) \\ &= P(Y_1 = y_1 | X_0 = n, Y_0 = m) \times \dots \times P(Y_{t+1} = 0 | X_t = x_t, Y_t = y_t) \\ &= \binom{n}{y_1} (1 - \alpha^m)^{y_1} (\alpha^m)^{n - y_1} \times \dots \times \binom{x_k}{0} (1 - \alpha^{y_t})^0 (\alpha^{y_t})^{x_t}.\end{aligned}\tag{4.3}$$

50 000 simulations of the Reed-Frost model as defined above produced Figures 4.1 and 4.2. The simulations were run with the same initial conditions of 10 infected and 2000 susceptible individuals. The probability of infection was varied with 4.1 being a more contagious disease than the one simulated in 4.2. With a low probability of infection, the epidemic last almost a month and just under a fifth of the population is infected at the peak. A higher probability of infection results in more than half the population infected in just a couple of days.

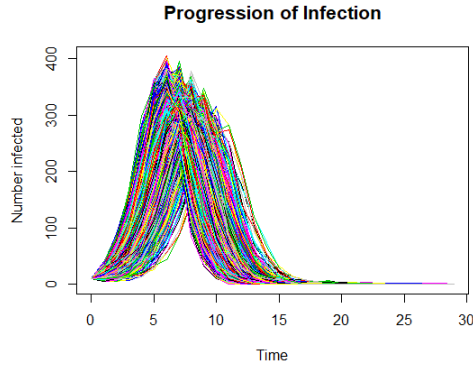


Figure 4.1: *Simulations of a Reed-Frost model with $p = 0.001$, $I_0 = 10$ and $S_0 = 2000$.*

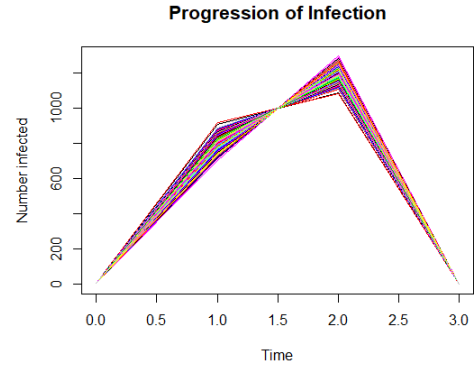


Figure 4.2: *Simulations of a Reed-Frost model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.*

Final Size of the Epidemic

The final size of an epidemic is defined by the total number of infected individuals. Let $z = \sum_{j \geq 1} Y_j$ define the final size of the epidemic. $P(Z = z | X_0 = n, Y_0 = m)$ can be found by summing all the probabilities ie, $|y| = \sum_{j \geq 1} y_i = z$.

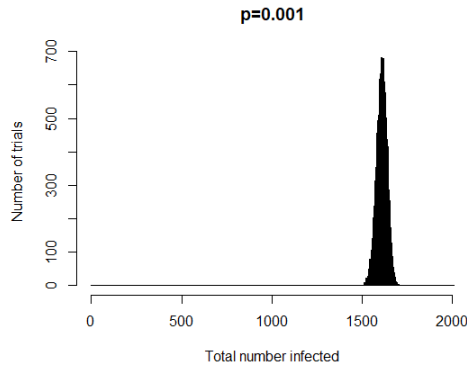


Figure 4.3: *Final size of a Reed-Frost model with $p = 0.001$, $I_0 = 10$ and $S_0 = 2000$.*

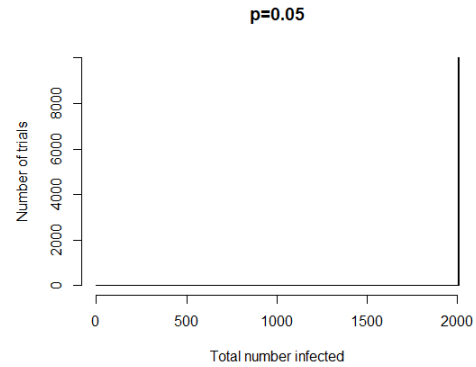


Figure 4.4: *Final size of a Reed-Frost model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.*

The final size also appears to depend on the probability of transmission, when it is very small, the distribution of the final size appears to be normally distributed with a mean just over 1500. The final size of a disease with a probability of infection of 0.05 will always lead to the whole population becoming infected at some stage.

4.1.2 The Greenwood Model

The Greenwood model, developed by Greenwood (1931) is similar to the Reed-Frost model, with the exception that the probability of infection is constant. The conditional probability of the Greenwood model is as follows:

$$\begin{aligned} P(Y_{t+1} = y_{t+1} | X_0 = x_0, Y_0 = y_0, \dots, X_t = x_t, Y_t = y_t) &= P(Y_{t+1} = y_{t+1} | X_t = x_t, Y_t = y_t) \\ &= \binom{x_t}{y_{t+1}} (1 - \alpha)^{y_{t+1}} (\alpha)^{x_t - y_{t+1}}. \end{aligned} \quad (4.4)$$

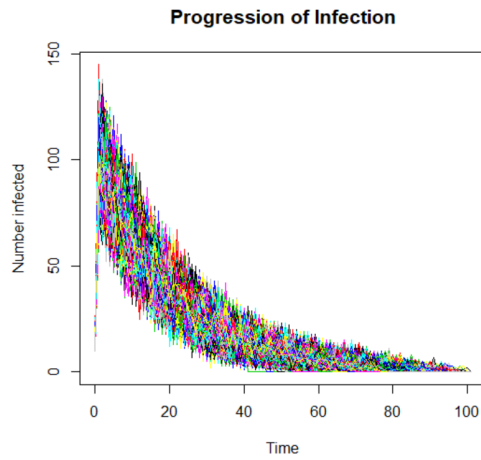


Figure 4.5: *Simulations of a Greenwood model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.*

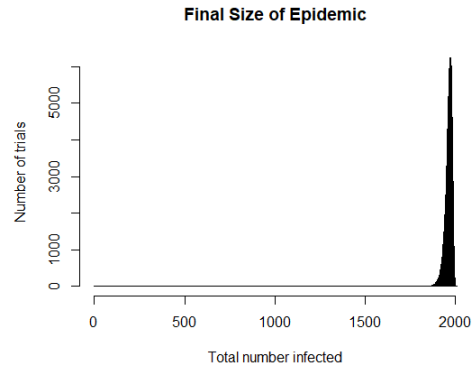


Figure 4.6: *Final size of a Greenwood model with $p = 0.05$, $I_0 = 10$ and $S_0 = 2000$.*

The Greenwood model behaves differently from the Reed-Frost model. Figure 4.6, showing the simulation for 50 000 trial of the Greenwood model looks very different from Figure 4.2 with the same probability of infection and initial conditions. Under the Greenwood model, the disease peaks earlier and then decreases almost monotonically over a longer span of time. The final size under the different models, are different. Although under the Greenwood model. the final size is almost the entire population, there is a bit of variation in the distribution unlike the Reed-Frost case.

4.1.3 Tuckwell-Williams Model

This was presented by Tuckwell and Williams (2007) as a simple SIR stochastic model. The model is outlined in their paper as follows:

Consider an SIR model with the following assumptions:

1. The total population size is fixed, there are no births, deaths, or migrations.
2. Time is discrete, $t = 0, 1, 2, \dots$. The unit for each time period is one day.

-
3. For individual i , $i = 1, 2, \dots, n$, there is a random process, $\{Y^i(t)|t = 0, 1, 2, \dots\}$, defined as:

$$Y^i(t) = \begin{cases} 1, & \text{if individual } i \text{ has been infected in } (t-1, t] \\ 0, & \text{otherwise} \end{cases}$$

Thus, the total number of infected individuals at time t is given by:
 $Y(t) = \sum_{i=1}^n Y^i(t)$, $t \geq 0$.

4. The total numbers of individuals met by individual i is represented by: $N_i(t) = n_i + M_i(t)$, where n_i is a fixed number of individuals randomly drawn from the population. $M_i(t)$ is a random amount of randomly chosen individuals in contact with individual i over time $(t, t+1]$. The model in this thesis is assumed to be a time homogenous model, which means that $M_i(t)$ is constant for all t .
5. If an individual becomes infected, it is in that state for Z consecutive time point, and in this stage they are capable of transmitting the disease. When it leaves this state it infers permanent immunity.
6. The probability of transmission during the period $(t, t+1]$ is denoted as p .
7. The probability that an individual is infected at time t is $f(t) = \frac{Y(t)}{n}$

The model formation allows for the number of contacts to vary over time and hence this can be adjusted for real-life interventions. It does not however, take into account the spatial nature of infectious diseases, which is what the extension includes.

4.1.4 Spatio-Stochastic Models

The Tuckwell and Williams (2007) model is extended in this thesis to produce a novel model which accounts for spatial heterogeneity. This is a versatile model which has been further extended to produce three variations: a spatio-stochastic model with no interaction between spatial units, a spatio-stochastic model which allows interaction but no movement between spatial units and finally a spatio-stochastic model with interaction and migration between spatial units.

The basic assumptions and framework is similar to those mentioned in 4.1.3, but include the following additions:

Defining the Model

The model is discrete in both time and space. Assume a population with n individuals. Assume the recovery time is M periods before an individual obtains immunity.

-
- Let $D \subseteq \mathbb{R}^d$ define the spatial domain. Assume that the domain D is fixed and discrete. The domain is divided into R non-overlapping spatial units with the following properties:

$$s = \{s_1, s_2, \dots, s_R\} \in D \subseteq \mathbb{R}^d \quad \text{and} \quad s_1 \cup s_2 \cup \dots \cup s_R = D.$$

The spatial units can represent any discrete non-overlapping definition such as households, wards, districts or countries.

- The total number of individuals met by individual i , is a function of both time and space and is represented by:

$$N_i(t, s_r) = N_{1i} + M_i(t, s)$$

where N_{1i} is the fixed number of individuals which is randomly drawn from the spatial unit that the individual is in. $M_i(t, s)$ is the random number of individuals in contact with individual i over the period $(t, t+1]$ and spatial units, s and $M_i(t, s) = \sum_{r=1}^R M_i(t, s_r)$.

- The probability of infection at time t and location s_r is

$$f(t, s_r) = \frac{Y(t, s_r)}{n(s_r)}$$

where $Y(t, s_r)$ represent the number of individuals who are infected at time t , $t \in \{1, 2, \dots, T\}$ and at location s_r , $s_r \in \{s_1, s_2, \dots, s_R\}$ and $n(s_r)$ is the number of people in spatial unit s_r and $\sum_{r=1}^R n(s_r) = n$.

Representation as a Markov Chain For individual i , $i = 1, 2, \dots, n$, there is a random process, $\{Y^i(t, s_r) | t = 0, 1, 2, \dots, T, s_r \in \{s_1, s_2, \dots, s_R\}\}$, defined as:

$$Y^i(t, s_r) = \begin{cases} 1, & \text{if individual } i \text{ is infected at time } t \text{ and location } s_r \\ 0, & \text{otherwise} \end{cases}$$

Thus, the total number of infected individuals at time t is given by:

$$Y(t) = \sum_{r=1}^R \sum_{i=1}^n Y^i(t, s_r), \quad t \geq 0. \quad (4.5)$$

and the total number of infected individuals at location s is given by

$$Y(s_r) = \sum_{t=1}^T \sum_{i=1}^n Y^i(t, s_r), \quad r \in \{1, 2, \dots, R\}. \quad (4.6)$$

$X^i(t, s_r)$ and $Z^i(t, s_r)$ are defined similarly for the susceptible and recovered respectively. Since an individual can either be susceptible, infected or recovered

at given time point, for all i , s and t :

$$X^i(t, s_r) + Y^i(t, s_r) + Z^i(t, s_r) = 1 \quad (4.7)$$

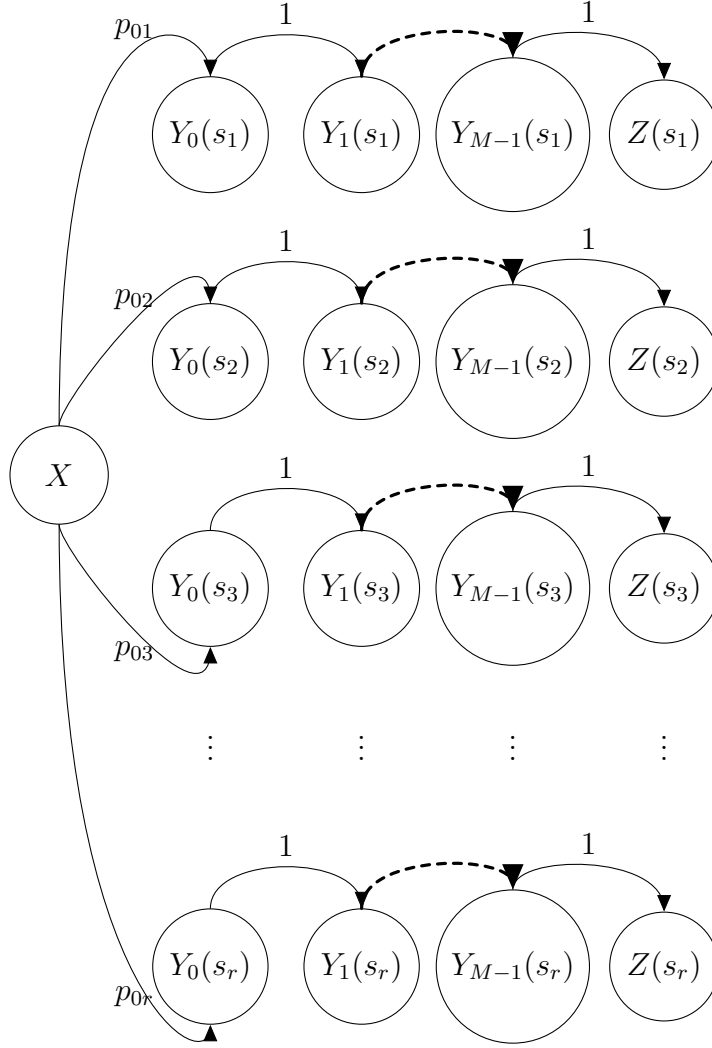


Figure 4.7: *Transition graph Spatio-Stochastic Model of an individual.*

Let $Y_k(t, s_r)$ represent the total number of individuals who have been infectious for exactly k time periods after t , $k \in \{0, 1, 2, \dots, Z - 1\}$ at time t in spatial unit s_r , and Z is the number of recovery periods. Assume that infected people cannot travel between spatial units. The only transition probability which is not zero or one is the transition from susceptible to infected, this is illustrated by Figure 4.7. This probability is denoted by p_{0r} and calculated below. Assume that n is larger than $N_i(t, s_r)$, so that the binomial approximation is valid. The probability of individual i meeting j , $j \in \{0, 1, \dots, N_i(t, s_r)\}$ infected individuals

in s_r is thus given by:

$$P_j^i(Y(t, s_r), N_i(t, s_r), n) \cong \binom{N_i(t, s_r)}{j} \left(\frac{Y(t, s_r)}{n-1} \right)^j \left(1 - \frac{Y(t, s_r)}{n-1} \right)^{N_i(t, s_r)-j}. \quad (4.8)$$

where $N_i(t, s_r)$ is the number of contacts of individual i at location s_r and time t .

The probability of becoming infected if there is contact with j infected individuals is

$$p_j = 1 - (1 - p)^j \quad (4.9)$$

where p is the probability of transmission of an individual during the period $(t, t + 1]$.

Then

$$\begin{aligned} p_{0r} &= P(Y_0^i(t+1) | N_i(t, s_r), X^i(t, s_r) = 1, Y(t, s_r)) \\ &= \sum_{j=1}^{N_i(t, s_r)} p_j P_j^i(Y(t, s_r), N_i(t, s_r), n) \\ &= \sum_{j=1}^{N_i(t, s_r)} (1 - (1 - p)^j) \binom{N_i(t, s_r)}{j} \left(\frac{Y(t, s_r)}{n-1} \right)^j \left(1 - \frac{Y(t, s_r)}{n-1} \right)^{N_i(t, s_r)-j} \\ &= 1 - \left(1 - \frac{pY(t, s_r)}{n-1} \right)^{N_i(t, s_r)}. \end{aligned} \quad (4.10)$$

Equation (4.10) can be obtained with the use of generating functions (Ng and Orav (1990)).

This model presented in section 4.1.4 includes the spatial heterogeneity in the number of contacts which a person makes over time. By including both the time and a spatial component, this model can represent the spread of infectious diseases more accurately. The number of contact that individual i makes in different units, $M_i(t, s)$, can follow any positive distribution. The nature of disease spread should carefully be considered when estimating a distribution for $M_i(t, s)$.

Extending the model to account for latent/exposed periods (SEIR)

The spatio-stochastic model can be further extended to account for a disease with a substantial latent period.

Assume a population with n individuals. Further assume that the latent period is M_1 periods before an individual becomes infectious and the recovery time is M_2 periods before an individual obtains immunity.

For individual i , $i = 1, 2, \dots, n$, there is a random process, $\{Y^i(t, s_r) | t =$

$0, 1, 2, \dots, T, s_r \in \{s_1, s_2, \dots, s_R\}$, defined as:

$$W^i(t, s_r) = \begin{cases} 1, & \text{if individual } i \text{ is exposed but not infectious at time } t \text{ and location } s_r \\ 0, & \text{otherwise} \end{cases}$$

Thus, the total number of exposed individuals at time t is given by:

$$W(t) = \sum_{r=1}^R \sum_{i=1}^n W^i(t, s_r), \quad t \geq 0. \quad (4.11)$$

and the total number of exposed individuals at location s is given by

$$W(s_r) = \sum_{t=1}^T \sum_{i=1}^n W^i(t, s_r), \quad r \in \{1, 2, \dots, R\}. \quad (4.12)$$

Since an individual can either be susceptible, exposed, infected or recovered at given time point, for all i, s and t :

$$X^i(t, s_r) + W^i(t, s_r) + Y^i(t, s_r) + Z^i(t, s_r) = 1 \quad (4.13)$$

Let $W_k(t, s_r)$ represent the total number of individuals who have been infectious for exactly k time periods after t , $k \in \{0, 1, 2, \dots, M_1 - 1\}$ at time t in spatial unit s_r , and Z is the number of recovery periods. Assume that infected people cannot travel between spatial units. The only transition probability which is not zero or one is the transition from susceptible to infected, this is illustrated by Figure 4.8. This probability is denoted by p_{0r} and calculated similarly to Equation 4.10, since the change in exposed is still a function of the number infected:

$$\begin{aligned} p_{0r} &= P(W_0^i(t+1) | N_i(t, s_r), X^i(t, s_r) = 1, Y(t, s_r)) \\ &= \sum_{j=1}^{N_i(t, s_r)} p_j P_j^i(Y(t, s_r), N_i(t, s_r), n) \\ &= \sum_{j=1}^{N_i(t, s_r)} (1 - (1 - p)^j) \binom{N_i(t, s_r)}{j} \left(\frac{Y(t, s_r)}{n-1} \right)^j \left(1 - \frac{Y(t, s_r)}{n-1} \right)^{N_i(t, s_r) - j}. \end{aligned} \quad (4.14)$$

The addition of the exposed compartment just increases the size of the Markov chains. This would result in a delayed peak of infection when comparing the model to the SIR formulation. This change makes sense and corresponds to the deterministic case. The exposed compartment is needed to accurately model infectious diseases with a substantial latent period. In this formulation of the SEIR model, the spatial location of the exposed individuals can also be included in the model.

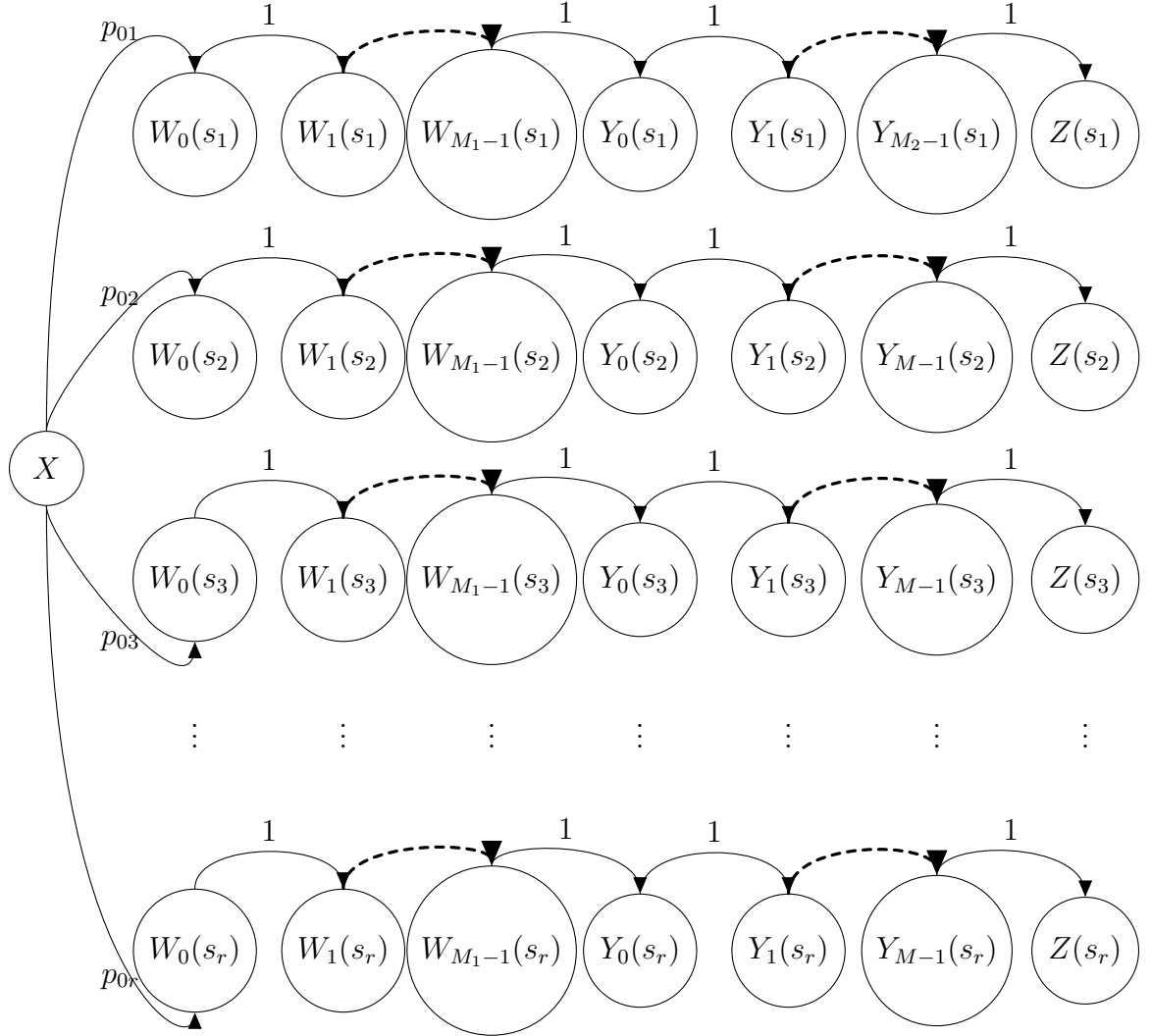


Figure 4.8: *Transition graph Spatio-Stochastic SEIR Model of an individual.*

Variations of the Spatio-Stochastic Model

The extension in section 4.1.4 includes the spatial component that is inherent in the spread of infectious diseases. In this thesis, the spatial model has been developed with three variations, increasing in complexity. Each of the variations can be applied under specific assumptions about the diseases dynamics:

1. Stochastic model with no interactions

In this model, the population is divided into R different sub-populations with a fixed number of individuals and where only local transmission is possible, i.e., individuals cannot leave the spatial unit nor can they infect or be infected by an individual from a different spatial unit. It is assumed that the number of contacts an individual only depends on time: $M_i(t, s) = M_i(t)$. This model should be used to model lockdown scenarios. A lockdown is a type of intervention implemented to prevent the rapid spread of an infectious diseases. Even though transmission from other spatial units are not permitted, the local spread per spatial unit is of interest to allocate resources where they are needed the most.

2. Stochastic models with interactions from other spatial units

Once again, the population is divided into R sub-populations with a fixed number of people in each spatial unit. Under this model, individuals in a spatial unit will remain there, however they can infect and be infected by individuals from other spatial units. The number of contacts, $M_i(t, s)$, now is a function of both time and spatial unit. A suitable spatial function would be one that includes the inverse relationship between distance and number of contacts, since individuals are more likely to be in contact with individuals who reside in closer spatial units. This model would be applicable when modelling the spread of people who reside in an area, but work or go to school in another area and can be infected via that connection.

3. Spatial model with interaction and immigration from other spatial units

The population is divided into R spatial units, but in this case, individuals can move around according to a pre-defined rule and infect people in that spatial unit. This model assumes that people migrate between different spatial units. This model is the most complex as it allows for individuals to move to a different spatial unit. The length of the stay and movement to and from a spatial unit can be chosen to follow any distribution and is included in the simulation update. This can be used to model long-term spread of the disease as migration does occur.

Analysis of Spatio-stochastic Model

The analysis of the stochastic model with no interactions

In this section we consider the simplest version of the spatio-stochastic model. Assume the spatial domain D is divided into R non-overlapping subgroups denoted by $s_r, r \in \{1, 2, \dots, R\}$. Let n_r denote the total number of individuals in spatial unit s_r . Let $N_i(t, s_r)$ denote the total number of contacts that an individual encounters at time t and at s_r . Assume that the number of contacts is constant at a time and unit, ie $N_i(t, s_r) = N$. This assumption would be valid

in the context of a strict lockdown and the number of contacts is reduced to those that you live with. For simplification of notation, let $\{y_j : j = 1, 2, \dots, k\}$ represent the number of infected individuals that an individual encounters at time t and spatial unit s_r . Then applying equation (1), we find that the probability of any susceptible getting infected at time t and location s_r is

$$p_r^* = 1 - \left(1 - \frac{pY(s_r)}{n_r - 1}\right)^N \quad (4.15)$$

where $Y(s_r)$ is the total number of infected individuals in $s_r, r \in \{1, 2, \dots, R\}$.

Let $Y_{1j}(t+1)$ be a random variable for the total number of newly infected individuals at time $t+1$ given that there were j infected individuals. $Y_{1j}(t+1)$ follows a binomial distribution with probability mass function:

$$P(Y_{1j}(t+1) = w) = \binom{N_j^*}{w} (p_r^*)^w (1 - p_r^*)^{N_j^* - w}, \quad (4.16)$$

where p_r^* is defined as equation (3) and $N_j^* = N - y_j$. The expected number of infected is thus $E[Y_{1j}(t+1)] = (N_j^*)(p_r^*)$ with variance $Var(Y_{1j}(t+1)) = (N_j^*)(p_r^*)(1 - p_r^*)$.

The analysis of the stochastic model with interaction from other spatial units

Let us now consider the variation of the stochastic model with interactions from other spatial units. Assume the spatial domain D is divided into R non-overlapping subgroups denoted by $s_r, r \in \{1, 2, \dots, R\}$. Let n_r denote the total number of individuals in spatial units s_r . In this case, the number of contacts is no longer constant and now is a function of the spatial units. Let the number on contacts made by individual i at time t be represented by $N(t, s) = N_r + M(t, s)$, where N_r is the constant number of contacts made in spatial unit s_r and $\{M(t, s) : s = (s_1, s_2, \dots, s_R)\}$ represents the number of contacts made from the other spatial units at time t . In this specific example, we will assume that the number of contacts over time remains constant and that the probability of getting infected in any spatial unit is equal. These assumptions would make sense in a setting where an individual moves to spatial units with similar infection rates. Let $\{y_j : j = 1, 2, \dots, k\}$ represent the number of infected individuals that an individual encounters at time t and spatial unit s_r , applying equation (1), we find that the probability of any susceptible at time t getting infected at time $t+1$ is

$$p(s_r) = 1 - \left(1 - \frac{pY(s_r)}{n_r - 1}\right)^{N(t, s)} \quad (4.17)$$

where $Y(s_r)$ is the total number of infected individuals in s_r .

Let $Y_1(s_r, t+1)$ be the random variable for the total number of newly infected

at time $t + 1$ and unit s_r . Then the distribution of $Y_1(s_r, t + 1)$ is

$$P(Y_1(s_r, t + 1) = w) = \binom{N(t, s)}{w} (p(s_r)) (1 - p(s_r))^{N(t, s) - w}, \quad (4.18)$$

where $0 < w < N(t, s)$. This is a compound binomial distribution.

The assumption of constant infection can be relaxed to derive another variation. By doing this, the number of contacts in each spatial unit is no longer constant and can be varied according to current disease dynamics. This model is useful to determine the transmission when there is interaction between hotspots or super-spreading events.

Assume $N(t, s_r), r \in \{1, 2, \dots, R\}$ contacts that are made at time t by an individual with individuals in other spatial units. Let $j(s_r), r \in \{1, 2, \dots, R\}$ denote the number of infected individuals in s_r . Let $j = \sum_{r=1}^R j(s_r)$ be the total number of infected individuals. The probability of meeting j infected individuals and getting infected is calculated as

$$\begin{aligned} P(\text{Meeting } j \text{ individuals}) &= \prod_{r=1}^R P(\text{Meeting } j(s_r) \text{ individuals}) \\ &= \prod_{r=1}^R \binom{N(t, s_r)}{j(s_r)} \left(\frac{Y(s_r)}{n_r - 1} \right)^{j(s_r)} \left(1 - \frac{Y(t, s_r)}{n_r - 1} \right)^{N(t, s_r) - j(s_r)}, \end{aligned} \quad (4.19)$$

where $Y(t, s_r)$ is the total number of infected people in s_r at time t . Following the same steps as before, we can arrive at the distribution of the number of infected. As one can see, the distribution becomes more complicated and involved, however, the use of simulations on this model allows us to obtain results without explicitly solving these equations.

The analysis of the stochastic model with interaction and migration between spatial units

In this last example, individuals are allowed to move freely. Assume the spatial domain D is divided into R non-overlapping subgroups denoted by $s_r, r \in \{1, 2, \dots, R\}$. Let $n_r(t)$ denote the total number of individuals in spatial unit s_r at time t with $\sum_{r=1}^R n_r(t) = n$. Note that the number of individuals per unit time is no longer constant as individuals can move at each time interval. Let $\sum_{r=1}^R N(t, s_r)$ denote the total number of infected individuals at time t and let $j(t, s_r)$ denote the number of infected individuals from s_r that an individual comes into contact with at time t , where $0 < j(t, s_r) < N(t, s_r), j(t) = \sum_{r=1}^R j(t, s_r)$. The probability that an indi-

vidual meets $j(t)$ infected individuals and gets infected is calculated as

$$\begin{aligned}
P(\text{Meeting } j(t) \text{ individuals}) &= \prod_{r=1}^R P(\text{Meeting } j(t, s_r) \text{ individuals}) \\
&= \prod_{r=1}^R \binom{N(t, s_r)}{j(t, s_r)} \left(\frac{Y(t, s_r)}{n_r(t) - 1} \right)^{j(t, s_r)} \left(1 - \frac{Y(t, s_r)}{n_r(t) - 1} \right)^{N(t, s_r) - j(t, s_r)}.
\end{aligned} \tag{4.20}$$

4.2 Simulation, Application and Analysis of Spatio-stochastic Model

In this section we look at the simulation of the models mentioned in section 4.3.1 including the developed spatio-stochastic model variants. Each model is run over several simulations and then the output is analysed and compared. In this thesis, the analysis of 1 000, 5 000, 10 000 and 15 000 simulations of the models are run. The simulation parameters for each model is defined in 4.2.1.

4.2.1 Simulation Setup

Each model is run on a population of 200 individuals with 2 individuals initially infected. Simulations have been run on the Reed-Frost, Tuckwell-Williams and the three spatio-stochastic models. Simulations were not done on the Greenwood model, since it is similar to the Reed-Frost model and the focus of this paper is on the developed binomial chain models. These parameters were chosen to investigate the differences in the model simulation output. The parameters were chosen so that the disease dynamics in the different models are the same, and hence the output is comparable.

- **Reed-Frost Model**

This is the model defined in 4.1.1. is simulated in the population with a probability of contact between a susceptible (β) of 10%. The other parameters are defined in Table 4.1.

Table 4.1: Parameters of the Reed-Frost Simulations.

Parameter	Value
Initial number of infected individuals	2
Initial number of susceptible individuals	198
Probability of transmission and infection(p)	0.01, 0.0075, 0.005

- **Tuckwell-Williams Model**

This model defined in 4.1.3 was simulated in the population. An individual was assumed to make 4 contacts per time unit. The Tuckwell-Williams model was run with the parameters defined in Table 4.2.

Table 4.2: Parameters of the Tuckwell-Williams Simulations.

Parameter	Value
Initial number of infected individuals	2
Initial number of susceptible individuals	198
Probability of transmission	0.1
Recovery Period	2
Number of fixed contacts per individual	4

- The Spatio-stochastic Models

The population is divided into 4 spatial units, each with an equal number of individuals. The Spatio-stochastic models were run with the parameters defined in Table 4.3.

Table 4.3: Parameters of the Spatio-stochastic Model Simulations.

Parameter	Value
Initial number of infected individuals	2
Initial number of susceptible individuals	198
Probability of transmission	0.1
Recovery Period	2
Number of fixed contacts per individual	4
Number of spatial units	4
Initial number of individuals per spatial unit	4

- *Spatio-Stochastic Model: Case I*

In this case, two spatial units have one infected individual each and there are no infected individuals in other two spatial units.

- *Spatio-Stochastic Model: Case II*

In this case, two spatial units have one infected individual each and there are no infected individuals in other two spatial units. It is further assumed that 1% of infections are obtained from other spatial units.

- *Spatio-Stochastic Model: Case III*

In this case, two spatial units have one infected individual each and there are no infected individuals in other two spatial units. It is assumed that there is an equal probability of migrating between spatial units and remaining in same unit.

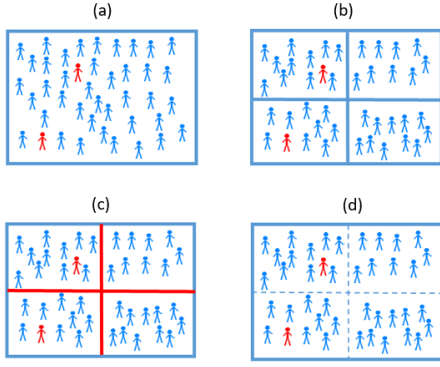


Figure 4.9: A diagrammatic representation of the population: (a) Reed-Frost and Tuckwell-Williams populations, (b) Spatio-Stochastic Case I population, (c) Spatio-Stochastic Case II population, (d) Spatio-Stochastic Case III population.

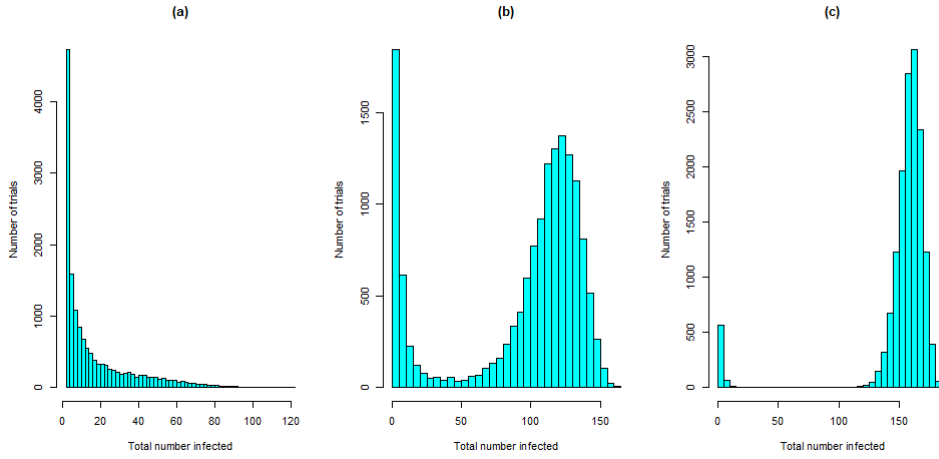


Figure 4.10: Final size of 15 000 simulations of Reed-Frost model varying the probability of transmission and infection: (a) $p = 0.005$, (b) $p = 0.0075$, (c) $p = 0.01$.

4.2.2 Simulation Results

The Reed-Frost Model

The progression of the infectious disease would depend on the probability that a successful infection occurs. When the probability is low, the disease rarely would become epidemic, however as you increase the probability, there are far more simulations with high final sizes. Figure 4.10 shows the effect of changing the probability.

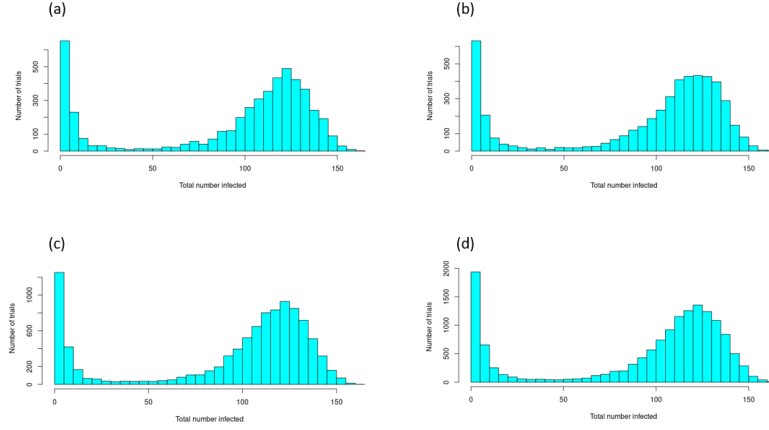


Figure 4.11: *Final size of Reed-Frost model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.*

With respect to the number of simulations, the observed distribution for the final size for 1000, 5000, 10 000, 15 000 all have the same shape and peak. As the number of simulations increases, the distribution does become smoother.

The Tuckwell-Williams Model

Figure 4.12 are plots of the final size of the population at each of the indicated number of simulations. From these figures, it can be seen that the distribution of the final size has the same shape and peak for all of the simulated runs. Almost a third of the simulations result in no infection occurring with the another peak at just over 150 individuals. The distribution is similar to that of the Reed-Frost model with the second peak being much lower ie a lot fewer cases result in epidemics.

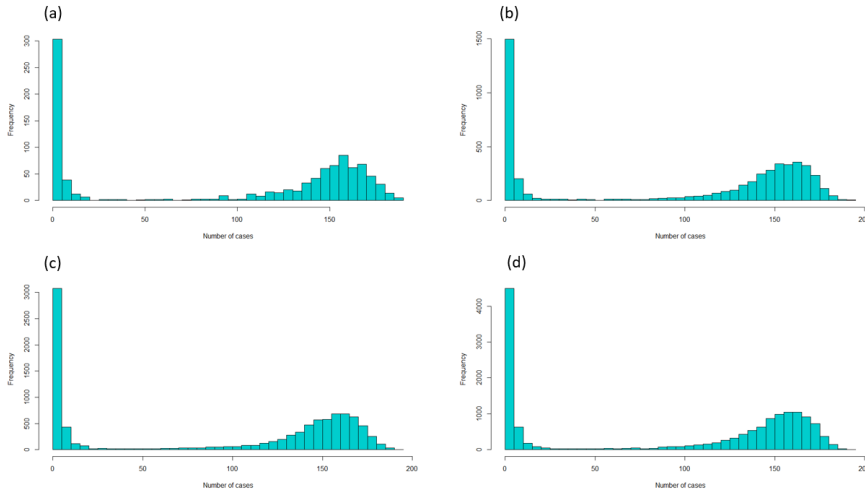


Figure 4.12: *Final size of Tuckwell-Williams model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.*

The Spatio-Stochastic Model

Spatio-Stochastic Model: Case I

In this case, there is no interaction or movement between spatial locations. This has been initialised with one individual in the first and second sub-population each. Figure 4.13 illustrates the final size over 1000, 5000, 10 000 and 15 000 simulations. As the number of simulations increases, the shape remains the same, but becomes smoother. The peak of the final size is around 40 individuals with the peak never exceeding 100 individuals. This makes sense as the disease would not be able to spread out of an infected sub-population to an infected one. The closed model indicates that a complete lockdown would prevent a wide-spread epidemic. About a fifth of the simulations result in the disease dying out and not infecting any more individuals.

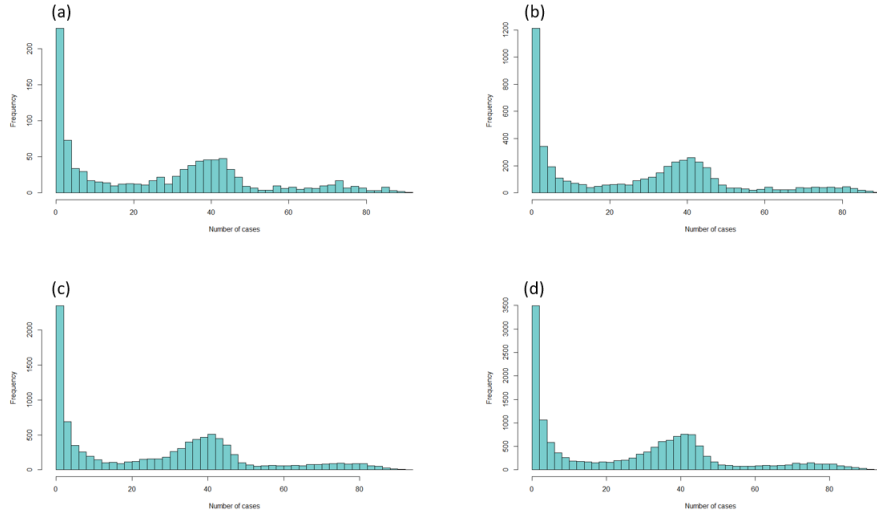


Figure 4.13: *Final size of Spatio-stochastic I model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.*

The final size of each sub-population for 15 000 simulations is in Figure 4.14. Only the first two sub-populations would have final sizes that vary over the simulations. The other two will have no infected the entire time as this is a closed population and only local transmission is possible.

Spatio-Stochastic Model: Case II

In this case we run the simulations over the second spatial model which allows for interaction between spatial units, but no movement. From Figure 4.15, it can be seen that the model converges over all iterations. The shape of the final size remains the same, and only becoming smoother as the number iterations increases.

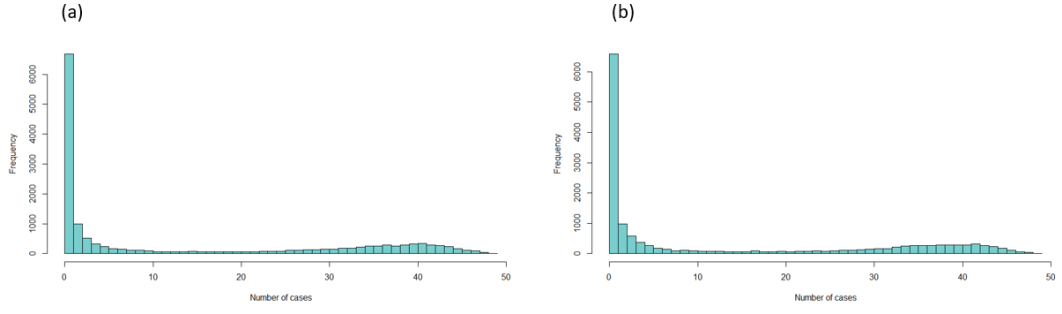


Figure 4.14: *Final size of Spatio-stochastic model I over 15 000 simulations in : (a) subpopulation 1, (b) subpopulation. 2*

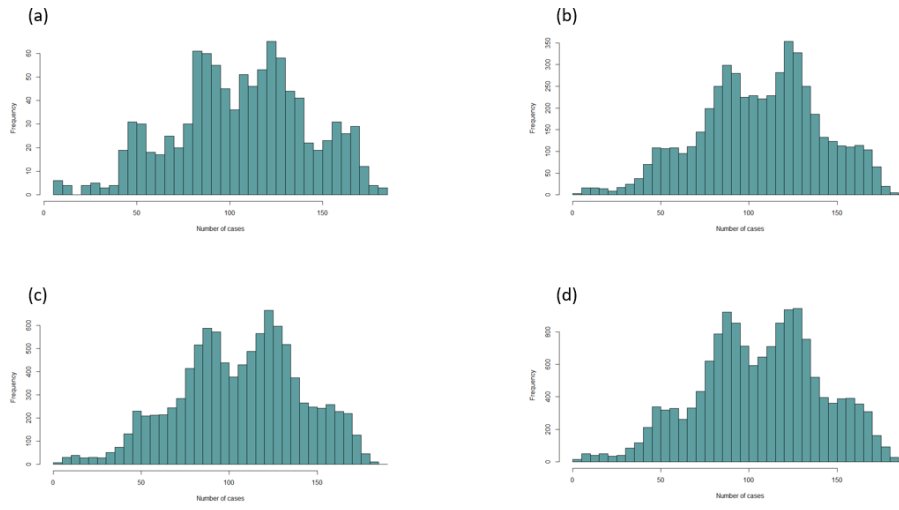


Figure 4.15: *Final size of Spatio-stochastic II model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.*

Spatio-Stochastic Model: Case III

In this case, individuals are allowed to move around and interact with other locations. For this example, we assume that each individual has a 25% of moving to any of the other locations or remaining in the same area. This is a simple assumption that can be adjusted for particular area dynamics. We assume that infected individuals cannot move since they are sick. This would apply well to a disease where there are no asymptomatic infections.

Figure 4.16 indicates the final sizes over each of the simulation runs. The shape stays the same for all runs.

The different spatial models provided different outcomes from each other. This is an indication that the circumstances under which a disease develops can yield different outcomes and one should carefully consider the particular case that they are dealing with when selecting a model.

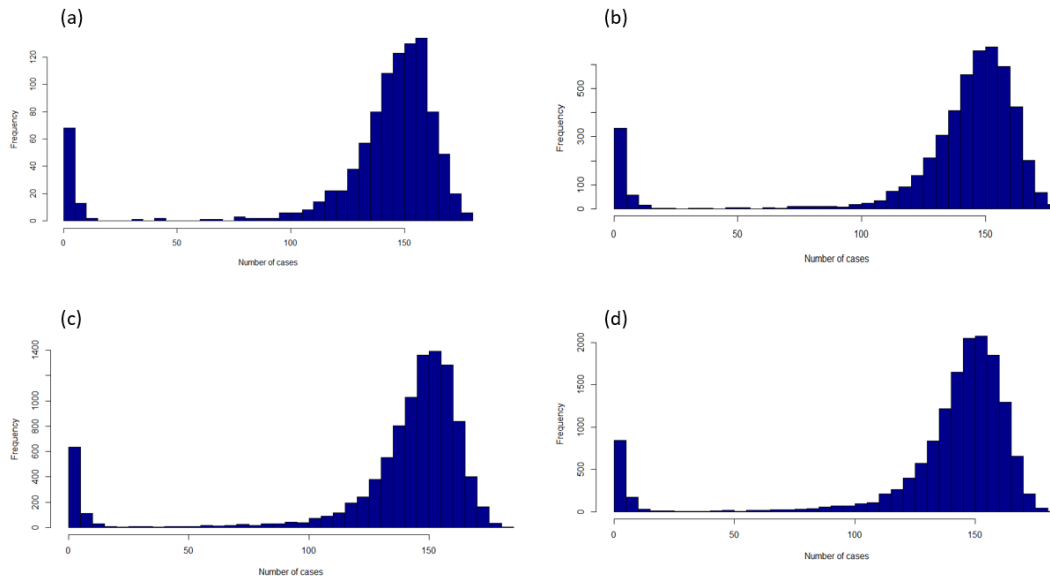


Figure 4.16: *Final size of Spatio-stochastic III model over: (a) 1000 simulations, (b) 5000 simulations, (c) 10000 simulations and (d) 15000 simulations.*

4.2.3 Application to Measles Data

The spatio-stochastic model is now used to analyse the spread of measles using a dataset from Hagelloch, a German village (Höhle (2007)). An SIR model is a suitable model to use to analyse the spread of measles since once an individual has recovered from measles they infer permanent immunity. In this model, we assume 188 susceptible individuals, which is the number of children who were still susceptible to measles. The population is divided into four spatial units each with 47 individuals. A spatio-stochastic model with interaction but no migration was fitted to the dataset. The probability of transmission of measles when in contact with an infected individual was assumed to be 80%, which was an adaption to the probability given by the CDC. The model was initialised with one infected individual in a spatial unit. It was assumed that an individual makes two fixed contact within the respective spatial units and the number of random contacts from other spatial locations was generated from a random uniform distribution. The spatio-stochastic model was implemented using a week for each time step. An individual was infectious one time-unit after they were infected which incorporates the incubation period, hence an SIR framework was still applicable. Four different models were fitted, the first model assumed that there was no additional mixing from other spatial units, the second model assumed that an individual made contact with 2.5% of individuals from other spatial units, the third model assumed contact with 3% of individuals from other spatial units and the fourth model assumed contact with 3.5% of individuals from other spatial units.

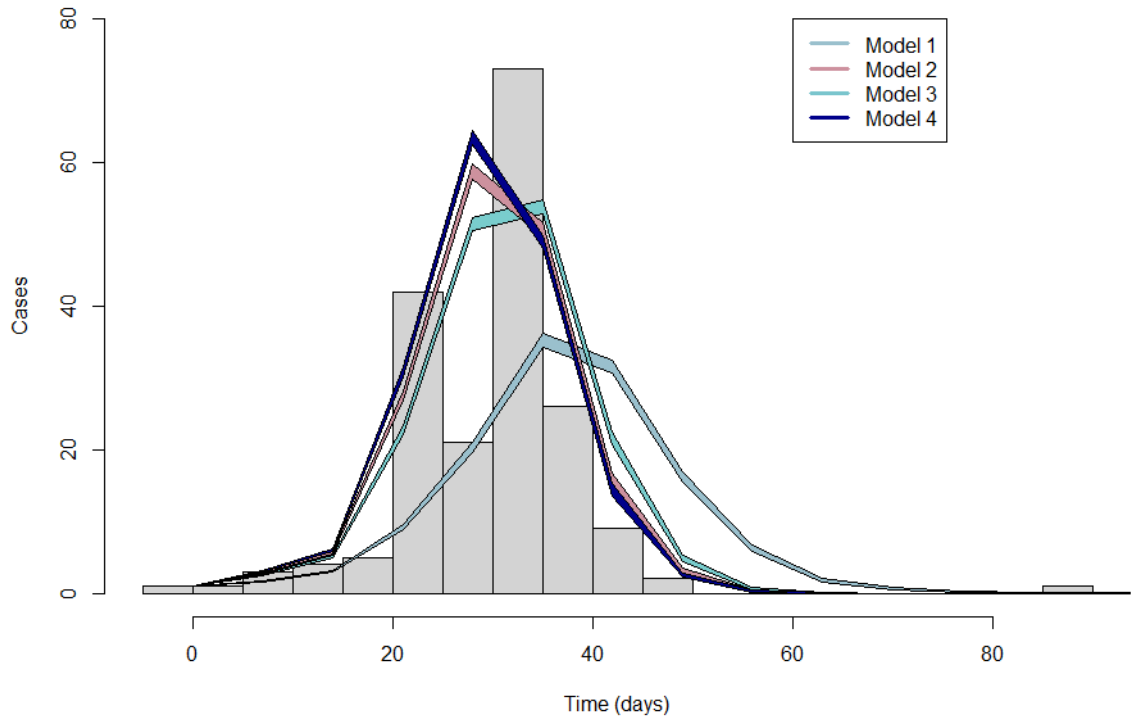


Figure 4.17: *95% Confidence intervals fit for Models 1,2,3 and 4 over 1000 simulations.*

Figure 4.17 shows the 95% confidence interval of each of the four models over the original data. The model which assumes no spatial mixing does not fit the real data very well. The more spatial mixing assumed, the better the fit of the model. Model 4 is the closest to the observed data. This example indicates the need for the spatial interaction in spatial models, as one is better able to capture the spread of an infectious disease.

4.3 Discussion

In this section, we considered an infectious disease with the same level of transmissibility using different models. In total, five different models were used: the Reed-Frost model, the Tuckwell-Williams model as well as the three variations of the spatio-stochastic model developed in this thesis. For the comparison, we will be using the results for each model run for 15 000 simulations.

Figure 4.18 indicates the final size of each of the three spatio-stochastic models over the 15 000 simulations. The third case of the spatio-stochastic model has

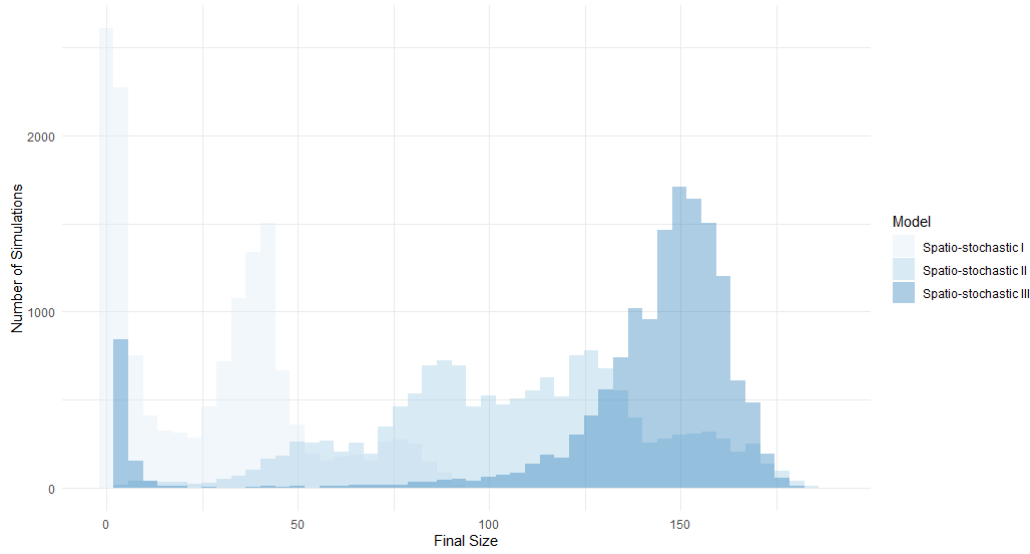


Figure 4.18: *Comparison of the final size of the Spatio-stochastic Models.*

the peak number with a final size around 150 individuals, however. The first spatial model has a peak at around 100 individuals, this makes sense as this model does not allow infected individuals to infect people in regions where there is no infection. The second spatial model is almost symmetrical with two peaks on either side of 100 individuals. The distribution of the final size of this distribution is very different and more spread out than the other models.

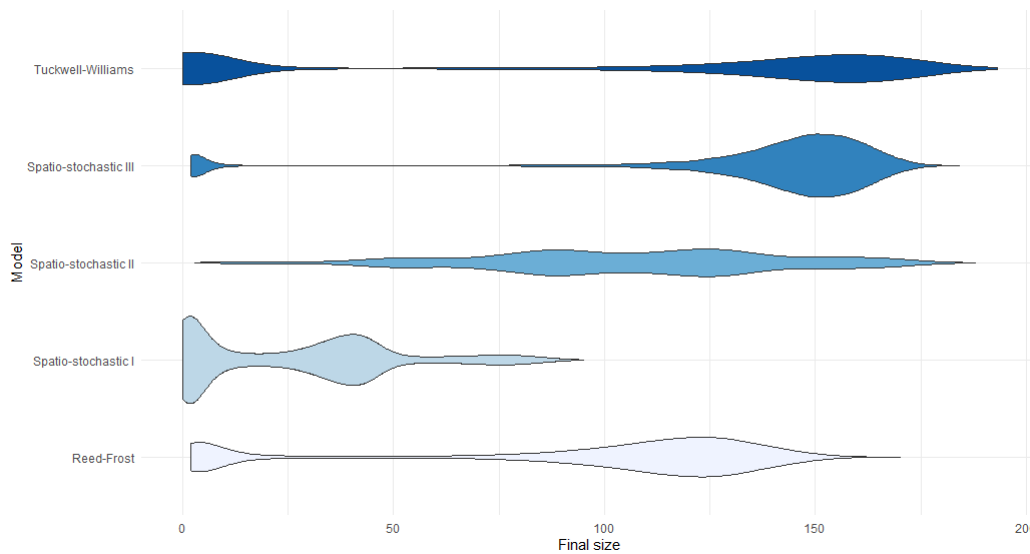


Figure 4.19: *Violin plot of the final size simulations for each of the models.*

From the violin plot (Figure 4.19), it is clear that each model's distribution is different. The Spatio-stochastic I model produces the shorted range with most final size of the disease being less than 100 individuals. The Tuckwell-Williams

Table 4.4: Summary of Final Size per Model.

Model	P(No Transmission)	Average Final Size
Reed-Frost	0.018	152.4
Tuckwell-Williams	0.0934	98.46
Spatio-Stochastic I	0.0949	28.84
Spatio-Stochastic II	6.67^{-5}	106.8
Spatio-Stochastic III	0.0269	135.3

model has a big range and more than half of the simulations result in the final size being less than 150 individuals. The three spatial models are completely unique. The first spatial model always has a final size of less than 100 individuals whereas the third spatial model rarely has less than 100 individuals. This emphasises the effect of migration and integration in the spread of infectious diseases. Table 4.4 summarises the probability of no further transmission and the average final size for each of the models over the 15 000 simulations. Ideally, a disease with a high probability of no transmission and a low number of average infected would be preferable. The Reed-Frost model paints the most dire situation of the disease and the first spatial model, the least. Spatio-stochastic II model has the lowest probability of no transmission and half the population is infected on average. A possible reason that the probability of no infection is lower than Spatio-stochastic III model is that even though individuals are allowed to move between spatial locations, the infections are only due to local cases.

Another observation that can be made on the effect of the parameter choice is illustrated by Figure 4.2. It can be seen that the choice of the parameter can determine extent and the timing of an epidemic. This has been illustrated with the use of the Reed-Frost model, but this result can be observed in all the spatial and non-spatial models. From this section it is clear that the assumptions and the choice of model play a big role in the analysis of the infectious diseases predictions. The models presented here are not competing models, and instead each applicable in certain conditions. One should understand the dynamics of the disease and the spread of it before selecting a model. The mode of transmission, contact-structure and spatial interactions are vital aspects that need to be correctly represented.

4.4 Conclusion

In this chapter an array of models adaptable for different situations have been developed. The modeller should carefully consider the dynamics of the disease before selecting which one of these stochastic models to analyse the spread of an infectious disease. The models presented in this thesis allow the modeller

to have control of the level and method that can be used to introduce the spatial component. Mobility and contact data can be used to better understand and build the parameters of the spatial component. This spatio-stochastic model allows you to incorporate the mobility data by allowing it to inform the movement patterns that are used. The spatio-stochastic models are flexible and can be tailored to the dynamics of a particular disease’s spatial spread. This aspect makes these models very useful in understanding and analysing the spatial spread of the infectious disease. The developed spatio-stochastic model can be applied to a wide variety of cases. The fact that it is stochastic allows not only the ability to predict average behaviour, but also to calculate probabilities related to vital questions in planning and prevention. Questions like what is the probability of outbreak in a particular area or which areas are likely to be hot spot areas. The different variations allow for modelling different stages of an infectious disease such as lock-downs, travel restrictions and interventions. The model has been set up as a Markov chain which allows us to apply well known concepts from stochastic processes for the simplification and analysis of the model.

The modelling of the spread of infectious diseases using a spatial model, like the model developed in this thesis, is important to inform decisions that need to be made on how to stop and prevent further infection. It is also useful in identifying the effects of different interventions like vaccines. Due to globalisation and the level of connection in this society, the spatial modelling is vital. The simulation study demonstrated the differences between other binomial-chain based models, namely the Reed-Frost and the Tuckwell-Williams model. The inclusion of spatial heterogeneity produced outcomes that varied from the two non-spatial models. Further, the spatial models can provide information on sub-populations which can be used to allocate resources accordingly.

This spatial model is not without any disadvantages. After the simulation study, it became evident that the amount of time it takes to run a model is quite significant. The size of the populations in the study was 200, with at most four spatial units and it took almost a day to run on a standard home computer. Given the rapidly improving capacity of computers and the availability of high computing and processing unit, larger areas and populations should not be a great problem. However, if run locally, the time could be substantial. One way to overcome this is to make the underlying algorithms more computationally efficient. This can be done with the use of generating functions to simplify the mathematical input.

Another limitation is that it is not a “one model fits all” approach given the difference in structures of infectious diseases. This model is applicable to those infectious diseases that can be classified as an SIR type. Other extensions that include more complex disease dynamics would need to be derived and added. The work done in the paper is the tip of the iceberg of all the further research that can be conducted. Firstly, the limitations can be addressed.

The models that are presented can be further simplified with the use of generation functions and the algorithm can be made more efficient to improve the running time. Further work can be done to extend this model to include more complex structure of infectious diseases such a drug-resistant diseases an co-infection stochastic models. Another area for further research is to design software packages which makes this model more accessible to users.

Chapter 5

Conclusion

5.1 Review of thesis

The modelling of infectious diseases is an important discipline to better understand infectious diseases and to plan preventative strategies in an efficient manner. These models can be developed in many ways to include the dynamics of infectious diseases. Important features in the disease modelling are the inclusion of the variability and spatial heterogeneity that is inherent in the spread of infectious diseases. The aim of this thesis was to develop a novel spatio-stochastic model which can be used to analyse the spread of an infectious disease. The model that was developed belonged to a family of models known as the compartmental or SIR models. In particular, the model belongs to the family of binomial chain models which are discrete-time stochastic models.

The thesis began with the introduction and background of the modelling of infectious diseases followed by the aims and objectives. The objectives of this thesis was to review and apply current compartmental models, to justify the need for a spatial component, to develop a spatio-stochastic model which was to be compared to other existing models and finally applied to a dataset to demonstrate the developed model. Some infectious disease terminology was defined to give those without a background in epidemiology context to the models and techniques.

The second chapter defined and demonstrated various deterministic and stochastic compartmental models which are used in the modelling of infectious diseases. This included reviewing and exploring compartmental models to ascertain the potential for improvement and enhancement of the models and techniques. Under the deterministic compartmental models, the basic SIR model was defined along with the assumptions of the model. This model was illustrated with a simplified COVID-19 example highlighting the impact of social distancing. The stratified model was then defined and demonstrated using a COVID-19 model that allows for vaccinations. As a novel infectious disease, there is much uncertainty that lies in model estimates due to the unknown

nature of the disease and the different mutations. To account for these uncertainties, we recommended to conduct a sensitivity analysis to understand the impact of changing parameters, to run different case scenarios so that you can account for some of the uncertainty and to predict over shorter intervals and to frequently update the model as new information arises.

The stochastic versions of the compartmental model was then defined and demonstrated. The stochastic compartmental model can be defined as Markov chains leading to a wealth of information from the theory of stochastic processes. The stochastic compartmental model was used to model HIV transmission to illustrate the usefulness of the underlying stochastic theory to solve the stochastic differential equations with the use of generating functions.

The second chapter emphasised the multidisciplinary nature of the modelling of infectious diseases and the wide range of techniques that can be used just within the space of one type of model.

In Chapter 3, the spatial autocorrelation of an infectious disease, namely TB, was studied. The spatial autocorrelation was measured with Moran's I under different spatial weights to understand the the spatial relationship, and how sensitive the spatial relationship is to the choice of weight.

The results indicated that there is spatial autocorrelation regardless of the choice of weight, however, the extent and interpretation is different. This justifies the need for a model that includes the spatial heterogeneity as well as the importance of correctly modelling and defining the relationship between spatial locations. Other notable conclusions that can be drawn from this study is that the quality and spatial resolution of the data is important for the analysis of spatial autocorrelation and that the inclusion of other types of data such as mobility or socio-economic data can be useful in understanding the drivers behind the spatial autocorrelation.

The final chapter contains the main and remaining objectives and is where the spatio-stochastic model is developed, compared and applied to data. The model was developed in discrete time and space as it allows for flexibility in the building of the model and can be simulated in a similar manner to how a disease develops in real life. A spatio-stochastic model in this space belongs to the binomial chain family of models. In order to develop a novel spatio-stochastic model within the binomial chain-space, current models are reviewed and then extended to include the spatial component. This begins with an introduction to binomial-chain models with well-established models such as the Reed-Frost and Greenwood model. The Tuckwell-Williams model was a notable model with the number of contacts being a function of time. This formulation was further extended to include the spatial component. The spatio-stochastic model was then defined, analysed and illustrated under different cases to fit the different scenarios that occur in real-life. As pointed out in the third chapter, the spatial relationship plays an important role in the results and hence the different cases were developed to allow flexibility in modelling. The modeller should carefully consider the dynamics of the disease

before selecting one of these stochastic models to analyse the spread of an infectious disease. The models presented in this thesis allow the modeller to have control of the level and method that you can introduce the spatial component. Mobility and contact data can be used to better understand and build the parameters of the spatial component. This spatio-stochastic model allows you to incorporate the mobility data by allowing it to inform the movement patterns that are used. The spatio-stochastic models are flexible and can be tailored to the dynamics of a particular disease's spatial spread. This aspect makes these models very useful in understanding and analysing the spatial spread of the infectious disease. The developed spatio-stochastic model is then compared to the previous binomial chain formulations. It is important to note that each formulation results in a different expected progression in the infectious disease. The simulation study demonstrated the differences between other binomial-chain based models, namely the Reed-Frost and the Tuckwell-Williams model. The inclusion of spatial heterogeneity produced outcomes that varied from the two non-spatial models. Further, the spatial models can provide information on sub-populations which can be used to allocate resources accordingly. Finally, the spatio-stochastic model was applied to a measles dataset to demonstrate the uses of the model and the added value of including the spatial component in an infectious disease model.

5.2 Significance and contributions of this study

In this thesis, a novel spatio-stochastic model for the spread of infectious diseases has been developed. This model belongs to the binomial-chain family of models with the added advantage that it allows the number of contacts made by an individual to vary over time and space. Unlike other spatial models, this formulation allows you to include various techniques and methods to model the spatial component. This is extremely useful as the prior binomial chain models have rates which are constant with respect to spatial location, which are unrealistic when modelling the spread of infectious diseases.

The developed spatio-stochastic model can be applied to a wide variety of cases and can be adapted to model different movement patterns of individuals. The model is formulated as an SIR and SEIR model which allows for the modelling of diseases with a significant latent period. This addition was not done in the preceding models. The fact that it is stochastic allows not only the ability to predict average behaviour, but also to calculate probabilities related to vital questions in planning and prevention. Questions like what is the probability of outbreak in a particular area or which areas are likely to be hot spot areas. The different variations allow for modelling different stages of an infectious disease such as lock downs, travel restrictions and interventions. The model has been set up as a Markov chain which allows us to apply well known concepts from stochastic processes for the simplification and analysis of the model.

The modelling of the spread of infectious diseases using a spatial model, like the model developed in this thesis, can be used to inform decisions that need to be made on how to stop and prevent further infection. It is also useful in identifying the effects of different interventions like vaccines. Due to globalisation and the level of connection in this society, the spatial modelling is vital.

In addition to the development of the spatio-stochastic model, other noteworthy contributions that come from this thesis are the development of a model for COVID-19 that addresses the fact that the vaccination is *leaky* and the variability that exists in a novel disease; the use of generating functions to reduce the complex stochastic equations into a solvable form for a model of HIV/AIDS and finally the demonstration of the effect of using different weight definitions on the calculation of spatial autocorrelation.

This thesis provides a broad demonstration of the many disciplines, techniques and considerations that need to be made when modelling an infectious disease.

5.3 Limitations of the study

This spatial model is not without disadvantages. After the simulation study, it became evident that the amount of time it takes to run a model is significant. The size of the population in the study was 200, with at most four spatial units and it took almost half a day to run on a standard home computer. Given the rapidly improving capacity of computers and the availability of high computing and processing power, running models over larger areas and populations should not be a great problem. However, if run locally, the time could be substantial. This is not necessarily a problem with the spatial model, but the way that the algorithm has been coded. One way to overcome this is to make the underlying algorithms more computationally efficient. This can be done with the use of generating functions to simplify the mathematical input or by using computational entities that run more efficiently. Another option to run the algorithm faster is to use different software. Though R has many benefits as a statistical programme, other programming software such as Python or Java could be more efficient.

Another limitation is that it is not a “one model fits all” approach given the difference in structures of infectious diseases. This model is applicable to those infectious diseases that can be classified as a SIR or SEIR model. As pointed out in the earlier chapters and the literature review, infectious disease models have more compartments and elements that need to be included to make the model realistic. Other extensions that include more complex disease dynamics would need to be derived and added. This leaves an area for further research and development. However, the simple spatio-stochastic models can be used to understand the spatial impacts of the model which can then inform decisions in a more comprehensive model.

There are some other limitations related to the contributions other than the spatio-stochastic model. The model development of the COVID-19 model can be improved with more data to estimate parameters instead of getting them from literature. Another limitation with regards to this model is the inability to predict far into the future accurately. Though the model can accurately model over one wavelength, it starts to diverge for periods longer than that. This is due to the ever-changing nature of the disease and the model should be built so that it is easily updated with new information. The HIV/AIDS model solution can be found using mathematical software and/or by working with specialists in differential equations to solve the equations optimally. Finally, for the justification of the autocorrelation, additional data related to the drivers of TB can be added to better understand the spatial autocorrelation patterns observed.

5.4 Further Work

The work done in the thesis is the tip of the iceberg of all the further research that can be conducted. Firstly, the limitations can be addressed. The models that are presented can be further simplified with the use of generation functions and the algorithm can be made more efficient to improve the running time. Admittedly, this is more an exercise in computer science and mathematics than it is of statistics. Further work can be done to extend this model to include a more complex structure of infectious diseases such as drug-resistant diseases and co-infection stochastic models.

Another area for further research is to design software packages which makes this model more accessible to users. This can be achieved by creating a user-friendly package in R which allows non-statisticians to use the model with relative ease. This should be done once further extensions on the base model are developed so that the researcher can have access to a wider range of modelling tools to build models that represent reality.

The last area for future work is to look at the integration of different mobility and network data types into the spatial model. As mentioned before, the spatio-stochastic model allows flexibility in the estimation of the spatial relationship. Instead of making assumptions of an individuals movement, this can be estimated and modelled from mobility data. It would be interesting to compare the performance of models based on different spatial data with each other.

Other areas for future research involve collaborations with researchers in epidemiology, mathematics and computer science to get their insight on areas of future development. Though the statistical models are developed with sound theory based in stochastic processes, to make the model more practical and useful to stakeholders making decisions, it is vital to obtain input from the

other disciplines. More areas for future research could include the following: (a) Use of spatio-temporal modelling of infectious diseases. (b) Multi-scale modelling of infectious disease systems.

This thesis has been an amazing opportunity to learn and grow in the space of stochastic modelling of the spread of infectious diseases. Although many parts were challenging, the overall experience was deeply engaging and fulfilling. I hope that the outcomes of the research forms the bedrock of the academic journey of myself and others.

References

- Allen, L. J. (2010). *An introduction to stochastic processes with applications to biology*. Chapman and Hall/CRC.
- Allen, L. J., Brauer, F., Van den Driessche, P., and Wu, J. (2008). *Mathematical Epidemiology*, volume 1945. Springer.
- Angulo, J., Yu, H.-L., Langousis, A., Kolovos, A., Wang, J., Madrid, A. E., and Christakos, G. (2013). Spatiotemporal infectious disease modeling: a bme-sir approach. *PloS one*, 8(9):e72168.
- Anselin, L. (1995). Local indicators of spatial association—lisa. *Geographical Analysis*, 27(2):93–115.
- Bailey, N. T. et al. (1975). *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.
- Barreto, M. L., Teixeira, M. G., and Carmo, E. H. (2006). Infectious diseases epidemiology. *Journal of Epidemiology & Community Health*, 60(3):192–195.
- Beichelt, F. (2018). *Applied Probability and Stochastic Processes*. CRC Press.
- Bisin, A. and Moro, A. (2020). Learning epidemiology by doing: The empirical implications of a spatial-sir model with behavioral responses. Technical report, National Bureau of Economic Research.
- Bivand, R., Keitt, T., and Rowlingson, B. (2018). *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*. R package version 1.3-3.
- Bivand, R. and Rundel, C. (2018). *rgeos: Interface to Geometry Engine - Open Source ('GEOS')*. R package version 0.3-28.
- Bivand, R. and Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3):716–748.
- Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, Second edition*. Springer, NY.
- Boutayeb, A. (2006). The double burden of communicable and non-communicable diseases in developing countries. *Transactions of the Royal society of Tropical Medicine and Hygiene*, 100(3):191–199.

-
- Castillo-Chavez, C. and Song, B. (2004). Dynamical models of tuberculosis and their applications. *Mathematical Biosciences and Engineering*, 1(2):361–404.
- Chemaitelly, H., Tang, P., Hasan, M. R., AlMukdad, S., Yassine, H. M., Benslimane, F. M., Al Khatib, H. A., Coyle, P., Ayoub, H. H., Al Kanaani, Z., et al. (2021). Waning of bnt162b2 vaccine protection against sars-cov-2 infection in qatar. *New England Journal of Medicine*, 385(24):e83.
- Chen, D., Moulin, B., and Wu, J. (2014). *Analyzing and modeling spatial and temporal dynamics of infectious diseases*. John Wiley & Sons.
- Coovadia, H., Jewkes, R., Barron, P., Sanders, D., and McIntyre, D. (2009). The health and health system of south africa: historical roots of current public health challenges. *The Lancet*, 374(9692):817–834.
- Daley, D. J. and Gani, J. (2001). *Epidemic modelling: an introduction*, volume 15. Cambridge University Press.
- Dietz, K. (1967). Epidemics and rumours: A survey. *Journal of the Royal Statistical Society. Series A (General)*, pages 505–528.
- Docrat, R. (2018). Generating functions in branching processes and birth and death processes.
- Dorrington, R. E., Moultrie, T. A., Laubscher, R., Groenewald, P. J., and Bradshaw, D. (2021). Rapid mortality surveillance using a national population register to monitor excess deaths during sars-cov-2 pandemic in south africa. *Genus*, 77(1):1–17.
- Dowdy, D. W., Golub, J. E., Chaisson, R. E., and Saraceni, V. (2012). Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. *Proceedings of the National Academy of Sciences*, 109(24):9557–9562.
- Drmotá, M. (2009). *Random trees: an interplay between combinatorics and probability*. Springer Science & Business Media.
- Eggo, R. M., Cauchemez, S., and Ferguson, N. M. (2011). Spatial dynamics of the 1918 influenza pandemic in england, wales and the united states. *Journal of the Royal Society Interface*, 8(55):233–243.
- Engle, S., Keppo, J., Quercioli, E. K. M., Smith, L., and Wilson, A. (2020). The behavioral sir model, with applications to the swine flu and covid-19 pandemics. Technical report, Working Paper, University of Wisconsin.
- Fabris-Rotelli, I., Holloway, J., Kimmie, Z., Archibald, S., Debba, P., Docrat, R., le Roux, A., Dudeni-Tlhone, N., Van Rensburg, C. J., Thiede, R., et al. (2021). A spatial seir model for covid-19 in south africa.
- Getis, A. and Ord, J. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3).

-
- Gomes, C. (2020). Report of the who-china joint mission on coronavirus disease 2019 (covid-19). *Brazilian Journal of Implantology and Health Sciences*, 2(3).
- Greenwood, M. (1931). On the statistical measure of infectiousness. *Epidemiology & Infection*, 31(3):336–351.
- Gutierrez, J. B., Galinski, M. R., Cantrell, S., and Voit, E. O. (2015). From within host dynamics to the epidemiology of infectious disease: scientific overview and challenges. *Mathematical Biosciences*, 270:143–155.
- Höhle, M. (2007). surveillance: An r package for the monitoring of infectious diseases. *Computational Statistics*, 22(4):571–582.
- Houben, R. M., Menzies, N. A., Sumner, T., Huynh, G. H., Arinaminpathy, N., Goldhaber-Fiebert, J. D., Lin, H.-H., Wu, C.-Y., Mandal, S., Pandey, S., et al. (2016). Feasibility of achieving the 2025 who global tuberculosis targets in south africa, china, and india: a combined analysis of 11 mathematical models. *The Lancet Global Health*, 4(11):e806–e815.
- Jacquez, J. A. and Simon, C. P. (1993). The stochastic si model with recruitment and deaths i. comparison with the closed sis model. *Mathematical Biosciences*, 117(1-2):77–125.
- Jing, W., Jin, Z., and Zhang, J. (2018). Low-dimensional sir epidemic models with demographics on heterogeneous networks. *Journal of Systems Science and Complexity*, 31(5):1103–1127.
- Karim, S. A. and Karim, Q. A. (2010). *HIV/Aids in South Africa*. Cambridge University Press.
- Keeling, M. J. and Gilligan, C. A. (2000). Metapopulation dynamics of bubonic plague. *Nature*, 407(6806):903–906.
- Kendall, D. G. et al. (1948). On the generalized” birth-and-death” process. *The Annals of Mathematical Statistics*, 19(1):1–15.
- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.
- Krämer, A., Kretzschmar, M., and Krickeberg, K. (2010). *Modern infectious disease epidemiology: Concepts, methods, mathematical models, and public health*. Springer.
- Lee, Y. and Ogburn, E. L. (2017). Testing for network and spatial autocorrelation. *arXiv preprint arXiv:1710.03296*.
- Manjoo-Docrat, R. (2022). A spatio-stochastic model for the spread of infectious diseases. *Journal of theoretical biology*, 533:110943.

-
- Maude, R. J., Socheat, D., Nguon, C., Saroth, P., Dara, P., Li, G., Song, J., Yeung, S., Dondorp, A. M., Day, N. P., et al. (2012). Optimising strategies for plasmodium falciparum malaria elimination in cambodia: primaquine, mass drug administration and artemisinin resistance. *PLoS One*, 7(5):e37166.
- McMichael, A. J. (2004). Environmental and social influences on emerging infectious diseases: past, present and future. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 359(1447):1049–1058.
- Meakin, S. R. and Keeling, M. J. (2018). Correlations between stochastic epidemics in two interacting populations. *Epidemics*.
- Meng, X., Zhao, S., Feng, T., and Zhang, T. (2016). Dynamics of a novel nonlinear stochastic sis epidemic model with double epidemic hypothesis. *Journal of Mathematical Analysis and Applications*, 433(1):227–242.
- Ming, R.-X., Liu, J., Cheung, W. K., and Wan, X. (2016). Stochastic modelling of infectious diseases for heterogeneous populations. *Infectious Diseases of Poverty*, 5(1):1–11.
- Mishra, S., Fisman, D. N., and Boily, M.-C. (2010). The abc of terms used in mathematical models of infectious diseases. *Journal of Epidemiology & Community Health*, pages jech–2009.
- Mollison, D. (1995). *Epidemic models: their structure and relation to data*, volume 5. Cambridge University Press.
- Moran, P. A. (1950). A test for the serial independence of residuals. *Biometrika*, 37(1/2):178–181.
- Nelson, K. E. and Williams, C. M. (2013). *Infectious disease epidemiology*. Jones & Bartlett Publishers.
- Ng, J. and Orav, E. J. (1990). A generalized chain binomial model with application to hiv infection. *Mathematical Biosciences*, 101(1):99–119.
- Ngonghala, C. N., Iboi, E., Eikenberry, S., Scotch, M., MacIntyre, C. R., Bonds, M. H., and Gumel, A. B. (2020). Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel coronavirus. *Mathematical Biosciences*, 325:108364.
- Ord, J. K. and Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27(4):286–306.
- Organization, W. H. (2018). World health statistics 2018: monitoring health for the sdgs, sustainable development goals.
- Organization, W. H. et al. (2016). Global health sector strategy on hiv 2016–2021. towards ending aids. Technical report, World Health Organization.

-
- Pebesma, E. J. and Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5(2):9–13.
- Pillay-van Wyk, V., Bradshaw, D., Groenewald, P., Seocharan, I., Manda, S., Roomaney, R., Awotiwon, O., Nkwenika, T., Gray, G., Buthelezi, S., et al. (2020). Covid-19 deaths in south africa: 99 days since south africa’s first death. *SAMJ: South African Medical Journal*, 110(11):1093–1099.
- Prem, K., Zandvoort, K. v., Klepac, P., Eggo, R. M., Davies, N. G., for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, C., Cook, A. R., and Jit, M. (2021). Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the covid-19 era. *PLoS computational biology*, 17(7):e1009098.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reingold, A. L. (2000). Infectious disease epidemiology in the 21st century: will it be eradicated or will it reemerge? *Epidemiologic Reviews*, 22(1):57–63.
- Riley, S. (2007). Large-scale spatial-transmission models of infectious disease. *Science*, 316(5829):1298–1301.
- Russell, C. D. (2011). Eradicating infectious disease: can we and should we? *Frontiers in Immunology*, 2:53.
- Sadoff, J., Gray, G., Vandebosch, A., Cárdenas, V., Shukarev, G., Grinsztejn, B., Goepfert, P. A., Truyers, C., Fennema, H., Spiessens, B., et al. (2021). Safety and efficacy of single-dose ad26. cov2. s vaccine against covid-19. *New England Journal of Medicine*, 384(23):2187–2201.
- Sen, A. (1976). Large sample-size distribution of statistics used in testing for spatial correlation. *Geographical Analysis*, 8(2):175–184.
- Shaweno, D., Karmakar, M., Alene, K. A., Ragonnet, R., Clements, A. C., Trauer, J. M., Denholm, J. T., and McBryde, E. S. (2018). Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. *BMC medicine*, 16(1):1–18.
- Shinde, V., Bhikha, S., Hoosain, Z., Archary, M., Bhorat, Q., Fairlie, L., Lalloo, U., Masilela, M. S., Moodley, D., Hanley, S., et al. (2021). Efficacy of nvx-cov2373 covid-19 vaccine against the b. 1.351 variant. *New England Journal of Medicine*, 384(20):1899–1909.
- Shrestha, S., Reja, M., Gomes, I., Baik, Y., Pennington, J., Islam, S., Faisel, A. J., Cordon, O., Roy, T., Suarez, P. G., et al. (2021). Quantifying geographic heterogeneity in tb incidence and the potential impact of geographically targeted interventions in south and north city corporations of dhaka, bangladesh: a model-based study. *Epidemiology & Infection*, 149.

-
- Siettos, C. I. and Russo, L. (2013). Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4):295–306.
- Silal, S., Pulliam, J., Meyer-Rath, G., Nichols, B., Jamieson, L., Kimmie, Z., and Moultrie, H. (2020). Estimating cases for covid-19 in south africa update: 19 may 2020. *Update*.
- Silal, S. P., Little, F., Barnes, K. I., and White, L. J. (2014). Towards malaria elimination in mpumalanga, south africa: a population-level mathematical modelling approach. *Malaria Journal*, 13(1):297.
- Silal, S. P., Little, F., Barnes, K. I., and White, L. J. (2016). Sensitivity to model structure: a comparison of compartmental models in epidemiology. *Health Systems*, 5(3):178–191.
- Stein, R. A. (2020). Covid-19 and rationally layered social distancing. *International Journal of Clinical Practice*, page e13501.
- Tennekes, M. (2018). tmap: Thematic maps in R. *Journal of Statistical Software*, 84(6):1–39.
- Toxvaerd, F. M. (2020). Equilibrium social distancing.
- Tuckwell, H. C. and Toubiana, L. (2007). Dynamical modeling of viral spread in spatially distributed populations: stochastic origins of oscillations and density dependence. *Biosystems*, 90(2):546–559.
- Tuckwell, H. C. and Williams, R. J. (2007). Some properties of a simple stochastic epidemic model of sir type. *Mathematical Biosciences*, 208(1):76–97.
- Waema, R. and Olowofeso, O. E. (2015). Mathematical modeling for human immunodeficiency virus (hiv) transmission using generating function approach.
- Wu, C., Yang, Y., Zhao, Q., Tian, Y., and Xu, Z. (2017). Epidemic waves of a spatial sir model in combination with random dispersal and non-local dispersal. *Applied Mathematics and Computation*, 313:122–143.

Appendix A

Mathematical Tools

This chapter defines and outlines the mathematical tools that are needed to simplify and solve the stochastic models developed. The two main mathematical tools used through out this thesis are generating functions, to simplify complicated expression and differential equations, which appear throughout the derivations.

A.1 Generating Functions

The following derivations, formulas and definitions were obtained from Drmota (2009)

A.1.1 Ordinary Generating Functions

The ordinary generating function (OGF) of a sequence $(a_n)_{n \geq 0}$ is the formal power series;

$$a(x) = \sum_{n=0}^{\infty} a_n x^n \quad (\text{A.1})$$

We use the notation $[x^n]a(x) = a_n$ to extract the coefficients of x^n in a generating function.

A generating function $a(x)$ represents an analytic function for $|x| < R$, where

$$R = \left(\limsup_{n \rightarrow \infty} |a_n|^{1/n} \right)^{-1} \quad (\text{A.2})$$

denotes the radius of convergence. Thus if $R > 0$, then we can either use a differentiation to represent the sequence

$$a_n = \frac{a^{(n)}(0)}{n!}, \quad n \geq 0 \quad (\text{A.3})$$

or we use Cauchy's Formula:

$$a_n = \frac{1}{2\pi i} \int_{\gamma} a(x) \frac{dx}{x^{n+1}} \quad (\text{A.4})$$

where γ is a closed curve inside the region of analyticity of $a(x)$ with winding +1 around the origin. Table A.1 is a summary of the combinatorial construc-

Combinatorial construction	OGF
$C = A + B$	$c(x) = a(x) + b(x)$
$C = A \times B$	$c(x) = a(x)b(x)$
$C = A^*$	$c(x) = \frac{1}{1-a(x)}$
$C = A(B)$	$c(x) = a(b(x))$

Table A.1: Combinatorial construction of operations used in counting

tion of some of the operations used in counting problems. The operations are between two OGFs. The first row is the sum, the second row is the product, the third row is the inverse and the last row is the composition between the OGFs. These constructions are useful as they aid computation of complex generating functions.

A.1.2 Bi-variate Generating Functions

Although the generating functions above are very useful, bi-variate generating functions are more useful when assessing functions where there are two parameters of interest.

The ordinary bi-variate generating function of a double indexed sequence a_{nk} is the series:

$$a(u, z) = \sum_{n \geq 0} \sum_{k \geq 0} a_{nk} u^k z^n \quad (\text{A.5})$$

A.1.3 Statistical Generating Functions

Assume X is a discrete random variable with state space $0, 1, 2, \dots$

Let f denote the probability mass function of X defined:

$$f(j) = \text{Prob}(X = j) = p_j, \quad j = 0, 1, 2, \dots$$

where

$$\sum_{j=0}^{\infty} p_j = 1.$$

The mean and variance of X are

$$\mu_X = E[X] = \sum_{j=0}^{\infty} j p_j \quad (\text{A.6})$$

and

$$\sigma_X^2 = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2 \quad (\text{A.7})$$

Probability generating function

The probability generating function (PGF) of the discrete random variable X is a function defined on a subset of the reals, denoted as P_X , and defined by:

$$P_x(t) = E[t^X] = \sum_{j=0}^{\infty} p_j t^j, \quad (\text{A.8})$$

for some $t \in \mathbb{R}$

As the name implies, the PGF generates probabilities associated with the distribution.

In general, the k^{th} derivative of the PGF of X is

$$P_X^{(k)}(0) = k! p_k \quad (\text{A.9})$$

Moment generating function

The moment generating function of the discrete random variable X with state space $0, 1, 2, \dots$ and probability function $f(j) = p_j$, $j = 0, 1, 2, \dots$ denoted $M_X(t)$ is defined as:

$$M_X(t) = E[e^{tX}] = \sum_{j=0}^{\infty} p_j e^{jt} \quad (\text{A.10})$$

for some $t \in \mathbb{R}$

The MGF generates moments $E[X^k]$ of the distribution of the distribution of random variable X , provided the summation converges. In general,

$$M_X^{(k)}(0) = E[X^k] \quad (\text{A.11})$$

Characteristic generating function

The characteristic generating function (CHF) of the random variable X is

$$\phi_X(t) = E[e^{itX}] = \sum_{j=0}^{\infty} p_j e^{ijt} \quad (\text{A.12})$$

where $i = \sqrt{-1}$

Cumulant generating function

The cumulative generating function (CGF) of the discrete random variable X ,

is the natural logarithm of the moment generating function denoted $K_x(t)$:

$$K_X(t) = \ln[M_X(t)] \quad (\text{A.13})$$

The generating functions for continuous random variables are obtained in a similar manner, just integrating instead of taking the sum.

The generating functions above are more efficient in calculating the moments and distribution of variables as it is more tractable than using direct methods. Both birth and death processes and branching processes make use of them in analysis.

A.2 Differential Equations

When working with stochastic processes, we often run into differential equations which need to be solved in order to obtain the desired solution. A differential equation is an equation which contains at least one term involving derivatives of one variable with respect to another. Unlike algebraic equations, the solutions of differential equations are functions and not just numbers. An ordinary differential equation involves a function of a single variable and some of its derivatives. The order of a differential equation is determined by the highest derivative in that equation. When working with generating functions of birth and death processes or branching processes, it is common to find differential equations of the first and second order.

A.2.1 First Order Differential Equations

A first order differential equation has the following form:

$$\frac{dy}{dx} = f(x, y) \quad (\text{A.14})$$

There are two methods which can be used to solve first order differential equations

1. Separation of variables

The separation of variables method is used when a differential equation can be written in the form

$$\frac{dy}{dx} = f(y)g(x) \quad (\text{A.15})$$

where f is a function of y only and g is a function of x only.

We solve this by rewriting the problem as

$$\frac{1}{f(y)} dy = g(x) dx \quad (\text{A.16})$$

and integrate both sides. We use the initial condition to find the constant of integration.

2. Integrating Factor

The integrating factor method is used when the differential equation is or can be arranged in the following form:

$$\frac{dy}{dx} + p(x)y = q(x), \quad (\text{A.17})$$

where p and q are functions of x only.

We solve this by rearranging the differential equation into the standard form and finding the integrating factor ($e^{\int p(x)dx}$). Then multiply through by the integrating factor and rewrite the left hand side as the derivative of $y^{\int p(x)dx}$. Integrating both sides gives the general solution.

Riccati Equations

Riccati looked at the approximation of the first order differential equations to the second degree, he considered equation of the type:

$$y' = P(x) + Q(x)y + R(x)y^2 \quad (\text{A.18})$$

We come across Riccati equations quite often when working with the generating function of birth and death processes. The Riccati equation defined by (2.11) can be solved with the use of the following theorem:

Theorem A.1

If the particular solution y_1 of a Riccati equation is known, then the general solution is given by

$$y = y_1 + u \quad (\text{A.19})$$

Substituting (A.12) into (A.11), we get

$$(y_1 + u)' = Q(x)(y_1 + u) + R(x)(y_1 + u)^2 + P(x)$$

Cancelling the particular solutions, a differential equation for $u(x)$ is obtained:

$$u' = R(x)u^2 + [2R(x)y_1 + Q(x)]u \quad (\text{A.20})$$

which is a Bernoulli equation.

Substituting $z = \frac{1}{u}$ into (2.14) results in a linear differential equation which can be integrated.

If the coefficients $R(x)$, $P(x)$ and $Q(x)$ are constant then the equation is reduced to a separable differential equation (see section A.2.1).

A.2.2 Second Order Differential Equations

The most general linear second order differential equation is in the form:

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y = G(x) \quad (\text{A.21})$$

In order to solve these, we will need to solve the homogeneous and in-homogeneous problem

Appendix B

Stochastic Processes

This chapter introduces the stochastic models typically used in the modelling the transmission of infectious diseases. Birth and death processes are defined along with properties that are later needed to simplify models for infectious diseases.

B.1 Birth and Death Processes

The theoretical framework found in this chapter was originally derived by Kendall et al. (1948), however this version has been compiled predominantly with the use of textbooks of Beichelt (2018) and Allen (2010) originally can be found in my master's thesis (Docrat (2018)).

B.1.1 Discrete-Time Birth and Death Processes

Definition

Let $\{X_n, n = 0, 1, 2, \dots, N\}$ denote a discrete Markov chain with transition probabilities:

$$p_{ji} = P\{X_{n+1} = j | X_n = i\} \quad (\text{B.1})$$

$$p_{ji} = \begin{cases} p_i, & \text{if } j = i + 1 \quad j = 0, 1, 2, \dots \\ q_i, & \text{if } j = i - 1 \quad j = 1, 2, \dots \\ r_i, & \text{if } j = i \quad j = 0, 1, 2, \dots \\ 0, & \text{otherwise,} \end{cases} \quad (\text{B.2})$$

where

$$p_i + q_i + r_i = 1. \quad (\text{B.3})$$

Then $\{X_n\}$ is a birth and death process with transition probabilities p_i and q_i . The transition matrix P for the finite Markov Chain ($N < \infty$), has the fol-

lowing form:

$$P = \begin{pmatrix} 1 & q_1 & 0 & \dots & 0 & 0 \\ 0 & r_1 & q_2 & \dots & 0 & 0 \\ 0 & p_1 & r_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & r_{N-1} & q_N \\ 0 & 0 & 0 & \dots & p_{N-1} & r_N \end{pmatrix} \quad (\text{B.4})$$

In general, so that P is a stochastic matrix, condition (B.2) must hold.

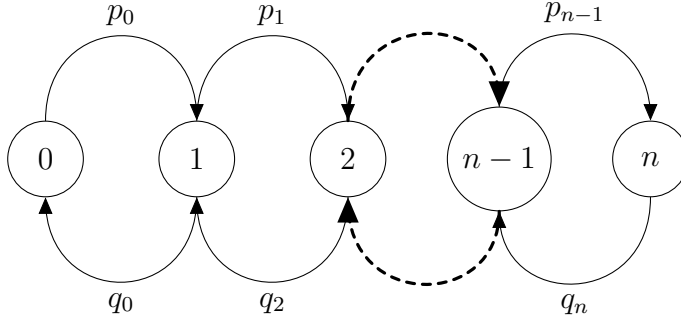


Figure B.1: *Transition graph of a birth and death process with finite state space on condition $r_i = 0$*

Time of First Passage

Let T_i denote the time for the first visit to state i , given $X_0 = 1$. Define the probability of first recurrence to i by the n^{th} step by

$$f_{ii}^{(n)} = P(X_n = i, X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i | X_0 = i) \quad (\text{B.5})$$

$$f_{ii}^{(n)} = P(T_i = n | X_0 = i) \quad (\text{B.6})$$

Further, let the probability of recurrence be

$$f_i = \sum_{n=1}^{\infty} f_{ii}^{(n)}, \quad (\text{B.7})$$

then state i is recurrent if $f_i = 1$ and transient if $f_i < 1$.

Theorem B.1

If $p_i > 0$ for $i = 0, 1, 2, \dots$ and $q_i > 0$ for $i = 1, 2, \dots$, then a birth and death process is transient iff

$$\sum_{j=1}^{\infty} \frac{q_j q_{j-1} \dots q_1}{p_j p_{j-1} \dots p_1} < \infty \quad (\text{B.8})$$

Proof

Let α_n denote the probability that a birth and death process starting at state $n \in \{0, 1, 2, \dots\}$ ever returns to state 0, then we have

$$\begin{aligned}\alpha_n &= P\{X_i = 0 \text{ for some } i \geq 1 | X_0 = n\} \\ \alpha_n &= \sum_k P\{X_i = 0 \text{ for some } i \geq 1 | x_1 = k\} P\{X_1 = k | X_0 = n\} \\ \alpha_n &= p_n \alpha_{n+1} + q_n \alpha_{n-1} + (1 - p_n - q_n) \alpha_n\end{aligned}\tag{B.9}$$

which yields the relation as follows:

$$(p_n + q_n) \alpha_n = p_n \alpha_{n+1} + q_n \alpha_{n-1}\tag{B.10}$$

Then we have the inductive relation:

$$\alpha_n - \alpha_{n+1} = \frac{q_n}{p_n} (\alpha_{n-1} - \alpha_n)\tag{B.11}$$

Iterating equation (B.5) yields:

$$\alpha_n - \alpha_{n+1} = \frac{q_1 \dots q_n}{p_1 \dots p_k} + 1\tag{B.12}$$

Finally, we have

$$\alpha_{n+1} = (\alpha_n - 1) \sum_{k=1}^{n \rightarrow \infty} \frac{q_1 \dots q_k}{p_1 \dots p_k} + 1\tag{B.13}$$

If the following term converges:

$$\sum_{k=1}^n \frac{q_1 \dots q_k}{p_1 \dots p_k} < \infty\tag{B.14}$$

Then, we let

$$\alpha_1 = \sum_{k=1}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k} \bigg/ \sum_{k=0}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k}\tag{B.15}$$

So that

$$\alpha_{n+1} = \frac{1}{\sum_{k=0}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k}} \sum_{k=1}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k} \rightarrow 0 \quad \text{as } n \rightarrow \infty\tag{B.16}$$

□

Expected Time to Extinction

Let τ_k denote the expected time until extinction for a population with initial size k . Then $\tau_0 = 0$ and the following relationship holds for τ_k , $k = 1, 2, \dots$:

$$\tau_k = p_k(1 + \tau_{k+1}) + q_k(1 + \tau_{k-1}) + (1 - (p_k + q_k))(1 + \tau_k) \quad (\text{B.17})$$

If the maximal population size is finite, then for $k = N$,

$$\tau_N = q_N(1 + \tau_{N-1}) + (1 - q_N)(1 + \tau_N). \quad (\text{B.18})$$

The difference equation can be simplified as follows:

$$q_k \tau_{k-1} - (p_k + q_k) \tau_k + p_k \tau_{k+1} = -1 \quad (\text{B.19})$$

for $k = 1, 2, \dots$

Theorem B.2

Suppose $\{X_n\}$ is a general birth and death process with $X_0 = m \geq 1$ satisfying $p_0 = 0 = q_0$, $p_i > 0$ for $i = 1, 2, \dots, N-1$ and $q_i > 0$ for $i = 1, 2, \dots, N$. The expected time until population extinction is

$$\tau_m = \begin{cases} \frac{1}{q_1} + \sum_{i=2}^N \frac{p_1 \dots p_i}{q_1 \dots q_i} & \text{if } m = 1 \\ \tau_1 + \sum_{s=1}^{m-1} \left[\frac{q_1 \dots q_s}{p_1 \dots p_s} \sum_{i=s+1}^N \frac{p_1 \dots p_{i-1}}{q_1 \dots q_i} \right] & \text{if } m = 2, 3, \dots, N \end{cases} \quad (\text{B.20})$$

Proof

For $k = 1, 2, \dots, N-1$, then $q_k \tau_{k-1} - (p_k + q_k) \tau_k + p_k \tau_{k+1} = -1$ can be solved recursively for $\tau_2, \dots, \tau_N$ to obtain the formulas

$$\tau_m = \tau_1 + \sum_{k=1}^{m-1} \frac{q_1 \dots q_k}{p_1 \dots p_k} \left[\tau_1 - \frac{1}{q_1} - \sum_{i=2}^k \frac{p_1 \dots p_{i-1}}{q_1 \dots q_i} \right] \quad (\text{B.21})$$

for $m = 2, \dots, N$. The second summation is zero when $k < 2$, then applying the following relation (2.6) when $k = N$:

$$q_N \tau_{N-1} - q_N \tau_N = -1 \quad (\text{B.22})$$

we get

$$\tau_N = \frac{1}{q_N} + \tau_{N-1} \quad (\text{B.23})$$

and equating the two values of τ_N , the following formula for τ_1 is obtained:

$$\tau_1 = \frac{1}{q_1} + \sum_{i=2}^N \frac{p_1 \dots p_{i-1}}{q_1 \dots q_i} \quad (\text{B.24})$$

Substituting τ_1 into (3.7) above, we get the formula stated in the theorem. □

B.1.2 Continuous-Time Birth and Death Processes

Let $\Delta X(t)$ denote the change in state of a stochastic process, $\{X(t), t \geq 0\}$ from t to $t + \Delta t$.

$$\Delta X(t) = X(t + \Delta t) - X(t) \quad (\text{B.25})$$

Homogeneous Birth and Death Process

A continuous-time Markov chain $\{X(t) : t \in [0, \infty]\}$ may have either a finite or infinite space $\{0, 1, 2, \dots, N\}$ or $\{0, 1, 2, \dots\}$ with infinitesimal transition probabilities:

$$p_{i+j,i}(\Delta t) = P\{\Delta X(t) = j | X(t) = i\} \quad (\text{B.26})$$

$$p_{i+j,i}(\Delta t) = \begin{cases} \lambda_i \Delta t + o(\Delta t) & \text{if } j = 1 \\ \mu_i \Delta t + o(\Delta t) & \text{if } j = -1 \\ 1 - (\lambda_i + \mu_i) \Delta t + o(\Delta t) & \text{if } j = 0 \\ o(\Delta t) & \text{if } j \neq -1, 0, 1 \end{cases} \quad (\text{B.27})$$

is a homogeneous continuous-time birth and death process.

The forward Kolmogorov differential equations for $p_{ji}(t + \Delta t)$ can be derived directly (B.9).

$$\begin{aligned} p_{ji}(t) &= p_{j-1,i}(t)[\lambda_{j-1} \Delta t + o(\Delta t)] + p_{j+1,i}(t)[\mu_{j+1} \Delta t + o(\Delta t)] \\ &\quad + p_{ji}(t)[1 - (\lambda_i + \mu_i) \Delta t + o(\Delta t)] + \sum_{k \neq -1, 0, 1}^{\infty} p_{j+k,i}(t) o(\Delta t) \\ p_{ji}(t + \Delta t) &= p_{j-1,i}(t) \lambda_{j-1} \Delta t + p_{j+1,i}(t) \mu_{j+1} \Delta t + p_{ji}(t)[1 - (\lambda_j + \mu_j) \Delta t] + o(\Delta t) \end{aligned} \quad (\text{B.28})$$

which holds for all i and j in the state space with the exception of $j = 0$ and $j = N$ (if the population size is finite). If $j = 0$, then

$$p_{0i}(t + \Delta t) = p_{1,i}(t) \mu_1 \Delta t + p_{0i}(t)[1 - \lambda_0 \Delta t] + o(\Delta t) \quad (\text{B.29})$$

If $j = N$ is the maximum population size, then

$$p_{Ni}(t + \Delta t) = p_{N-1,i}(t) \lambda_{N-1} \Delta t + p_{Ni}(t)[1 - \mu_N \Delta t] + o(\Delta t) \quad (\text{B.30})$$

where λ_N and $p_{kN}(t) = 0$ for all $k > N$. Subtracting $p_{ji}(t)$, $p_{0i}(t)$ and $p_{Ni}(t)$ from the preceding three equations respectively, dividing by Δt and taking the limit as $\Delta t \rightarrow 0$, yields the forward Kolmogorov differential equations for the general birth and death process,

$$p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \quad (\text{B.31})$$

$$p'_j(t) = \lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), \quad j = 1, 2, \dots \quad (\text{B.32})$$

$$p'_N(t) = \lambda_{N-1} p_{N-1}(t) - \mu_N p_N(t) \quad (\text{B.33})$$

To further determine state probabilities, we use their respective probability generating functions.

$$M(t, z) = \sum_{i=0}^{\infty} p_i(t) z^i \quad (\text{B.34})$$

under the initial conditions of type

$$p_n(0) = P(X(0) = n) = 1 \quad (\text{B.35})$$

In terms of the PGF, this condition is equivalent to

$$M(0, z) = z^n, \quad n = 0, 1, \dots \quad (\text{B.36})$$

Furthermore, partial derivatives of the z -transform will be needed:

$$\frac{\partial M(t, z)}{\partial t} = \sum_{i=0}^{\infty} p'_i(t) z^i \quad (\text{B.37})$$

and

$$\frac{\partial M(t, z)}{\partial z} = \sum_{i=1}^{\infty} i p_i(t) z^{i-1} \quad (\text{B.38})$$

Linear Birth and Death Process

$\{X(t), t \geq 0\}$ is called a linear birth and death process if it has transition rates:

$$\lambda_i = i\lambda \quad \mu_i = i\mu \quad (\text{B.39})$$

for $i = 0, 1, \dots$

In what follows, this process is analysed on the condition that

$$p_1(0) = P(X(0) = 1) = 1 \quad (\text{B.40})$$

The system of differential equations becomes

$$\begin{aligned} p'_0(t) &= \mu p_1(t), \\ p'_j(t) &= (j-1)\lambda p_{j-1}(t) - j(\lambda + \mu)p_j(t) + (j+1)\mu p_{j+1}(t); \quad j = 1, 2, \dots \end{aligned} \quad (\text{B.41})$$

Multiplying the j th differential equation by z^j and summing from $j = 0$ to $j = \infty$, taking the partial derivatives of the PGF into account yields the following linear homogeneous partial differential equation in $M(t, z)$:

$$\frac{\partial M(t, z)}{\partial t} - (z-1)(\lambda z - \mu) \frac{\partial M(t, z)}{\partial z} = 0 \quad (\text{B.42})$$

The corresponding characteristic differential equation is called the Riccati differential equation with constant coefficients:

$$\frac{dz}{dt} = -(z-1)(\lambda z - \mu) = -\lambda z^2 + (\lambda + \mu)z - \mu \quad (\text{B.43})$$

a) **The death rate and birth rate is not equal** ($\lambda \neq \mu$)

By the separation of variables, the above can be written in the form

$$\frac{dz}{(z-1)(\lambda z - \mu)} = -dt \quad (\text{B.44})$$

Integrating both sides of this relationship yields:

$$-\frac{1}{\lambda - \mu} \ln \left(\frac{\lambda z - \mu}{z - 1} \right) = -t + C \quad (\text{B.45})$$

The general solution $z = z(t)$ of the characteristic differential equation in implicit form is therefore given by

$$c = (\lambda - \mu)t - \ln \left(\frac{\lambda z - \mu}{z - 1} \right) \quad (\text{B.46})$$

where c is an arbitrary constant. Thus the general solution $M(t, z)$ has structure

$$M(0, z) = f \left((\lambda - \mu)t - \ln \left(\frac{z - 1}{\lambda z - \mu} \right) \right) = z, \quad (\text{B.47})$$

f must have structure

$$f(x) = \frac{\mu e^x - 1}{\lambda e^x - 1} \quad (\text{B.48})$$

Thus $M(t, z)$ is

$$M(t, z) = \frac{\mu \exp \left\{ (\lambda - \mu)t - \ln \left(\frac{\lambda z - \mu}{z - 1} \right) \right\} - 1}{\lambda \exp \left\{ (\lambda - \mu)t - \ln \left(\frac{\lambda z - \mu}{z - 1} \right) \right\} - 1} \quad (\text{B.49})$$

After simplification, $M(t, z)$ becomes

$$M(t, z) = \frac{\mu[1 - e^{(\lambda - \mu)t}] - [\lambda - \mu e^{(\lambda - \mu)t}]z}{[\mu - \lambda e^{(\lambda - \mu)t}] - \lambda[1 - \mu e^{(\lambda - \mu)t}]z} \quad (\text{B.50})$$

This representation of $M(t, z)$ allows its expansion as a power series in z . The coefficient of z^j is the desired state probability of $p_j(t)$. Letting $\rho = \frac{\lambda}{\mu}$ yields:

$$p_0(t) = \frac{1 - e^{(\lambda - \mu)t}}{1 - \rho e^{(\lambda - \mu)t}} \quad (\text{B.51})$$

$$p_j(t) = (1 - \rho) \rho^{j-1} \frac{[1 - e^{(\lambda - \mu)t}]^{j-1}}{[1 - \rho e^{(\lambda - \mu)t}]^{j+1}} e^{(\lambda - \mu)t}, \quad j = 1, 2, \dots \quad (\text{B.52})$$

Since state 0 is absorbing, $p_0(t)$ is the probability that the population is extinct at time t

$$\lim_{t \rightarrow \infty} p_0(t) = \begin{cases} 1 & \text{for } \lambda < \mu \\ \frac{\mu}{\lambda} & \text{for } \lambda > \mu \end{cases} \quad (\text{B.53})$$

Thus, in the cases where $\lambda < \mu$, the distribution function of the lifetime L of the population is

$$P(L \leq t) = p_0(t) = \frac{1 - e^{(\lambda - \mu)t}}{1 - \rho e^{(\lambda - \mu)t}}, t \geq 0 \quad (\text{B.54})$$

Hence, the population will survive interval $[0, t]$ with probability

$$P(L > t) = 1 - p_0(t) \quad (\text{B.55})$$

From this, we get

$$E[L] = \frac{1}{\mu - \lambda} \ln \left(2 - \frac{\lambda}{\mu} \right) \quad (\text{B.56})$$

The trend function $m(t) = E[X(t)]$ is principally given by

$$m(t) = \sum_{j=0}^{\infty} j p_j(t) \quad (\text{B.57})$$

which can also be obtained by the PGF:

$$m(t) = \left. \frac{\partial M(t, z)}{\partial z} \right|_{z=1} \quad (\text{B.58})$$

Multiplying the j th differential equation of the process by j and summing from $j = 0$ to ∞ yields the following first-order differential equation

$$m'(t) = (\lambda - \mu)m(t) \quad (\text{B.59})$$

Taking into account the initial condition $p_1(0) = 1$, its solution is

$$m(t) = e^{(\lambda - \mu)t} \quad (\text{B.60})$$

Multiplying the j th differential equation by j^2 and summing from $j = 0$ to ∞ , a second order differential equation in $Var(X(t))$ is obtained. The solution is

$$Var(X(t)) = \frac{\lambda + \mu}{\lambda - \mu} [1 - e^{-(\lambda - \mu)t}] e^{2(\lambda - \mu)t} \quad (\text{B.61})$$

b) **The birth rate and death rate are equal ($\lambda = \mu$)**

In this case the characteristic differential equation simplifies to

$$\frac{dz}{\lambda(z - 1)^2} = -dt \quad (\text{B.62})$$

Integration yields

$$c = \lambda t - \frac{1}{z - 1} \quad (\text{B.63})$$

where c is an arbitrary constant. Therefore $M(t, z)$ has structure:

$$M(t, z) = f\left(\lambda t - \frac{1}{z - 1}\right) \quad (\text{B.64})$$

where f is a continuously differentiable function. Since $p_1(0) = 1$, f satisfies

$$f\left(-\frac{1}{z - 1}\right) = z \quad (\text{B.65})$$

Hence, the function f is given by

$$f(x) = 1 - \frac{1}{x}, \quad x \neq 0 \quad (\text{B.66})$$

The corresponding PGF is

$$M(t, z) = \frac{\lambda t + (1 - \lambda t)z}{1 + \lambda t - \lambda t z} \quad (\text{B.67})$$

Expanding $M(t, z)$ as a power series in z yields the absolute state probabilities:

$$\begin{aligned} p_0(t) &= \frac{\lambda t}{1 + \lambda t}, \\ p_j(t) &= \frac{(\lambda t)^{j-1}}{(1 + \lambda t)^{j+1}}, \end{aligned} \quad (\text{B.68})$$

for $j = 1, 2, \dots, t \geq 0$.

An equivalent form of the state probabilities is

$$\begin{aligned} p_0(t) &= \frac{\lambda t}{1 + \lambda t}, \\ p_j(t) &= [1 - p_0(t)]^2 [p_0(t)]^{j-1}, \end{aligned} \quad (\text{B.69})$$

for $j = 1, 2, \dots, t \geq 0$.

Mean value and variance of $X(t)$ are

$$E[X(t)] = 1 \quad (\text{B.70})$$

$$\text{Var}(X(t)) = 2\lambda t \quad (\text{B.71})$$

Stationary Distribution

If $\{\pi_i, i \in Z\}$ is a stationary distribution of the Markov chain $\{X(t), t \geq 0\}$, then this special absolute distribution must satisfy the Kolmogorov equations in (2.10). Since the π_i are constant, all the left hand sides of the equations are equal to 0. Therefore, the system of the differential equations simplifies to a system of linear algebraic equations in the unknown π_i .

In the case of their existence the stationary distribution of a birth and death process satisfies the following system of equations:

$$\begin{aligned} \lambda_0 \pi_0 - \mu_1 \pi_1 &= 0 \\ \lambda_{i-1} \pi_{i-1} - (\lambda_i + \mu_i) \pi_i + \mu_{i+1} \pi_{i+1} &= 0, \quad \text{for } i = 1, 2, \dots \\ \lambda_{n-1} \pi_{n-1} - \mu_n \pi_n &= 0, \quad \text{for } n \leq \infty \end{aligned} \quad (\text{B.72})$$

Theorem B.3

Suppose the continuous-time Markov chain $\{X(t) : t \in [0, \infty)\}$ is a homogeneous birth and death process with infinitesimal transition probabilities given by the equations (2.1). If the state space is infinite, $\{0, 1, 2, \dots\}$, a unique

positive stationary probability distribution π exists iff

$$\mu_i > 0 \text{ and } \lambda_{i-1} > 0 \text{ for } i = 0, 1, 2, \dots \quad (\text{B.73})$$

and

$$\sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i} < \infty \quad (\text{B.74})$$

The stationary probability distribution equals

$$\pi_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i} \pi_0, \quad i = 1, 2, \dots, \quad (\text{B.75})$$

and

$$\pi_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i}} \quad (\text{B.76})$$

If the state space is finite, $\{0, 1, 2, \dots, N\}$, then a unique positive stationary probability distribution π exist iff

$$\mu_i > 0 \text{ and } \lambda_{i-1} > 0 \text{ for } i = 0, 1, 2, \dots, N$$

The stationary probability is given by the above, where the summation goes from $i = 0, 1, 2, \dots, N$

Proof

The explicit equations for the stationary distribution are as (B.24). These equations can be recursively solved:

$$\pi_1 = \frac{\lambda_0}{\mu_1} \pi_0. \quad (\text{B.77})$$

Then

$$\mu_2 \pi_2 = (\lambda_1 + \mu_1) \pi_1 - \lambda_0 \pi_0 \quad (\text{B.78})$$

$$\mu_2 \pi_2 = \left[\frac{(\lambda_1 + \mu_1) \lambda_0}{\mu_1} - \lambda_0 \right] \pi_0 \quad (\text{B.79})$$

$$\pi_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} \pi_0 \quad (\text{B.80})$$

Applying the induction hypothesis, assume π_j is given by

$$\pi_i = \frac{\lambda_1 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i} \pi_0 \quad (\text{B.81})$$

for $j = 1, 2, \dots, i$.

Then

$$\mu_{i+1}\pi_{i+1} = (\lambda_i + \mu_i)\pi_i - \lambda_{i-1}\pi_{i-1} \quad (\text{B.82})$$

$$= \frac{\lambda_0\lambda_1\ldots\lambda_{i-1}}{\mu_1\mu_2\ldots\mu_{i-1}} \left[\frac{\lambda_i + \mu_i}{\mu_i} - 1 \right] \pi_0 \quad (\text{B.83})$$

$$\mu_{i+1}\pi_{i+1} = \frac{\lambda_0\lambda_1\ldots\lambda_{i-1}}{\mu_1\mu_2\ldots\mu_{i-1}} \left[\frac{\lambda_i + \mu_i}{\mu_i} - 1 \right] \pi_0 \quad (\text{B.84})$$

$$\pi_{i+1} = \frac{\lambda_0\lambda_1\ldots\lambda_i}{\mu_1\mu_2\ldots\mu_{i+1}}\pi_0 \quad (\text{B.85})$$

For the infinite case, we apply an additional constraint, $\sum_{i=0}^{\infty} \pi_i = 1$ or $\pi_0(1 + \sum_{i=1}^{\infty} \frac{\pi_i}{\pi_0}) = 1$. Then,

$$\pi_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \frac{\lambda_0\lambda_1\ldots\lambda_{i-1}}{\mu_1\mu_2\ldots\mu_i}}$$

A unique positive stationary distribution exists iff the conditions (B.25) and (B.26) are satisfied.

□

Appendix C

Codes for the Spatio-Stochastic Models

The codes for the models build can be found on the following Rstudio Cloud workspace: <https://rstudio.cloud/project/4212932>