

# **A spatial model with vaccinations for COVID-19 in South Africa**

**By**

**Claudia Josephina Dresselhaus**

**14285739**

**Supervisor: Inger Nicolette Fabris-Rotelli**

Submitted in partial fulfillment of the requirements for the degree

Masters of Science in Advanced Data Analytics

in the Faculty of Natural and Agricultural Science

Department of Statistics University of Pretoria, Pretoria

February, 2023



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

## Declaration

I, *Claudia, Josephina Dresselhaus*, declare that the mini-dissertation, which I hereby submit for the degree *MSc Advanced Data Analytics* at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

.....

Claudia Josephina Dresselhaus

Date: .....

## Acknowledgements

I thank God for his love, grace and the adventure of doing a Masters- a certificate for those who stand the test of perseverance under trial. Testing produces perseverance and perseverance makes us mature and complete.

Powerful people do not give up and persevere like long distance runners. Prof Inger Fabris-Rotelli - thank you for not giving up and thank you for finishing the race with me. Thank you for believing in me. The world prospers with the amount of care you give to us students.

Thanks to Carel, who was my cornerstone in times of trouble, when I did not know further. Thank you for the support, your problem-solving suggestions and time you invested to make this research possible.

Thank you to my mom, Prof Cama Brandt, whose upbringing thought me to think and learn. She explained why things are the way they are and not because she said so. She helped me choose the right goals to go for and supported me. All three of my her children have at least a Masters degree. I hope that I will be the fourth and last one to achieve this degree as well.

Many thanks to Dr Fritz Dresselhaus and Prof Kobie Kleynhans, who proof-read this mini-dissertation and helped me assemble a paper.

Thank you to my best friend, Deidré Bredenkamp, who helped me get through my Masters.

We acknowledge the contributions of the SEPIMOD research team, who assisted with design and implementation of the model.

We also thank Prof Bruce Medallo from the University of the Witwatersrand and the Gauteng Department of Health for great support in data access. The UP Postgraduate Masters Coursework Bursary of the University of Pretoria towards this research is hereby acknowledged.

We thank the NAS Research Ethics Committee that approved our ethics application with the unique ethics number NAS227/2022.

# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	The field of epidemiology . . . . .	9
1.2	The problem . . . . .	11
1.3	Aims and objectives of this research . . . . .	12
1.4	Structure of the Thesis . . . . .	12
<b>2</b>	<b>Literature Review</b>	<b>13</b>
2.1	Historical Perspective on the prevalence of Diseases and Pandemics . . . . .	13
2.2	Geospatial perspectives . . . . .	15
2.3	Vaccinations for COVID-19 . . . . .	18
2.4	Conclusion . . . . .	20
<b>3</b>	<b>Methodology</b>	<b>21</b>
3.1	The extended SEIRDV Model . . . . .	21
3.1.1	Vaccination and susceptible compartments . . . . .	23
3.1.1.1	Infection transition rates . . . . .	24
3.1.2	Exposed compartment . . . . .	24
3.1.2.1	Spatial weight matrices . . . . .	25
3.1.3	Asymptomatic and mild compartments . . . . .	25
3.1.4	Hospitalisation compartments . . . . .	26
3.1.5	Recovered compartment . . . . .	26
3.1.6	Death compartment . . . . .	26
3.2	System of differential equations . . . . .	27
3.3	Estimating a wave of infectious disease . . . . .	28
3.3.1	Wave Detection . . . . .	28
3.4	Sensitivity Analysis . . . . .	28
3.4.1	Facilitating reliability and validity of SEIRDV through the Sensitivity Analysis . . . . .	30
3.5	Concluding remarks . . . . .	31

<i>CONTENTS</i>	4
<b>4 Discussion and Results</b>	<b>32</b>
4.1 Data Gathering . . . . .	32
4.2 Gauteng and descriptive statistics . . . . .	33
4.3 Simulation study . . . . .	36
4.3.1 Estimating parameters . . . . .	36
4.3.1.1 Estimating starting values . . . . .	36
4.3.1.2 Homogeneous parameters . . . . .	38
4.3.1.3 Spatial parameters . . . . .	39
4.3.1.4 Spatial temporal parameters . . . . .	39
4.3.2 Results of the SEIRDV model simulation study . . . . .	45
4.4 Concluding remarks . . . . .	47
<b>5 Conclusion and further research</b>	<b>48</b>

# List of Figures

3.1	The extended spatial SEIRDV model indicating the movement of exposed spatial areas for $i = 1, 2, 3, \dots, n$ .	22
3.2	A correlation plot, showing positive correlation in a red scale, and negative correlation in a blue scale. Simply put, the correlation between the input and output values. The values were obtained from the runs of the MC sensitivity analysis. The input variables are the columns of this figure and the outcomes are the rows. The correlation is in ascending order, according to high positive or negative correlation value.	30
4.1	A map depicting the Gauteng districts	34
4.2	A heatmap, that visualises the approximate proportion of fully vaccinated individuals for each district in Gauteng (February 21, 2020 – October 19, 2022).	35
4.3	Five subplots for each district of Gauteng. Each subplot shows the daily number of tested positive cases (displayed by bars) together with the 7 day moving average (represented by a solid line). Every subplot also shows dashed lines with relating start dates of the COVID-19 waves.	35
4.4	A pie chart illustrating the numerical proportion of Pfizer and Johnson & Johnson dosages in South Africa.	40
4.5	Boxplots of the vaccination rates $v_i$ grouped by $i$ districts and coloured in 5 waves	41
4.6	Boxplots of the General Ward Recovery Rate $\gamma_{3i}$ grouped by districts and categorised in 5 waves	42
4.7	Boxplots of the ICU Recovery Rate $\gamma_{4i}$ grouped by $i$ districts and coloured in 5 waves	43
4.8	Boxplots of the GW death rate $\delta_{3i}$ grouped by $i$ districts and coloured in 5 waves	43
4.9	Boxplots of the ICU death rate $\delta_{4i}$ grouped by $i$ districts and coloured in 5 waves	44
4.10	The movement range maps data (March 1, 2020 to May 31, 2022) relative to a baseline calculated in the month of February 2020.	45
4.11	Actual mild cases (solid line) vs the estimated mild cases in a (dashed line), and its confidence bounds.	46
4.12	Actual hospitalised cases (solid line) vs the estimated hospitalised cases in a dashed line, with confidence bounds. The hospitalised cases is a sum of general ward cases and ICU cases.	46

# List of Tables

2.1	Government intervention dates [66]	17
3.1	Description of the different compartments used in the SEIRDV model.	22
4.1	Descriptive Statistics to compare the districts according to the total population numbers, the total population proportion of Gauteng, the total proportion of fully vaccinated individuals and the total proportion of tested positive cases over all time.	34
4.2	The estimated parameters of each district of each wave.	37
4.3	Government intervention dates with the associated scaling factor $\kappa$ [66].	38
4.4	Parameters that can be estimated using their associated equations and data.	40

# List of Abbreviations

**COVID-19** Coronavirus disease 2019.

**HIV/AIDS** Human Immunodeficiency Virus/ Acquired Immunodeficiency Syndrome.

**ICU** Intensive Care Unit.

**MCMC** Markov Chain Monte Carlo.

**mRNA** Messenger Ribonucleic Acid.

**NICD** The National Institute For Communicable Diseases.

**SARS-COV** Severe acute respiratory syndrome coronavirus.

**SEIR** Susceptible, Exposed, Infected and Recovered compartments.

**SEIRDV** Susceptible, Exposed, Infected, Recovered, Diseased and Vaccinated compartments.

**SIR** Susceptible, Infected and Recovered compartments.

**WHO** World Health Organisation.



# Abstract

Rarely has the world undertaken a public health effort equal in scale or scope to the one it faced in response to the COVID-19 pandemic. Countries around the world implemented government interventions such as lockdowns and national vaccination campaigns to gain control of the COVID-19 pandemic that tore across the globe in 2020. Despite best effort and determination, thousands within the global population continued to suffer and die from COVID-19 day-after day. Nevertheless, much was learned about designing mass vaccination plans and implementing mass vaccination roll-outs throughout the world. When analysing cause and effect of the pandemic and when proposing intervention and prevention mechanisms to counter the pandemic, analysts in the health sector often apply mathematical models. Within the context described above, the main objective is to improve on the previously published spatial SEIR model for South Africa by including a compartment for spatial vaccination. The study further aims to assess validity, reliability and accuracy of the new model, given a socially heterogeneous and mobile population. The conclusion of this study is that the proposed model shows promising results in predicting the number of cases as well as the peak point and longevity of the wave. The study further concludes that factors such as immunity, lockdown levels, infectiousness and virulence are the main drivers of the spread of COVID-19.

# Chapter 1

## Introduction

Rampant diseases have always been part of human history [13, 46]. The Black Plague, which has spread via the Silk Route from China throughout Europe, caused the death of one-third of the European population between 1346 and 1350 [13, 65]. The plague recurred regularly for more than 300 years, of which notably the Great Black Plague of London took more than 15% of its population in 1665-1666, followed by the Plague of Milan, and years later the Plague of Marseilles in France. It then gradually disappeared from Europe.

The Silk Route was not the only trade route that brought diseases such as the Black Plague from other countries. As the Dutch East India Company rose to account for half of the world's trade during the early 1800s, their ships made it easier for diseases such as cholera, measles and smallpox to cause devastating pandemics during these times [118]. All these disease outbreaks have been recorded as devastating historical events [46, 83]. They have led to death, societal unrest, and economic disruption [65]. It is therefore of utmost importance to enquire in more depth on how to respond to such devastating outbreaks, inter alia: how an outbreak may progress and how to reduce its effects. To assist in these efforts, epidemiological mathematical models serve as valuable resources [51].

### 1.1 The field of epidemiology

The study of epidemics, general disease and the analysis of health outcomes in a population is called epidemiology [54]. Epidemiology coins terms such as epidemic, endemic, outbreak and pandemic [13, 54]. The term epidemic describes the occurrence of a health condition compared to its predicted rate as well as to its spread in spatial areas [13]. An outbreak occurs, when there has been an unprecedented increase in the case numbers of a disease or if cases are reported in a new spatial area [54]. An epidemic is an outbreak that spreads to larger spatial areas [65]. A pandemic is an epidemic that has spread globally [118]. The spread of diseases can cause many deaths before disappearing, sometimes diseases recur years later, but generally become less severe as populations develop some immunity [13, 65]. The term endemic describes a disease that is spreading among a population at a predictable rate. Diseases that become endemic (always present) are primarily harmful [73].

Diseases such as malaria, HIV/AIDS, cholera and tuberculosis are endemic in many parts of the world [98]. These diseases have considerable effects on mean life span as well as a social and economical burden to more than one billion people [90].

Mathematical models have been used to get a better understanding of how an infectious disease spreads [30]. The book in [30] is called "Epidemic Modelling: An introduction". This book introduces the most common approaches to modelling epidemics. The types of models this book illustrates are empirical methods [53], deterministic models [74] and stochastic models [37]. Since the emergence of the novel COVID-19 virus pandemic in December 2019, numerous mathematical models were published to assess the transmission dynamics of the disease, predict its future course and analyse the effect of different control measures. The reviews [51, 78, 132] report that the Susceptible-Exposed-Infected-Recovered compartmental model is a popular epidemiological model. The majority of modellers used compartmental models (SEIR-type models) for predicting the spread of a disease. These models are well-known for being able to take the dynamics of the spread into account, by utilising a set of differential equations [78]. These differential equations can be adapted to include numerous parameter estimates affecting spread like vaccinations, hospitalisation and interventions [51]. If the data availability is scarce and the intervention planning requires a great deal of detail, SEIR-type models become favourable [1, 148]. This is because they are good at predicting worst-case scenarios and can estimate the impact of interventions with a combination of expert advice and limited data [132]. Examples of researchers that modelled the potential impact of lockdown, vaccinations or social distancing on the COVID-19 epidemic in South Africa with a SEIR model are in [101, 111].

Many of these SEIR models are based on trying to understand how individuals move between the compartments and at what speed they move from one compartment to the next: individuals are either susceptible (S), have been exposed (E); have become infected (I) and then either recover (R) or die [1]. The basic reproduction number ( $R_0$ ) is of particular importance, as  $R_0$  estimates the speed at which a disease is capable of spreading in a population [21, 31, 88].  $R_0$  is the average number of secondary infections a single infectious individual causes [2]. For SARS-COV-2, the  $R_0$  value for South Africa was estimated to be between 1.7 and 2.5, according to the NICD [24]. This means that one infected person will on average infect 1.7 to 2.5 additional people. On the contrary, the effective reproduction number  $R_e$  is  $R_0$  compromised under the influence of immunisation, vulnerability and protection measures [86]. To complicate matters further, new mutations and variants of the virus have appeared that vary in infectiousness. A substantial increase in number of mild cases and hospitalisation caused by the delta variant proves that variants differ in  $R_0$ . An improved mathematical model is required to take this factor into account. NICD reports in 2022 show that South Africa intervention to stop the spread of COVID-19 manages to keep their effective reproduction  $R_e$  to approximately 1 [25].

## 1.2 The problem

The simplest SEIR models make the basic assumption that everyone has equal chances of catching the virus from an infected person, because the population is homogeneous [1]. A population with an equal social structure, in which individuals uniformly mix, is said to be homogeneous [34]. This assumption becomes problematic for diverse countries, that aggregate heterogeneous COVID-19 outbreaks in local areas [69]. Local predictions are therefore necessary, to adequately assess the spatial-temporal clustering pattern of COVID-19 cases [64, 122] in a diverse country.

South Africa is a large and diverse country with deep-rooted inequalities causing stark differences in access to basic municipal services such as health care, running water, sanitation, housing and social amenities [68]. The historical spatial planning of the colonial and apartheid system has the most profound influence of the cities that exist in South Africa [68]. Apartheid's fragmented spatial planning was used as an instrument to place discriminated people in townships on the outskirts of cities [68]. The past policies of spatial segregation has left a legacy of poverty and inequality, which are reflected in poor communities located in former homelands and townships [143]. Considering South Africa's spatial segregation, it is unlikely that the COVID-19 epidemic will take the same course in all areas of South Africa, even when important factors such as age distribution are considered [1]. To understand better the risks COVID-19 poses to vulnerable communities, the National Disaster Management Centre together with the Council of Scientific and Industrial Research, have built the COVID-19 Vulnerability Dashboard<sup>1</sup>. The dashboard spatially maps vulnerability to COVID-19 of areas of South Africa, thereby generating better insight into the challenge of protecting entire communities against COVID-19 [81].

Even though South Africa introduced a strict lockdown on onset of the COVID-19 epidemic, impoverished individuals kept on travelling from settlements into the cities during this time, due to many living day-to-day for food. The effect of former Apartheid policies and the ensuing formation of townships and consequent movement mainly with taxis from townships to suburbs and to the centre of towns will inevitably determine mobility and movement patterns [151, 152]. This mobility is reflected in the mobility data available for this research<sup>2</sup>. A hard lockdown was therefore not feasible and spatial interaction between individuals from different areas continued, while the fear of contracting COVID-19 was overridden by the basic need for food [120]. Considering the mobility rates in the model is therefore an important factor in South Africa, even under lockdown measures [66].

In [66], the heterogeneity in COVID-19 prevalence and its transmission rate in South Africa is higher than in other areas of South Africa and hence,  $R_0$  is heterogeneous spatially in South Africa [140]. To establish the variety of symptom levels and transmission rate in regions or areas, Thiede et al. in [140] used a spatial model to estimate the spatial spread of COVID-19. The model in [66] included vulnerability patterns in South Africa through a vulnerability index of areas that is based on its socioeconomic and health susceptibility characteristics. These indexes were produced by [81]. The spatial perspective was further enriched in [66] with mobility patterns be-

<sup>1</sup>Council of Scientific and Industrial Research. COVID-19 Vulnerability Dashboard, 2020. URL <https://bit.ly/3iFU4Zo>. (Accessed on 2020-07-09)

<sup>2</sup><https://data.humdata.org/dataset/movement-range-maps> (Accessed October 2022)

tween spatial areas. The mobility patterns between areas were calculated in [120]. One aspect that [66] did not address is the status of immunity in an area or region, following adjustments for vulnerability and mobility. Hence the current spatial model as described in [66] must be extended to include this aspect.

### 1.3 Aims and objectives of this research

The aims and objectives are the following:

1. The aim of this study is to extend the SEIR model in [66] by including the vaccination compartment.
2. The study should analyse the new model in terms of validity, reliability and accuracy.

### 1.4 Structure of the Thesis

Chapter 1 sets the context for the study by setting the background for the study and it continues to describe various problems that arise from the background. Research questions then ensue from set problems. Thereafter Chapter 1 defines the scope the study and the aims and objectives the study should reach.

Chapter 2 reviews previous literature published in the domain of COVID, especially the work done in the area of mathematical models. Chapter 2 continues to identify gaps in the literature and specific problems that remain and why they should be solved.

Chapter 3 reports on how the research was conducted with specific reference to the gathering of data, visual description of the data and statistical analysis of the data set. Chapter 3 then continues with an indication of trends, identification of any outliers and other significant observations, where applicable.

Chapter 4 recapitulates the problem statement and reflects on the objectives and research questions, which are mentioned in Chapter 1. Chapter 4 then continues to show the results of the studies and describes which research questions have not been achieved. It also reflects on the limitations of the study.

Chapter 5 summarises the findings of this work and gives recommendations for further work.

## Chapter 2

# Literature Review

The literature review in this chapter will take a historical perspective on the causes and responses to epidemics and pandemics. A historical analysis will help to identify key themes, principles, and concepts, which in turn will inform and provide the basis for this research. The emergent role of past mathematical models is of particular importance and will be dealt with in detail.

In addition, this literature review will also take into consideration present developments in the international and the South African context with the goal to reveal controversy and further research that is needed. The application of mathematical models in the modern context is an important tool in the research of epidemics and pandemics. Questions do arise on the validity and the reliability and the robustness of present day mathematical models - hence these must be investigated as well. Lastly the literature review will describe the further improvements that currently exist and map out a possible way forward.

### 2.1 Historical Perspective on the prevalence of Diseases and Pandemics

The study of infectious disease data began in the book called "Natural and Political Observations made upon the Bills of Mortality", authored by John Graunt in 1662 [27]. The book captured weekly records of death cases and cause of deaths in London between the years of 1592-1603 [134]. Graunt used the cases and causes of deaths to create so-called life tables, which entailed first approaches of extrapolating survival probabilities based on various causes including extrapolating the survival probabilities of the Black Plague in comparison to other diseases [53].

The first mathematical model used in epidemiology to model the life expectancy of vaccinated compared to non-vaccinated individuals was introduced by Daniel Bernoulli [10]. The Swiss mathematician published probability models in 1693, to predict the life expectancy of vaccinated compared to non-vaccinated babies during the "smallpox" or syphilis pandemic [10]. Heated discussions took place whether inoculation of a mild strain of smallpox would be beneficial [26]. Bernoulli's work, published in [10] outlined the competing risks and the increase in life expectancy if smallpox could be eliminated as a cause of death. The question of competing risks

followed in 1766 by a more complete exposition in [11]. This exposition had the intention to justify vaccination roll-outs to the general public and not only to the wealthiest social classes [26]. Bernoulli's description of age-specific prevalence of immune individuals was later revisited and then generalised by [33].

Another valuable contribution to understand infectious diseases is the study of the temporal and spatial pattern of cholera cases [116]. A book called "The Ghost Map" [70] describes the studies of temporal and spatial pattern, but lack of understanding about disease transmission process. The book talks about the 1855 Cholera epidemic in London, featuring John Snow, who pinpointed the Broad Street water pump as the source of the infection [136]. Similar studies of the temporal and spatial pattern of typhoid cases by William Budd in 1873 [16] and later in 1840 by William Farr [41] had the goal of discovering the laws that underlie the rise and fall of epidemics.

The knowledge about the disease transmission process by contact through living microorganisms as agents for disease was only inferred after the invention of the microscope [80]. Leeuwenhoek revealed, with the aid of a microscope, the concept of microorganisms that are invisible to the naked eye [20]. It was only in the early twentieth century that the definitions of immunity and germ theory and the relevant mathematics were established by researchers such as Robert Koch and Louis Pasteur [20]. This relatively new discovery highlights that little is known about microorganisms and even less regarding their applied mathematics.

Using the knowledge from germ theorists, Sidney Harmer [58] defined a mass action law for the rate of new infections, which is still the basic assumption in compartmental models since that time [145]. Harmer's mass action law in 1908, has mathematically shown that the spread of infection strictly depends on the number of susceptible species as well as the number of infected species called hosts [58].

A few years later, the British Nobel Prize winner and medical doctor Ronald Ross found a particularly instructive example of Harmer's mass action law [6]. Ross demonstrated in [126], that when killing enough mosquitos, such that the mosquito population falls below a critical level, is mathematically sufficient to eliminate malaria. Field trials supported this conclusion and led to some important successes in malaria control [91]. This work was the first theory of the basic reproduction number  $R_0$ , which has been a central idea in mathematical epidemiology since that time.

Similar basic compartmental models were later developed by Kermack and McKendrick in 1927 [74]. The basic compartmental models to describe the transmission of communicable diseases are formulated in a system of differential equations, found in a sequence of three papers [74–76]. The compartmental model was coined SIR, because it splits up the population into compartments namely Susceptible (S), Infected (I) and Recovered (R). Each compartment is associated with formulas that define the number of people in the compartment at a given time interval. The original mathematical formulas work in time intervals  $t$ , where infections only occur in the transition from one interval to the next and not during the time intervals themselves, thereby creating a discrete model. The SIR is a system of pre-defined mathematical functions and thus a deterministic model. Simple mathematical models described in [74] - [76] are used to highlight general behaviours or rough quantitative estimations [50]. The simplicity of the SIR model is accompanied by stark assumptions raised by Kermack and

McKendrick in their sequence of papers. One of the major shortcomings is the assumption of homogeneous mixing of the infected and susceptible populations [66]. The serious shortcomings caused by the assumptions of the SIR model, led researchers in either creating other disease models or adjusting the SIR model of Kermack and McKendrick [93, 105].

Erdos and Renyi introduced another way of disease modelling as a stochastic branching process in three consecutive articles [38–40]. Their abstract idea was to explain the beginning of a disease outbreak, by visioning the contact of individuals as a graph network. Individuals of the population are represented by vertices and the contacts between individuals are represented by edges. Graph theory has recently become more prominent in some areas, such as social media networks, cellular networks and the spread of communicable diseases [99, 128, 149]. Another frequently used epidemic model is the stochastic model of Reed and Frost [37]. The chain binomial model in [37] has been used widely [30, 55]. The Reed-Frost model and some extensions was described in the book called "Epidemic Modelling: An introduction" [30].

Other researchers continued to develop on the foundations of the SIR compartmental models and hence have grown the compartmental model research rapidly since the early publications of Kermack and McKendrick. A remarkable contribution of demonstrating extensions of these models are due to Hethcote that published a series papers in 1978 – 1989 [59–61]. Hethcote argued in [59] that children infected by the whooping cough can immediately be re-infected again after recovery. Therefore, Hethcote introduced a SIS model in [60]. Hethcote was also interested in overcoming the assumption of fixed population sizes by designing a SIR model with births and deaths. The SIR model with births and deaths in [61] shows improvements in accuracy. The series of papers [59] - [61] demonstrate that compartmental models for diseases can be customised according to properties as well as population size, to improve accuracy of the results.

From the descriptions above it can be concluded that mathematical models have come a long way and they fulfill an important role in the management and prevention of diseases. From all these models the SIR model has emerged as a prominent and recognised model. As such it will form an important part of this research.

## 2.2 Geospatial perspectives

The climate crisis in the 21<sup>st</sup> century, together with agricultural land use, livestock farming and urbanisation are changing the habitats of pathogens [135]. High contact between humans and animals through breeding, hunting as well as trade in exotic delicacies increases the risk of spreading zoonotic disease transmissions to humans. Example of zoonotic diseases is Influenza (flu), which is a respiratory infection in mammals and birds [7]. Often they are named after the animal that the virus originates from, such as bird flu outbreaks caused by chicken farms, swine flu viruses discovered at Mexican pig farms [87].

Other examples of viruses found in avian and mammalian species are severe acute respiratory tract infections (SARS), such as COVID-19 [144]. Researchers established the origin of the SARS corona virus being in bats sold at exotic wet markets in China during 2002, 2003 and later in 2019. Every year, Influenza and SARS viruses



cause 3 to 5 million cases of severe illness and approximately 500,000 deaths worldwide [47].

Since SARS epidemics emerged, a variety of extended SIR models have been designed to gain valuable insights into the transmission dynamics of infectious diseases. In order to improve the accuracy of the basic reproduction number, these models incorporate information such as age, differences in mobility rates and government interventions [32]. Since the spread of the disease is also influenced by the severity [110], Lowth and Jarvis in [77] proposed a model, which categorised the infected compartment into sub-categories called asymptomatic, mild, severe and critical, to incorporate the severity levels of a disease. Additionally, Putrino et al. [121] have proven that the spread of the virus is effected by the number of casualties it causes. For example, Ebola studies in [147] used SEIRD (where D denotes the death compartment) to demonstrate that the Ebola virus was too deadly to spread. Hence, studies such as [147] included a death compartment to take this effect into account, when extending the SIR model. They extended the model to SEIRD.

Furthermore, studies such as [139] have shown overwhelming evidence that vaccinations reduce severe and critical infections of COVID-19 and hence reduce the hospitalisation rate. This was reported in most countries world-wide [115]. Therefore, studies including [49, 95, 129], have extended the SEIRD model further by adding vaccinated compartment, denoted by V. Their proposed SEIRDV model has shown that all compartments have significant effects on the spread of the disease they were analysing.

Whilst SIR-type models have been powerful tools in predicting the spread of infectious diseases, they often struggle to predict the extent of local breakouts [156]. Communicable diseases naturally occur in clusters and outbreaks are more common in dense spaces [17, 34, 66]. Incorporating spatial details or realistic population mixing structures have led to more accurate transmission rate predictions [22]. It makes sense that a virus can spread faster in areas with high density or high levels of mobility [34, 45, 52, 155]. Examples of works that consider spatial clustering of disease outbreaks are included in the research on eradicating the foot and mouth disease in cattle. In [142] it is shown that vaccinating all cow herds that fall within a 10 kilometre radius from an infected cow had the best results in taking control of the virus. Often location, time and to whom the vaccination should be provided, were the dimensions used to evaluate vaccination allocation strategies [57, 115]. Ignoring these spatiotemporal dimensions may lead to an inefficient vaccination distribution strategy [156]. Therefore, information regarding the vaccination administrations in different areas were incorporated in [43, 146].

Further when considering the social interactions between individuals, human or cow, it is disadvantageous to assume that these interactions occur randomly. Gonzales et al. [52] argue that the use of simple SIR-type models are too simple to capture the complexities of human interactions, and would often incorrectly predict infection numbers. They further argue that social interactions are more prevalent between individuals in certain age groups. Interactions which occur between any two individuals in a population randomly is known as homogeneous mixing. The assumption of homogeneous mixing in a SIR-type model is hence a disadvantage [52].

To overcome the assumption of homogeneous mixing, Pooley et al. [119] incorporated age-related factors when modelling the spread of diseases. Alternatively, Zhou et al. [156] incorporated spatio-temporal informa-

tion, such as mobility predictions based on social media interactions and mobile phone networks. These measures prove to provide important information about the movements of individuals and the social interactions between them [4]. Work done by [96] use age and mobility as features in their model, but reason differently than the SIR model in [52]. They assume that the removed compartment is a mixture between recovered (R), vaccinated (V) or diseased (D) individuals (SEIRVD). Since vaccinated and recovered individuals are considered removed from the study, this approach assumes that vaccinated and recovered individuals will be permanently immune against the COVID-19 disease. This is however an invalid assumption, after studies in [79] reported that both recovered and vaccinated individuals have been reinfected during the COVID-19 pandemic. The study in [56] compared stochastic, age-dependent and spatial models when modelling the spread of measles. They concluded that adapting the model to be spatial heterogeneous, without increasing the dimensions of the system beyond tractability could lead to accurate disease modelling in medical research.

Several researchers have incorporated spatial heterogeneous information through taking the spatial density and mobility of the population into account [14]. For example, Carroll et al. [17] show that modelling spatial waves of Ebola infections can significantly improve models used for analysing the effects of vaccines. A SEIR approach to modelling the Ebola outbreak has also highlighted that infection counts cannot accurately be predicted without including spatio-temporal information [48]. Others such as Maidana [89] have implemented reaction-diffusion to incorporate the aviation transmission of mosquitoes in the spread of diseases such as Malaria.

Alternatively to reaction-diffusion equations, other works have used contact matrices to incorporate the spatial information of human interaction in SIR-type models [127]. Human contact drives the spread of communicable diseases [4]. Mobility of humans implies a greater exposure to the virus and this increases the probability of infections at an exponential rate [102]. For this reason, [102] explains that lockdown measures and track and trace efforts were implemented in many countries to monitor and curtail the spread of COVID-19. In South Africa, the government intervened by making use of lockdown levels<sup>1</sup>. The intervention dates are summarised in Table 2.1. This is supported in [101], that concludes that current social distancing measures to reduce contacts is key and must remain key in controlling the infection in the absence of vaccines and other therapeutics.

Level	Date
Hard lockdown - level	26 March - 30 April 2020
Level 4 lockdown (L4)	1 May - 30 May 2020
Level 3 lockdown (L3)	1 June - 17 August 2020
Level 2 lockdown (L2)	18 August - 20 September 2020
Level 1 lockdown (L1)	21 September - 29 December 2020
Lifting of National State of Disaster	4 April 2022

Table 2.1: Government intervention dates [66]

The "Spatial Model for COVID-19 in South Africa" [66] is a SEIR model, that includes spatial probability

<sup>1</sup><https://www.gov.za/COVID-19/resources/regulations-and-guidelines-coronavirus-COVID-19> (Accessed October 2022)

matrices calculated in [120]. The spatial weight matrices were used in the exposed compartment to imitate the movement of individuals from one geographical area to the next.

Next to spatial mobility matrices, le Roux et al. [81] have shown that vulnerability in South Africa differs across spatial areas in terms of socioeconomic and health susceptibility characteristics. Based on [81], Fabris-Rotelli et al. [66] incorporated vulnerability levels that prove to be significantly different across spatial areas.

With this approach, Fabris-Rotelli et al. discovered so-called "hot-spots", which are spatial clusters that have a significantly higher number of infected cases than other spatial areas. It is important to note that the model was defined at the beginning of the pandemic, when COVID-19 vaccinations were not yet available and not considered. Since COVID-19 vaccinations are available since February 2021, this study aims to add a vaccination compartment to the SEIR model. COVID-19 vaccinations information will help to formulate the vaccination compartment in this study.

## 2.3 Vaccinations for COVID-19

At the beginning of many COVID-19 vaccination programs, vaccinations were generally not evenly distributed. Strategies involving priority setting of vaccination varied from country to country [72]. However, most countries were in unity that vaccination should be administered according to age, comorbidity and frontline workers. Other countries vaccinated crucial members of society to keep the risk of hospital admission and mortality from COVID-19 in adults at minimum [23, 57, 94]. After vaccinating the cohort groups, most vaccination policies usually started vaccinating using age categories in decreasing order [44, 67]. However, the particulars of the vaccination strategies, such as vaccination types, priority setting, vaccine delivery (especially transport of mRNA vaccines, which proved to be a challenge), logistics, vaccine safety and side-effect monitoring differed from country to country [72]. For example, the first priority groups that were vaccinated in Germany and the United Kingdom were the elderly and their care workers, whilst the United States of America, South Africa and China started vaccinating the frontline health care workers [72]. China additionally prioritised vaccinations to staff of delivery and transportation companies [154]. Despite these different approaches, the following common factors and denominators are considered, there are a number of factors to consider, which we discuss next.

**Type of Vaccine** The World Health Organisation (WHO) listed 172 vaccines candidates in clinical development on the 21<sup>st</sup> of October 2022<sup>2</sup>. In the development of a COVID-19 vaccine, many manufacturers have primarily relied on new methods, such as the so-called mRNA vaccination or vector vaccines. They are among the first vaccines approved worldwide and are already in use. South Africa rolled out vaccines manufactured by Johnson & Johnson and Pfizer [137]. Depending on the type of COVID-19 vaccine, most individuals either need 1 or 2 dosages to be seen as fully vaccinated [106].

<sup>2</sup><https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines> (Accessed October 2022)

**Effectiveness of vaccines against variants of concern** A plethora of effectiveness figures were published by research groups that base their results on trial data and hence it is difficult to compare them [29, 57, 63, 79]. Pfizer BioNTech claim their vaccinations to be 95% effective after the second shot [114], Moderna 90% [71] and AstraZeneca 66,7 – 90% [109]. Johnson & Johnson states their vaccination to be 67% effective after the first shot [114]. All vaccine effectiveness must suffice throughout all three stages in the trials, and the vaccine manufactured according to international standards, which are also known as the the current Good Manufacturing Practice (cGMP) [18]. Vaccine efficiency on mild cases caused by the omicron variant reduced the efficiency of Pfizer to approximately 67% [103].

**Vaccination of vulnerable individuals** COVID-19 is known to affect different sub-groups within the population differently, where older people are more at risk than younger people, females more affected by hospitalisations than males, and those with comorbidity being more at risk of severe disease and death than those without comorbidity [5, 35]. South Africa exhibits a large degree of poverty and informal settlements located on fringes of cities [66]. These social-economic and geographic factors increase disease transmissions, often caused by lack of sanitation and densely packed travelling minibuses as transporting facilities [5, 67, 94, 152]. Relying on an under-sourced public health care system in South Africa and fighting additional epidemics such as TB and HIV, raises the need for shielding measures against the detrimental effects that the pandemic might have on the society and economy at large, specifically those most vulnerable [97, 124]. The distribution of vaccines was done taking the most vulnerable individuals of society into account.

**Distribution of Vaccines in South Africa** The vaccination rollout plan<sup>3</sup> was a three-phase approach that begins with the most vulnerable in South Africa's population.

1. **Phase 1** focused on frontline healthcare workers
2. **Phase 2** focused on vaccinating essential workers, individuals in congregate settings, are over 60 years or have comorbidities.
3. **Phase 3** focused on persons older than 18 years, targeting 22,500,000 of the population<sup>4</sup>

**Reinfections** Several studies have suggested that reinfections with COVID-19 are possible after recovering from COVID-19 [103] or after a vaccine-induced immunity [19].

**Stigma and communication challenges** The public health entities face notable challenges to address vaccine hesitancy in a compressed timeline [29, 85]. Religion, fake news, fear of needles, suspicion in vaccine safety and mistrust in the government are main deterrents for South Africans not to get vaccinated against COVID-19 [8, 29, 35, 85, 92, 146]. To ensure vaccine uptake, it is important to communicate positive emotions, enhance trustworthiness and increase the impact of the message [12, 35].

<sup>3</sup><https://www.gov.za/covid-19/vaccine> (Accessed October 2022)

<sup>4</sup><https://www.gov.za/covid-19/vaccine> (Accessed October 2022)

All features thus far have an influence on the vaccination policies, that are usually designed by vaccination and immunisation committees [67, 123, 130, 155]. The reviews in [5, 42, 85, 123] report on machine learning models, that estimate the infection risk of an individual, to assist policy makers to write vaccination policies. Most of these machine learning models did not take the location into account, even though other epidemiological models, such as [5, 34, 125] have proven that location is important. Epidemiological models in [34, 45] show that outbreaks naturally occur in clusters labelled as "hotspots" at locations. These "hotspots" highlight that the risk of infection in a specific location is higher than any other location. These "hotspots" are not often included in the risk models [107, 123], giving room for improvement. We address these weaknesses through the use of a mobility index [120] that indicates how much individuals move between spatial areas in our study.

## 2.4 Conclusion

From the literature review, it is clear that disease modelling is important and has been done for centuries. Unfortunately, micro-organisms such as bacteria and viruses were only discovered in the 20<sup>th</sup> century. This calls attention to the urgency to explore the little known viral and bacterial transmission dynamics and even more so its mathematics. The key models were originally mostly deterministic models due to lack of data and the SIR-type model remained to be the dominant model. Researchers later extended the SIR model to improve its validity and accuracy, but most did not include the geo-spatial characteristics in these models due to its taxing complexity. This gap is also highlighted in modern literature, where geo-spatial effects continue to be an important topic. In the next chapter we propose a spatial SIR-type model which incorporates the impact of vaccinations on the disease spread.

## Chapter 3

# Methodology

Chapter 2 demonstrated the emergence, importance and the continued development of mathematical models from the past and into the present [132]. It was further shown that the SIR-type models emerged as a prominent and generally accepted instrument to understand the spread of diseases, epidemics and pandemics. Furthermore, the SIR-type model has enabled the community of scientists to better understand the dynamics of disease transmission. However, it was shown that the SIR-type models do not adequately address geo-spatial factors, especially in heterogeneous and mobile populations, such as in South-Africa. Chapter 3 will therefore seek to describe the geo-spatial perspective in more detail within an extended SEIRDV model.

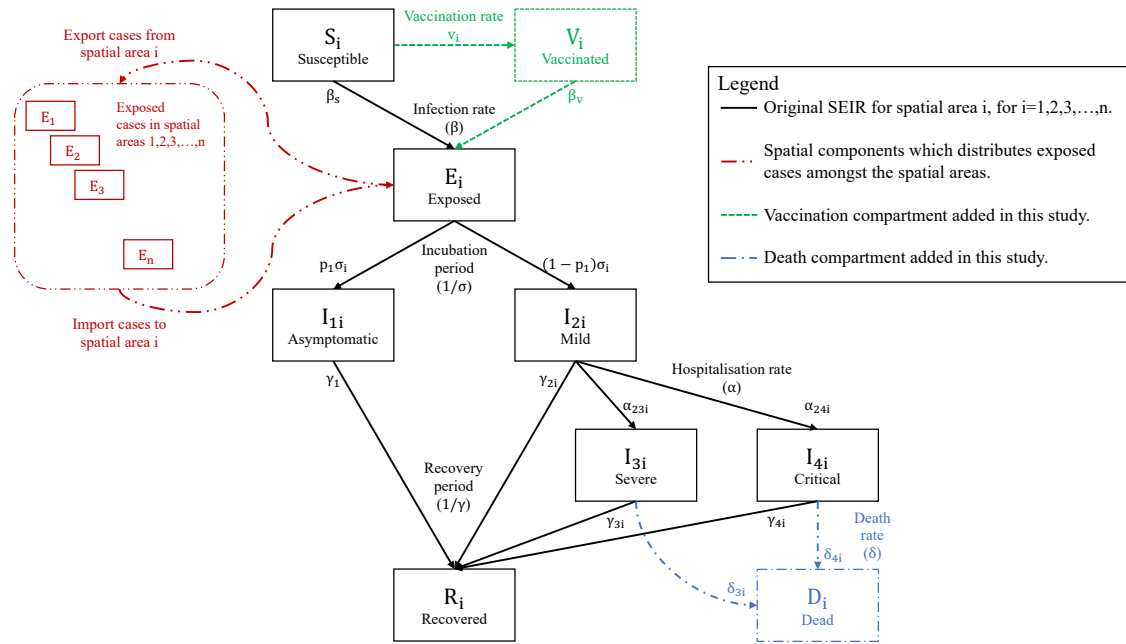
### 3.1 The extended SEIRDV Model

In this study we propose the spatial SEIRDV model, displayed in Figure 3.1. The proposed model consists of nine compartments called Susceptible, Vaccinated, Exposed, Asymptomatic, Mild, Severe, Critical, Recovered and Dead. Table 3.1 lists the abbreviated compartments with a corresponding short description. The course of the infection was modelled at a daily time-step with an SEIRDV compartment model to simulate the epidemiology of the disease. In other words, at each time step  $t$  for each spatial area  $i$ , the SEIRDV visualised in Figure 3.1 was calculated. The black and red dashed compartments are from the SEIRD model in [66]. The blue and green compartments are the compartments that we add in this research of [66].

Starting from the susceptible compartment, individuals in this compartment are susceptible to infection of COVID-19 and have no immunity against the disease. Over time, some individuals get vaccinated and move to the vaccination compartment at a vaccination rate  $\nu_i$ . The vaccinated individuals are less likely to become infected as they have some immunity. Thus they have a different infection rate,  $\beta_{vi}$ , from the susceptible infection rate,  $\beta_{si}$ . Over time, both susceptible and vaccinated individuals may encounter an infected individual, with a corresponding rate denoted as  $\beta_{si}$  and  $\beta_{vi}$ . Since they had close contact to an infected person they are now exposed.

Compartment	Description
$S_i$	Number of susceptible individuals in spatial area $i$
$V_i$	Number of vaccinated individuals in spatial area $i$
$E_i$	Number of exposed individuals in spatial area $i$
$I_{1i}$	Number of individuals with an asymptomatic infection in spatial area $i$
$I_{2i}$	Number of individuals with a mild infection in spatial area $i$
$I_{3i}$	Number of infected individuals admitted into a general ward (severe cases) in spatial area $i$
$I_{4i}$	Number of infected individuals in ICU (critical cases) in spatial area $i$
$R_i$	Number of recovered individuals in spatial area $i$
$D_i$	Number of dead individuals in spatial area $i$

Table 3.1: Description of the different compartments used in the SEIRDV model.

Figure 3.1: The extended spatial SEIRDV model indicating the movement of exposed spatial areas for  $i = 1, 2, 3, \dots, n$ .

The exposed people are still able to move from one spatial area to another and hence can transfer the virus. This is incorporated in the model by allowing exposed individuals to transition from one spatial area to another using the probabilities,  $w_{ij}$  for  $i, j \in 1, 2, \dots, n$ . The probabilities,  $w_{ij}$ , originate from the mobility matrices produced in [120]. The mobility of exposed cases are visualised as a red dotted compartment. The model only allows exposed individuals to either become asymptomatic or mild infected cases. Hence  $\beta_{si}$  and  $\beta_{vi}$  are also called infection rates, since the SEIRDV model only allows exposed individuals to transition to either to asymptomatic or to mild compartment, with transition rates  $p_1\sigma$  or  $(1 - p_1)\sigma$  respectively.

Hospitalisation is represented by severe and critical compartments and mild cases transition to either a gen-

eral ward (severe) or a intensive care unit (critical) with rates corresponding to  $\alpha_{23i}$  and  $\alpha_{24i}$ . These parameters also vary with  $i$ , to incorporate heterogeneity. The number of severe ward or critical patients that recover is set by the transition recovery rates  $\gamma_{i3}$  and  $\gamma_{i4}$ , also dependent on  $i$ . Similarly, the number of general ward or critical patients that die are determined by transition death rates  $\delta_{i3}$  and  $\delta_{i4}$ . Only severe or critical compartments allow for individuals to die. The mild cases that do not get admitted to hospital will recover after some time, which is determined by transition recovery rates  $\gamma_{i1}$  and  $\gamma_{i2}$ , also dependent on  $i$ .

Not all variables vary in each spatial area. Some effects are still homogeneous across all spatial areas. For example, the basic reproduction number  $R_0$ , the government lockdown levels  $\kappa$ , the incubation period  $\frac{1}{\sigma}$  and the proportion of asymptomatic cases  $p_1$  are the same for all individuals, regardless of the spatial area.

Altogether, the population numbers of each spatial area is denoted as  $N_i$ , we can write the equation as  $N_i = S_i + E_i + I_{1i} + I_{2i} + I_{3i} + I_{4i} + R_i + V_i + D_i$ , for spatial area  $i$ . Figure 3.1 provides an overview of the compartments of the SEIRDV and their relative transition rates. The next subsections explain every compartment of the model in detail.

### 3.1.1 Vaccination and susceptible compartments

The daily vaccination rate is indicated by  $v_i$ , that is written above the green solid line in Figure 3.1. This vaccination rate depends on  $i$ . Depending on the COVID-19 vaccine type, most individuals either need 1 or 2 dosages. In general, an individual on average receives about  $E[d]$  dosages and is defined in 3.1

$$E[d] = 1P[d = 1] + 2P[d = 2] \quad (3.1)$$

Using  $E[d]$ , we can estimate the expected number of individuals that are fully vaccinated per week for  $i$ . The estimated weekly number of fully vaccinated individuals in a spatial area  $i$  is given in Equation 3.2.

$$V_i^w = \frac{D_i^w}{E[d]} \quad (3.2)$$

To estimate the daily number of fully vaccinated individuals, we divide  $V_i^w$  by the total number of days in a week. This is shown in Equation 3.3

$$V_i = \frac{V_i^w}{7} \quad (3.3)$$

We estimate the daily vaccination rate, which is the transition rate from the susceptible to vaccinated compartment using Equation 3.4.

$$v_i = \frac{V_i}{S_i} \quad (3.4)$$

The rates of transition from the vaccinated or susceptible compartments to the exposed compartment are called the infection transition rates. These rates are discussed in more detail in the remainder of this subsection.



### 3.1.1.1 Infection transition rates

In [66], the transition rate from susceptible to exposed is calculated as the daily amount of infections  $R_{i0}$  divided by the infectious period  $(\gamma_{1i})^{-1}$ . The infectious period  $(\gamma_{1i})^{-1}$  is obtained from a  $\Gamma(2, \frac{7}{2})$  distribution [82, 104, 133]. The daily amount of infections  $R_{i0}$  can also be seen as the effective reproduction number, described by Equation 3.5.

$$R_{i0} = R_0 \kappa \psi_i, \text{ for } i \in \{1, 2, \dots, n\}. \quad (3.5)$$

The effective reproduction number ( $R_{i0}$ ) is  $R_0$  scaled by lockdown measures  $\kappa$  (described in Table 2.1) and vulnerability rates  $\psi_i$ . The basic reproduction number is assumed to have a  $\Gamma(57.2, 0.05)$  distribution with mean 2.86 [66]. The vulnerability rates incorporates the health susceptibility using data about age cohorts, presence of comorbidities and high poverty levels issued. The vulnerability rate  $\psi_i$  depends on  $i$  and is based on the weighted average of the normalised vulnerability indices described in [81] for a spatial area  $i$ .  $R_{i0} (\gamma_{1i})^{-1}$  are then used to formulate  $\beta_{si}$  and Equation 3.6.

$$\beta_{si} = \frac{R_{i0}}{(\gamma_{1i})^{-1}} = R_{i0} \gamma_{1i} \quad (3.6)$$

The transition rate  $\beta_{vi}$  is similar to  $\beta_{si}$  but takes the estimated vaccine efficiency  $g$  into account. In other words, the transition rate of vaccinated individuals to the exposed compartment symbolised by  $\beta_v$  is the  $\beta_s$  multiplied by  $g$ , to imitate the protection power that a vaccination has against COVID-19.

$$\beta_{vi} = g \beta_{si} \quad (3.7)$$

The vaccine efficiency is implemented in Equation 3.7 for  $\beta_v$ . The estimated  $g$  value is a simplified representation of the actual vaccine efficiency. The estimated  $g$  is the average efficiency of all vaccination types that were given to the population.

Since there is a possibility of re-infection [79], the SEIRDV model allows for both completely susceptible and vaccinated individuals to be exposed to and infected by the disease.

### 3.1.2 Exposed compartment

In this SEIRDV model, only the exposed compartment  $E$  incorporates the mobility matrices. This means that only exposed individuals move between districts and spread the disease. The SEIRDV model assumes that infected individuals quarantine themselves and therefore do not spread the disease. The SEIRDV model also does not consider movements of susceptible or vaccinated individuals, as they did not get exposed to the disease yet. Thus, only including the mobility matrices in the exposed compartment makes sense and was also done in [66]. We use probabilities of the mobility matrix to calculate the transition of people between spatial areas. The higher the mobility probabilities, the more exposed individuals would travel from one spatial area to the

next. As shown in [120], mobility probabilities are higher in major cities compared to the mobility probabilities in rural areas. This means that individuals in major cities encounter more exposed individuals than those living in rural settings. This model assumes that all exposed individuals will have some incubation period,  $\frac{1}{\sigma}$ , and once infected, have a constant infectious periods for each infected compartment,  $\gamma_{1i}$  for asymptomatic up to  $\gamma_{4i}$  for critical, respectively.

The spatial weight matrices that calculate the probability of individuals moving from one spatial area to the next is discussed in the next subsection.

### 3.1.2.1 Spatial weight matrices

The use of mobility data in SEIRDV is essential to capture the intrinsic spread through the population. Spatial weight matrices are useful, since they can represent mobility between spatial areas at a given time [9].

Potgieter et al. [120] discuss four methods to construct spatial weight matrices. The fourth method was chosen due to the data available only from Facebook<sup>1</sup>. This data was used to create a so-called spatial weight matrix. Simply put, the spatial weight matrix used is a  $n \times n$  matrix, which is constructed in such a way so that it quantifies  $w_{ij}$  as the amount of spatial influence that spatial area  $i$  exerts on spatial area  $j$ . The entries of the spatial weight matrix are given by Equation 3.8, where  $F_i^{(t)}$  is the mobility of spatial unit  $i$  at time  $t$  and  $d_{ij}$  is the standardised Euclidean distance between the centroids of spatial units  $i$  and  $j$ .

$$w_{ij}^{(t)} = (1 + F_i^{(t)})e^{-d_{ij}} \quad (3.8)$$

The  $n \times n$  mobility matrices are included in the exposed compartment, to simulate the spread of the virus [66]. This assumption is based on the COVID-19 regulations, that enforced a quarantine on individuals who tested positive against the COVID-19 virus.

Adding the mobility indexes to the exposed compartment should significantly cover the mobility of individuals in South Africa between spatial areas under various restrictions imposed by the government.

### 3.1.3 Asymptomatic and mild compartments

The transition rate from exposed to asymptomatic or mild cases are indicated by the black solid lines and  $\sigma$  symbols in Figure 3.1. The sole incubation period,  $\frac{1}{\sigma}$ , is the period of days between exposure to an infection and the appearance of the first symptoms [15]. The value is given in units of days and is obtained from a  $\Gamma(2, 1)$  distribution [117]. In [84, 108, 133], it was concluded that  $p_1 = 0.75$  of COVID infections have asymptomatic symptoms (compartment  $I_1$ ). This means that only 0.25 of the exposed individuals actually show mild symptoms of a COVID-19 infection. The transition rate of exposed to asymptomatic cases is estimated as  $p_1\sigma$ .

<sup>1</sup><https://data.humdata.org/dataset/movement-range-maps> (Accessed October 2022)

### 3.1.4 Hospitalisation compartments

The SEIRDV model assumes that severe or critical cases are either admitted to the general ward or to the Intensive Care Unit (ICU) of a hospital. Thus, the compartments  $I_{3i}$  and  $I_{4i}$  can also be seen as the general ward and ICU compartment respectively. The transition rate from the mild  $I_{2i}$  to the severe compartment is denoted as  $\alpha_{23i}$ . Similarly, the transition rate from the mild to the critical compartment is denoted as  $\alpha_{24i}$ .

### 3.1.5 Recovered compartment

The recovered compartment is denoted by  $R$ . The transition rates  $\gamma_1$ ,  $\gamma_{2i}$ ,  $\gamma_{3i}$  and  $\gamma_{4i}$  are transition rates from the asymptomatic, mild, severe or critical to the recovered compartment respectively, dependent on  $i$ .

The transition rate from asymptomatic to the recovered compartment is given by  $\gamma_1$ , which is independent of  $i$ . In [104, 133, 138], the recovery period is  $\frac{1}{\gamma_1}$  and is assumed to have a  $\Gamma(2, 1)$ .

The transition rate from the severely infected to recovered ( $\gamma_{3i}$ ) or from the critical to recovered ( $\gamma_{4i}$ ) for  $i$  can be estimated from the data set. Let  $I_{3i_{alive}}$  be the number of cases that were discharged alive from the general ward of a hospital in spatial area  $i$  on a given day. Then  $\gamma_{3i}$  can be calculated as

$$\gamma_{3i} = \frac{I_{3i_{alive}}}{I_{3i}} \quad (3.9)$$

Likewise, let  $I_{4i_{alive}}$  be the number of cases that were discharged alive from the ICU of a hospital in spatial area  $i$  on a given day. Then  $\gamma_{4i}$  can be calculated as

$$\gamma_{4i} = \frac{I_{4i_{alive}}}{I_{4i}} \quad (3.10)$$

The transition rate  $\gamma_{2i}$  is calculated as

$$\gamma_{2i} = 1 - \gamma_1 - \gamma_{3i} - \gamma_{4i} \quad (3.11)$$

The data set also reports death cases, which will be discussed in the next subsection.

### 3.1.6 Death compartment

For simplicity, the SEIRDV model only allows hospitalised individuals to die of COVID-19. The transition rates to the death compartment  $D_i$  are indicated by blue dashed lines in Figure 3.1. The transition rates  $\delta_{3i}$  and  $\delta_{4i}$  define the transition rate from hospitalised or ICU compartments to the death compartment,  $D$ , respectively.

Therefore,  $D_{3i_{dead}}$  be the number of cases that died in the general ward of a hospital in spatial area  $i$  on a given day. Then  $\delta_{3i}$  can be calculated as

$$\delta_{3i} = \frac{D_{3i_{dead}}}{D_{3i}} \quad (3.12)$$

Likewise, let  $D_{4i_{dead}}$  be the number of cases that died in the ICU of a hospital in spatial area  $i$  on a given day. Then  $\delta_{4i}$  can be calculated as

$$\delta_{4i} = \frac{D_{4i_{dead}}}{D_{4i}} \quad (3.13)$$

### 3.2 System of differential equations

The number of individuals in each compartment of the model is calculated for a timepoint  $t$  in a spatial area  $i$ , by using a set of differential equations. Simply put, the number of individuals that transitions into a compartment are additions in the associated differential equation. The number of individuals that transitions out of a compartment are removals in the associated differential equation. The system of differential equations of the model are given as follows:

$$\frac{dS_i}{dt} = -v_i S_i - \frac{\beta_s S_i (\rho I_{1i} + I_{2i} + I_{3i} + I_{4i})}{N_i} \quad (3.14)$$

$$\frac{dV_i}{dt} = v_i S_i - \frac{\beta_{v_i} S_i (\rho I_{1i} + I_{2i} + I_{3i} + I_{4i})}{N_i} \quad (3.15)$$

$$\frac{dE_i}{dt} = \frac{(\beta_v + \beta_s) S_i (\rho I_{1i} + I_{2i} + I_{3i} + I_{4i})}{N_i} - \sigma E_i \quad (3.16)$$

$$\frac{dI_{1i}}{dt} = p_1 \sigma E_i - \gamma_1 I_{1i} \quad (3.17)$$

$$\frac{dI_{2i}}{dt} = (1 - p_1) \sigma E_i - \gamma_2 I_{2i} - \alpha_{23i} I_{2i} - \alpha_{24i} I_{2i} \quad (3.18)$$

$$\frac{dI_{3i}}{dt} = \alpha_{23i} I_{2i} - \gamma_3 I_{3i} - \delta_{3i} I_{3i} \quad (3.19)$$

$$\frac{dI_{4i}}{dt} = \alpha_{24i} I_{2i} - \gamma_4 I_{4i} - \delta_{4i} I_{4i} \quad (3.20)$$

$$\frac{dR_i}{dt} = \gamma_1 I_{1i} + \gamma_2 I_{2i} + \gamma_3 I_{3i} + \gamma_4 I_{4i} \quad (3.21)$$

$$\frac{dD_i}{dt} = \delta_{3i} I_{3i} + \delta_{4i} I_{4i} \quad (3.22)$$

The system of differential equations models the course of the infection at a daily timestep  $t$ . We use the system of differential equations to model a wave of infectious disease for a spatial area  $i$ .

### 3.3 Estimating a wave of infectious disease

The following algorithm was used to model a COVID-19 wave in South Africa for each spatial area  $i$

---

**Algorithm 1** SEIRDV algorithm to model a COVID-19 wave

---

- 1: Where applicable, fit distributions to the input parameters of the SEIRDV model.
  - 2: Sample input parameters of the fitted distributions.
  - 3: Solve the system of linear equations using an ordinary differential equations solver [62].
  - 4: Repeat step 3 for  $T$  time units.
  - 5: Repeat step 2 – 4 for multiple seeds.
  - 6: Compute the mean and confidence interval of the model predictions.
- 

The SEIRDV model can estimate the number of individuals in a compartment for  $T$  time points, but cannot detect when the wave starts. The ability to detect or preempt an imminent wave of COVID-19 is an important task in any modelling or prediction initiative. With advanced warning, local and national governments can better prepare for and manage the challenges associated with these waves.

#### 3.3.1 Wave Detection

To be able to make accurate predictions, knowledge of the start of a wave can also lead to better outcomes in modelling the severity and duration of the waves. The detection of waves was carried out using the method proposed in [112]. In this method, generalized additive models (GAMs) are fitted to daily cases in an "add-one-in" fashion using the package "mgcv" in R [150]. An Early Warning Signal (EWS) is triggered if the standard deviation, the autocorrelation function (ACF) and the return-rate,  $r$ , exceed two standard deviations from their expanding (or cumulative) mean. The EWS detector identifies a disruption in the variability of the case numbers or in the prevailing pattern (autocorrelation) of daily cases numbers.

The EWS detector and the SEIRDV algorithm to model a COVID-19 wave are used to estimate the start date and reproduce the actual case numbers respectively.

### 3.4 Sensitivity Analysis

A construct validity test via a sensitivity analysis is done to test the robustness and validity of the SEIRDV model output. Recall the SEIRDV diagram illustrated in Figure 3.1. All parameters may influence the case number and longevity of stay of individuals in each compartment of the model. In societies, however, the number of casualties and hospitalisations a disease creates is what causes devastation in communities and families. Especially during the COVID-19 pandemic, government interventions, social distancing, vaccinations and mask-wearing were defined as the easiest and best prevention mechanisms to keep the cumulative sum of hospitalisation and

casualties as low as possible. Thus it makes sense to use these three compartments as output of interest for this sensitivity analysis.

Input parameters that drive infection, hospitalisation and casualties because of COVID-19 are the different factors responsible for its transmission. To analyse the extend to which input parameters effect hospitalisation and casualties, a method called the sensitivity analysis was conducted. Sensitivity analysis is a process that analyses how variation in the input variables affect the model outputs and highlights input variables that cause the most variation in the outputs.

A form of sensitivity analysis that avoids having to sample every combination of values for each parameter, while ensuring a roughly even distribution across the parameter space over time, is called probabilistic sensitivity analysis [28]. An example of a probabilistic sensitivity analysis is called the Monte Carlo (MC) sampling, which works as follows:

---

**Algorithm 2** MC algorithm to perform a probabilistic sensitivity analysis of the SEIRDV model.

---

- 1: Assign an assumed distribution for each input variable.
  - 2: Sample a value for each input variable using its assumed distribution.
  - 3: Enter the sampled values into the SEIRDV model
  - 4: Solve the system of linear equations using a ordinary differential equations solver [62].
  - 5: Repeat steps 1 – 4  $N$  times, using a new seed for each repeat.
- 

The advantage of using probabilistic sensitivity analysis is the flexibility of allowing us to steer how likely each parameter will take on a certain value, by choosing its underlying distribution. Results in the MC sampling analysis done by [100] and [153] show that sampling completely at random from a probability distribution for each parameter is inefficient and causes in-exhaustive coverage of the parameter space. This is due to homogeneous traits consequently inherited from sampling completely randomly in simple MC sampling. The methodology review for sensitivity analysis for disease models in [113] shows that the global sensitivity analysis (GSA) show more reliable results than the traditional sensitivity analysis. A GSA method in [131] proposes the so-called Latin hyper-cube sampling. The Latin hyper-cube sampling is a process that divides the population in  $N$  equal probability sections and then performs a simple MC sample on each probability section. The sensitivity analysis method chosen for this study is Monte Carlo (MC) sampling method, but different than usual MC method as we stratify this analysis across different spatial areas. This follows from knowing that the SEIRDV model is heterogeneous, and so the analysis should also take the different sub-population into account. Thus, a simple MC sample would also give sufficient insight about how different inputs of each parameter could result in a different output for death and hospitalised cases across different spatial areas. The different input parameters vary using pre-defined underlying distributions in each MC sample. These MC samples are used to run the model to simulate the maximum outcome of deaths and hospitalisations after a given period. All codes are available on a Github repository<sup>2</sup>.

---

<sup>2</sup>[https://github.com/ItsClaudiPie/SEIRDV\\_model](https://github.com/ItsClaudiPie/SEIRDV_model)

### 3.4.1 Facilitating reliability and validity of SEIRDV through the Sensitivity Analysis

The results in Figure 3.2 of the global sensitivity analysis show a correlation plot. In the correlation plot, we indicate the correlation between the parameter of the model and its outcome. Additionally to this, we overlay a heat-map to illustrate how strong the correlation is between the parameters of the model and the outcome. Warm colours such as red and orange indicate strong positive correlation, whereas cool colours such as dark blue colours indicate strong negative correlation. The input variables in the correlation plot are particular transition rates, whereas the outputs are the number of mild, hospitalised and deaths respectively.

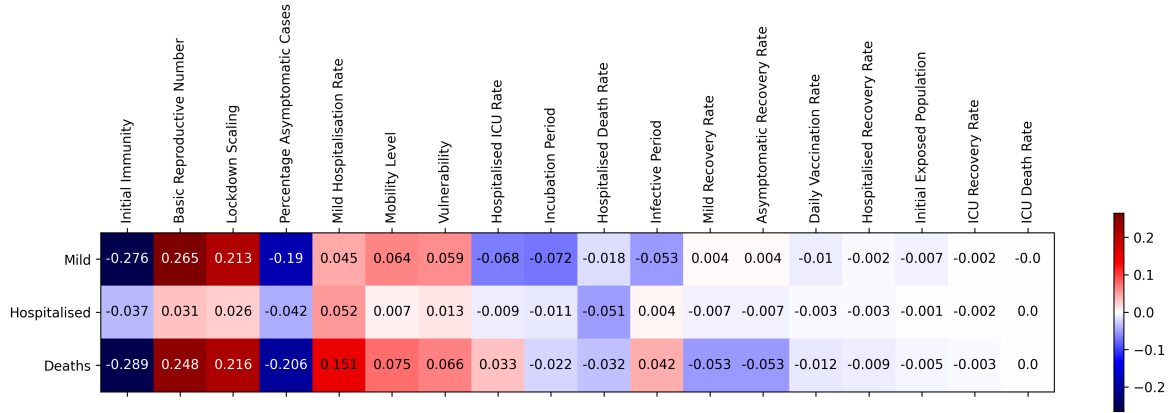


Figure 3.2: A correlation plot, showing positive correlation in a red scale, and negative correlation in a blue scale. Simply put, the correlation between the input and output values. The values were obtained from the runs of the MC sensitivity analysis. The input variables are the columns of this figure and the outcomes are the rows. The correlation is in ascending order, according to high positive or negative correlation value.

The correlation plot in Figure 3.2 shows that there are strong positive and negative correlations (portrayed in vivid colours), but also weak correlations (portrayed in pale colours) for the majority of the input parameters. This shows that the SEIRDV model is robust, since only four input variables effect the variability of the output generated by the SEIRDV model significantly. Positive, negative and weak correlation are discussed below

**Negative Correlation** The input variable that has the strongest negative correlation amongst all other variables is the immunity level. Figure 3.2 indicates that COVID-19 cannot spread effectively, if the vast majority of the population is immune to it. Another input variable, that reduces the number of mild, hospitalised and death cases, is the percentage of asymptomatic cases  $p_1$ . The higher the value of  $p_1$ , the lower the number of mild, hospitalised and death cases. This result was expected, as  $p_1$  is the proportion of the individuals that become asymptomatic and recover, whereas  $1 - p_1$  is the proportion that is not. Therefore a negative correlation between  $p_1$  and mild, hospitalised and dead cases should exist. The correlation between  $p_1$  and hospitalisations is the strongest. This means that the higher the proportion of asymptomatic cases, the lower the number of hospitalisations. This result is valid, because  $p_1$  directly effects the virulence of COVID-19 and therefore a strong negative correlation between  $p_1$  and hospitalisations must exist.

**Positive Correlation** The correlation plot shows that the basic reproduction number and the lockdown scale have the highest positive correlation. It makes sense, because an increase in  $R_0$ , increases the number of infections, which increases the number of mild, hospitalised and death cases. Similarly, the lockdown scale has a positive correlation, since a higher lockdown scale implies low government intervention and low lockdown scale a high government intervention. Hence, the higher the lockdown scale, the more social interactions may occur and the higher the infections of COVID-19. As predicted from the literature review, if the basic reproduction number increases, so do the number of infections.

**Weak Correlations** Some transition rates show little to no effects on the number of mild, hospitalised or death cases. For example, the daily ICU or hospitalisation recovery rates have little to no effect on the outcome. Another example is the vaccination rate, as the increase of vaccination rate during or just before the start of a wave does not allow enough time for people to build immunity. The least influential variable is the initial exposed population. This could be due to COVID-19 being a virus and hence can spread fast enough that it is arbitrary how many completely susceptible people were initially exposed to the virus. This implies that it may only take one exposed to COVID-19 case, to start the COVID-19 pandemic.

### 3.5 Concluding remarks

Chapter 3 provided the proposed methodology of the extended SEIRDV model. Figure 3.1 describes each compartment and its transition rates. Validating the construct of the extended SEIRDV model was done through a sensitivity analysis. The result in Figure 3.2 showed reliability and construct validity of the SEIRDV model. However, the sensitivity analysis did not show whether the SEIRDV model can reproduce the actual case numbers observed. Therefore a second construct validity test via a simulation study will be applied in Chapter 4, to validate the accuracy and validity of the model extensively.



## Chapter 4

# Discussion and Results

Health care and consumer expectations are demanding new standards, beyond healthcare regulations, to understand how business decisions are being driven by artificial intelligence (AI). Health care is becoming more inclusive in modernising their systems through technology initiatives to empower teams with decision-making authority. There is increasing recognition that lowering the cost to serve requires empowering ecosystem actors in new ways. Digital solutions are generating more cost-effective care recommendations utilising contextually aware devices, and AI-driven solutions to elevate caregivers, pharmacists, and other supporting roles with previously untapped potential.

In order to consult healthcare stakeholders, it is vital to produce a model that is responsive, explainable, fair, safe and secure to use. This will gain the trust of its stakeholders. For the model to be effective, it must be accurate. The SEIRDV model offers explainability, security and fairness, simply because it is a deterministic model (a system of differential equations) and does not learn bias from data. However, reliability and accuracy needs to be tested, since any parameters in this model changes the outcome.

Infectious disease compartmental models have been vital in helping decision makers anticipate the waves of COVID-19 infections. Given that the government lockdown intervention stopped and the nation returns to business as usual, there is an imperative to look at the past data retrospectively to understand and learn from the historic data. In this application, we specifically focus on the five districts within the Gauteng province of South Africa. The model in this application uses multiple spatial factors that must be taken into consideration. This inevitably led to the circumstance that the scope is focused on the districts of Gauteng, due to spatial data only being available for Gauteng districts.

### 4.1 Data Gathering

To build the SEIRDV model for COVID-19 in South Africa, we need the transition rates of people that move from one compartment to the next. Transition rates were either calculated from data or reference from other literature. Data used in this study are the daily reported positively tested COVID-19 cases, the vaccination ad-

ministered data, vulnerability data and mobility data. These five data sets are described as the following:

1. The hospital data was given by the Gauteng Health department for research purposes such as this study to obtain approximate values for the recovery and death rates for hospitalised patients in Gauteng. The hospital data recorded the daily number of people in compartments general ward, ICU, recovered and death in Gauteng respectively. Therefore, this application will focus on Gauteng only.
2. Other data needed for the SEIRDV model, such as daily reported positively tested COVID-19 cases where downloaded on the National Institute For Communicable Diseases (NICD) dashboard situated at<sup>1</sup>.
3. The weekly vaccination data were obtained from the the latest vaccine statistics published by the Ministry of Health in South Africa<sup>2</sup>.
4. Mobility data were obtained from Facebook Data For Good<sup>3</sup>, to deduct current location and movement across the various spatial areas.
5. Vulnerability data were obtained from the Vulnerability Dashboard<sup>4</sup>.

The Susceptible-Exposed-Infectious-Removed-Dead-Vaccinated (SEIRDV) model is applied to different infection waves that occurred in Gauteng. The waves all are different scenarios. This is beneficial, because we can both analyse the models performance and explain the effects and outcome that some variables such as the vaccination rates have. However, SEIRDV model is a model, meaning that it is a simplistic view of reality and should not fully explain all effects that caused the waves perfectly.

Nevertheless, these simplistic predictions are important for the country and even more so for populous provinces like Gauteng, where spatial predictions also help in directing intervention resources to areas of greatest demand or priority. The SEIRDV model was used to guide the governments response to the COVID-19 pandemic. This model includes the spatial SEIRD model of Fabris-Rotelli et al. [66] and the added vaccination compartment. The model incorporates movement between spatial units as well as population vulnerability. The spatial units used are at district municipality level, because as vaccination data and case numbers are not publicly available at a lower spatial resolution.

## 4.2 Gauteng and descriptive statistics

Gauteng (GP) is a province in South Africa that is split up in to five districts. Figure 4.1 is a map of the five districts<sup>5</sup>.

<sup>1</sup><https://www.nicd.ac.za/surveillance-reports/national-covid-19-daily-report/> (Accessed October 2022)

<sup>2</sup><https://sacoronavirus.co.za/latest-vaccine-statistics/> (Accessed October 2022)

<sup>3</sup><https://data.humdata.org/dataset/movement-range-maps> (Accessed October 2022)

<sup>4</sup><https://bit.ly/3iFU4Zo>. COVID-19 Vulnerability Dashboard, CSIR. (Accessed on 2020-07-09)

<sup>5</sup><https://ndagis.nda.agric.za> (Accessed October 2022)

<sup>6</sup>[https://www.statssa.gov.za/?page\\_id=964](https://www.statssa.gov.za/?page_id=964) (Accessed October 2022)

<sup>7</sup><https://sacoronavirus.co.za/latest-vaccine-statistics> (Accessed October 2022)

<sup>8</sup><https://www.nicd.ac.za/national-covid-19-daily-report> (Accessed October 2022)

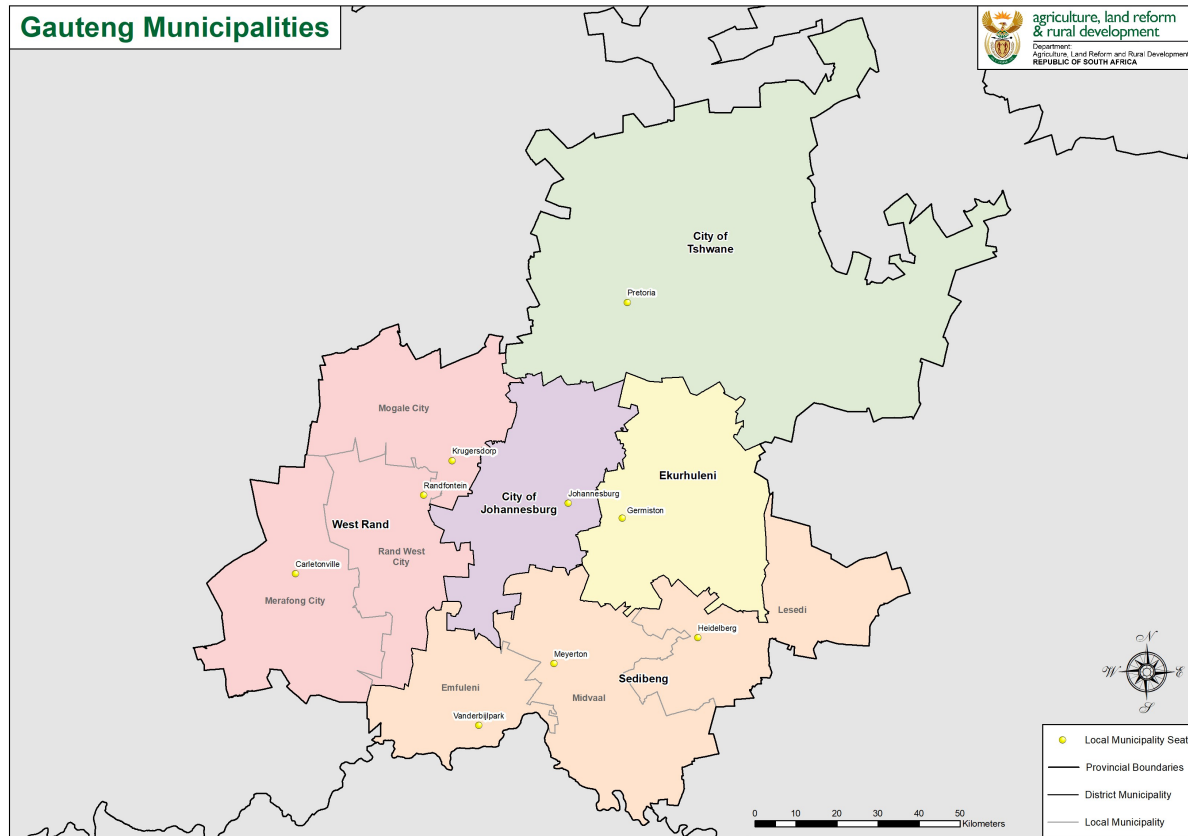


Figure 4.1: A map depicting the Gauteng districts

Measure of comparison	City of Johannesburg	City of Tshwane	Ekurhuleni	Sedibeng	West Rand
Population number <sup>6</sup>	5538596	3522325	3781377	952102	922640
Population proportion of Gauteng	0.38	0.24	0.26	0.06	0.06
vaccination coverage <sup>7</sup>	0.41	0.36	0.33	0.29	0.43
Total proportion of tested positive cases <sup>8</sup>	0.07	0.07	0.05	0.07	0.06

Table 4.1: Descriptive Statistics to compare the districts according to the total population numbers, the total population proportion of Gauteng, the total proportion of fully vaccinated individuals and the total proportion of tested positive cases over all time.

The table of comparison in 4.1 reports that the population numbers are larger for urban and sub-urban areas such as the City of Johannesburg, City of Tshwane and Ekurhuleni in compare to rural spatial areas, such as Sedibeng and West-Rand. The table shows that the largest proportions of individuals live in cities.

The approximate vaccination coverage displayed in Figure 4.2 of each district is significantly greater than zero, giving reason to include vaccinations into the SEIRDV model. The vaccination coverage for City of Johannesburg and West Rand are above average vaccination coverage in Gauteng, whereas City of Tshwane, Ekurhuleni and Sedibeng have roughly the same vaccination coverage. Hence, it is important to adjust the model spatially to take heterogeneous vaccine coverage into account.

Next to the vaccination coverage, the total proportion of tested positive cases were mostly 0.07, which shows that the number of positive tested cases are severely under-reported.

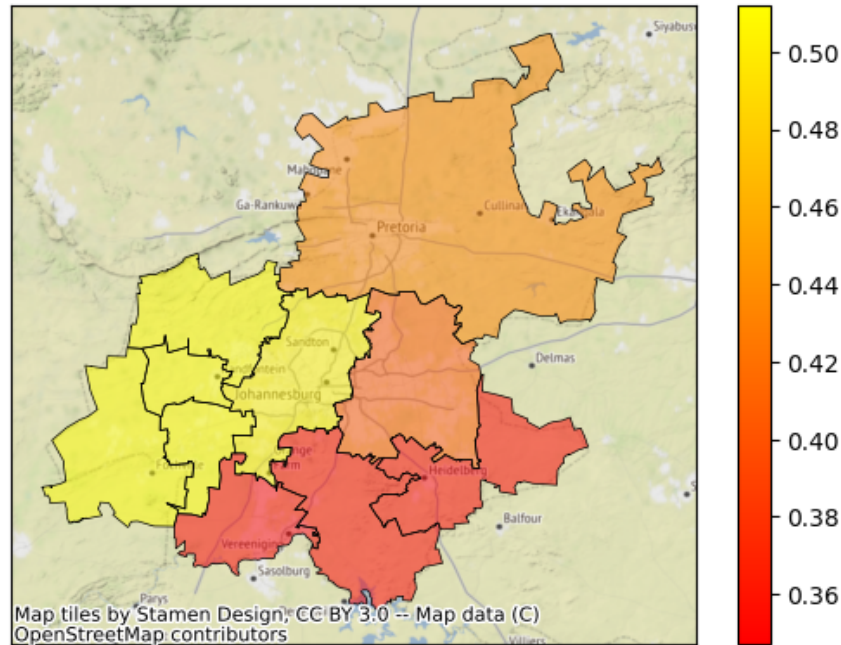


Figure 4.2: A heatmap, that visualises the approximate proportion of fully vaccinated individuals for each district in Gauteng (February 21, 2020 – October 19, 2022).

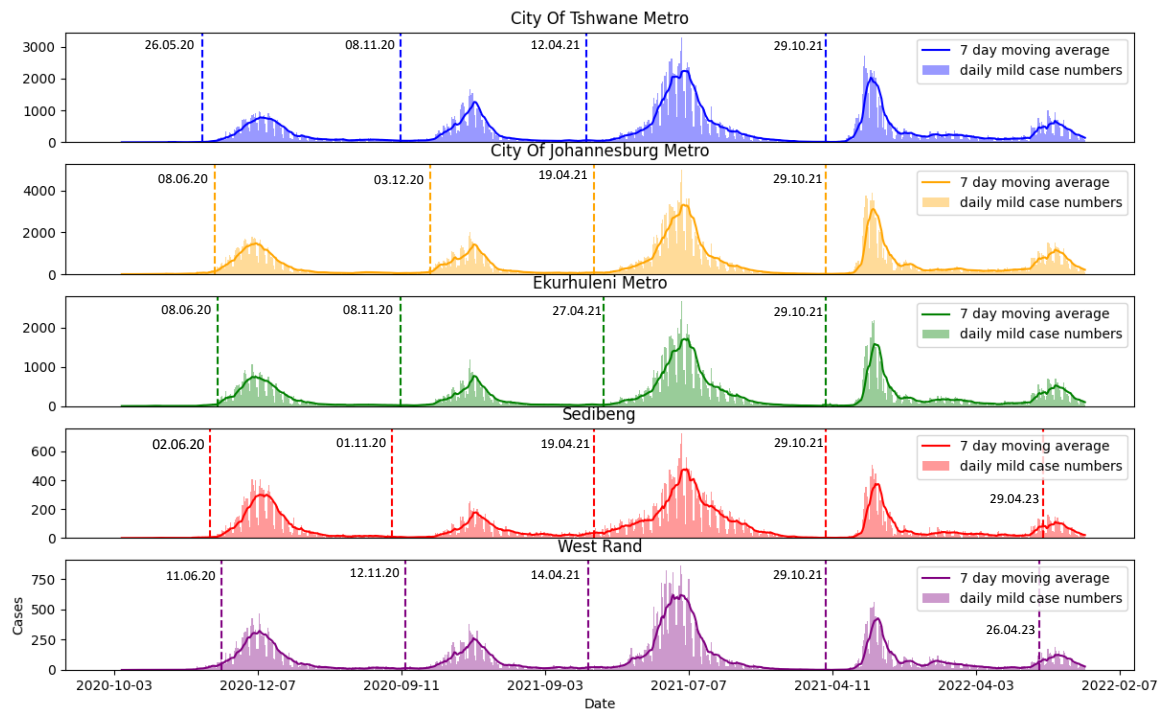


Figure 4.3: Five subplots for each district of Gauteng. Each subplot shows the daily number of tested positive cases (displayed by bars) together with the 7 day moving average (represented by a solid line). Every subplot also shows dashed lines with relating start dates of the COVID-19 waves.

In application, when the daily number of tested positive cases are plotted it shows surges in new cases followed by declines. The series of infections the coronavirus pandemic showed so far are known as COVID-19 waves. These waves also emerge when plotting the daily number of tested positive cases for each district, as seen in Figure 4.3. The waves are outlined by the 7 day moving average for each district. The tested number of positive cases (displayed by the y-axis) for each district shows different number of tested positive cases. For City of Johannesburg and Tshwane the cases vary from 0 up to 3000 or 4000, whereas for smaller districts such as West Rand and Sedibeng, the cases only vary from 0 up to 600 or 750 respectively.

The dashed vertical lines are the estimated start dates for each wave and district. The estimated start dates were calculated using early warning detector proposed in [112]. The start dates are mostly different for each district. Early warning signals (EWS) for wave 5 were only detected for Sedibeng and West Rand. Furthermore, Figure 4.3 shows that the EWS detected the fifth waves for Sedibeng and West Rand later than the actual start of the fifth waves in these districts. No EWS occurred for Johannesburg, Tshwane and Sedibeng.

Figure 4.3 indicates that the dynamics of the spread of COVID-19 does not only differ between districts, but also over time. This observation gives an imperative to predict the waves in Figure 4.3 with the SEIRDV model. We need to estimate the input parameters, when predicting each wave by the differential equations of the SEIRDV model. Furthermore, we can utilise a simulation study to determine whether the SEIRDV model is valid in specific real-life applications. The estimated input parameters are reported in Table 4.2 and will be discussed next.

### 4.3 Simulation study

To validate whether the model can re-produce the case numbers observed, we used the estimated transition rate distributions, to simulate multiple runs of the model using 50 different seeds. With the multiple runs, we obtained a confidence interval of the models predictions. The overall runs were averaged to get the average model prediction. Based on this, we can evaluate how well the model can reproduce the data that was observed. The SEIRDV algorithm to model a COVID-19 wave (Algorithm 1) is used to estimate each COVID-19 wave separately.

#### 4.3.1 Estimating parameters

The parameters for the SEIRDV model for each district and each wave is discussed in this subsection. The estimates are summarised in Table 4.2

##### 4.3.1.1 Estimating starting values

The simulation study used one exposed case to start a wave. For simplicity, the simulation study assumes there are no cases in the infected, recovered or dead compartment at the beginning of a wave. The number of individuals in the vaccination compartment at the beginning of wave 1 is 0. This is based on the assumption that no cases recovered from the COVID-19 nor received a COVID-19 vaccination at the beginning of wave 1.

District	Input parameter	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
All	$S_i$ starting value	$N_i - E_i - V_i$				
	$E_{1i}$ starting value	1				
	$I_{1i}$ starting value	0				
	$I_{2i}$ starting value	0				
	$I_{3i}$ starting value	0				
	$I_{4i}$ starting value	0				
	$R_i$ starting value	0				
	$D_i$ starting value	0				
	$V_i$ starting value	0	$0.3N_i$	$0.4N_i$	$0.58N_i$	$0.9N_i$
	$R_0$	$\Gamma(57.2, 0.05)$				
	$\rho$	0.75				
	$\frac{1}{\sigma}$	$\Gamma(2, 1)$				
	$p_1$	0.75				
	$\alpha_{23i}$	0.002272727	0.003125	0.006	0.004	0.002
	$\alpha_{24i}$	0.001282051	0.002083333	0.002173913	0.001612903	0.002380952
Tshwane	$g$	0.80	0.67			
	$N$	3522325				
	$\Psi_i$	19.67188339				
	$\gamma_{31}$	$\Gamma(0.012, 8.035)$	$\Gamma(0.077, 1.296)$	$\Gamma(0.002, 61.717)$	$\Gamma(0, 46353.52)$	$\Gamma(0.001, 272.619)$
	$\gamma_{41}$	$\Gamma(0.012, 8.035)$	$\Gamma(0.077, 1.296)$	$\Gamma(0.002, 61.717)$	$\Gamma(0, 46353.52)$	$\Gamma(0.001, 272.619)$
	$\delta_{31}$	$\Gamma(0.007, 3.164)$	$\Gamma(0.013, 1.339)$	$\Gamma(0.001, 10.255)$	$\Gamma(0.002, 67.534)$	$\Gamma(0.009, 0.202)$
	$\delta_{41}$	$\Gamma(0.001, 49.589)$	$\Gamma(0.049, 1.288)$	$\Gamma(0.064, 0.891)$	$\Gamma(0.001, 5.805)$	$\Gamma(0.02, 1.352)$
	$v_i$	$\Gamma(0, 0)$	$\Gamma(0.001, 0.156)$	$\Gamma(0.002, 0.575)$	$\Gamma(0.001, 1.213)$	$\Gamma(0.001, 0.646)$
Johannesburg	$N$	5538596				
	$\Psi_i$	15.36825768				
	$\gamma_{32}$	$\Gamma(0, 2139.376)$	$\Gamma(0.096, 1.506)$	$\Gamma(0, 1134.626)$	$\Gamma(0.01, 3.816)$	$\Gamma(0, 14704.91)$
	$\gamma_{42}$	$\Gamma(0.011, 7.655)$	$\Gamma(0.077, 1.272)$	$\Gamma(0.077, 1.272)$	$\Gamma(0, 12689.14)$	$\Gamma(0.003, 50.091)$
	$\delta_{32}$	$\Gamma(0.001, 2.884)$	$\Gamma(0.004, 1.304)$	$\Gamma(0.004, 1.304)$	$\Gamma(0.002, 51.794)$	$\Gamma(0.003, 0.182)$
	$\delta_{42}$	$\Gamma(0.001, 76.195)$	$\Gamma(0.05, 1.277)$	$\Gamma(0.05, 1.277)$	$\Gamma(0, 5.941)$	$\Gamma(0.054, 0.721)$
	$v_i$	$\Gamma(0, 0)$	$\Gamma(0.001, 0.153)$	$\Gamma(0.003, 0.627)$	$\Gamma(0.003, 1.180)$	$\Gamma(0.001, 0.644)$
Ekurhuleni	$N$	3781377				
	$\Psi_i$	16.46564581				
	$\gamma_{33}$	$\Gamma(0, 42947.5)$	$\Gamma(0.105, 1.508)$	$\Gamma(0.105, 1.508)$	$\Gamma(0.016, 2.877)$	$\Gamma(0, 124948.9)$
	$\gamma_{43}$	$\Gamma(0.012, 7.869)$	$\Gamma(0.079, 1.297)$	$\Gamma(0.079, 1.297)$	$\Gamma(0, 52661.49)$	$\Gamma(0.001, 85.659)$
	$\delta_{33}$	$\Gamma(0.003, 3.091)$	$\Gamma(0.003, 1.252)$	$\Gamma(0.003, 1.252)$	$\Gamma(0.02, 5.719)$	$\Gamma(0.002, 0.132)$
	$\delta_{43}$	$\Gamma(0.001, 69.372)$	$\Gamma(0.047, 1.298)$	$\Gamma(0.047, 1.298)$	$\Gamma(0.001, 3.706)$	$\Gamma(0.014, 4.338)$
	$v_i$	$\Gamma(0, 0)$	$\Gamma(0.001, 0.198)$	$\Gamma(0.002, 0.524)$	$\Gamma(0.001, 1.194)$	$\Gamma(0.001, 0.652)$
Sedibeng	$\gamma_{34}$	$\Gamma(0, 5618.201)$	$\Gamma(0.106, 1.508)$	$\Gamma(0.106, 1.508)$	$\Gamma(0.003, 15.411)$	$\Gamma(0, 15930.77)$
	$\gamma_{44}$	$\Gamma(0.022, 4.087)$	$\Gamma(0.079, 1.231)$	$\Gamma(0.079, 1.231)$	$\Gamma(0, 40739.05)$	$\Gamma(0.011, 10.557)$
	$\delta_{34}$	$\Gamma(0.003, 3.336)$	$\Gamma(0.004, 1.379)$	$\Gamma(0.004, 1.379)$	$\Gamma(0.04, 3.431)$	$\Gamma(0, 0)$
	$\delta_{44}$	$\Gamma(0.002, 44.392)$	$\Gamma(0.045, 1.438)$	$\Gamma(0.045, 1.438)$	$\Gamma(0.001, 1.211)$	$\Gamma(0, 0)$
	$v_i$	$\Gamma(0, 0)$	$\Gamma(0.001, 0.184)$	$\Gamma(0.002, 0.532)$	$\Gamma(0.002, 1.156)$	$\Gamma(0.001, 0.660)$
	$N$	952102				
	$\Psi_i$	13.49495407				
West Rand	$\gamma_{35}$	$\Gamma(38.351, 0.002)$	$\Gamma(38.351, 0.002)$	$\Gamma(38.351, 0.002)$	$\Gamma(38.351, 0.002)$	$\Gamma(38.351, 0.002)$
	$\gamma_{45}$	$\Gamma(38.35, 0.002)$	$\Gamma(38.35, 0.002)$	$\Gamma(38.35, 0.002)$	$\Gamma(38.35, 0.002)$	$\Gamma(38.35, 0.002)$
	$\delta_{35}$	$\Gamma(0.017, 0.204)$	$\Gamma(0.005, 1.279)$	$\Gamma(0.005, 1.279)$	$\Gamma(0, 0.242)$	$\Gamma(0, 0)$
	$\delta_{45}$	$\Gamma(0.02, 3.879)$	$\Gamma(0.05, 1.476)$	$\Gamma(0.05, 1.476)$	$\Gamma(0, 0.406)$	$\Gamma(0.033, 0.775)$
	$v_i$	$\Gamma(0, 0)$	$\Gamma(0.001, 0.185)$	$\Gamma(0.003, 0.430)$	$\Gamma(0.013, 0.793)$	$\Gamma(0.001, 0.625)$
	$N$	922640				
	$\Psi_i$	16.23844228				

Table 4.2: The estimated parameters of each district of each wave.

For simplicity, we place a proportion of the population in the vaccinated compartment to imitate immunity, that increases wave after wave. This simplicity refers to [48, 95], which were previously covered in Section 2.2. Assuming that 30% of the population was immune after recovering from COVID-19 during wave 1,  $0.3N_i$  were placed in the vaccinated compartment at the start of wave 2. At the start of wave 3, we assumed that 40% of the population was immune towards COVID-19, 58% at the start of wave 4 and 90% at the start of wave 5. Since the SEIRDV model assumes that the population numbers are fixed, the remaining number of individuals are in the

compartment  $S_i$  and hence

$$S_i = N_i - E_i - V_i \quad (4.1)$$

Thus the starting value of  $S_i$  depends on  $i$  and  $t$  due to  $V_i$  differing at the starting time point of any given wave. However, some parameters, such as the incubation period, are assumed to be independent of  $i$  and  $t$  and are thus homogeneous.

#### 4.3.1.2 Homogeneous parameters

Homogeneous parameters, that are assumed to be independent of  $i$  or  $t$  are following:

- $R_0 \sim \Gamma(57.2, 0.05)$  with mean 2.86 from [66]
- $(\gamma_{1i})^{-1} \sim \Gamma(2, \frac{7}{2})$  from [82, 104, 133]
- $\rho = 0.75$  from [3]
- $p_1 = 0.75$  from [84, 108, 133]
- $\frac{1}{\sigma} \sim \Gamma(2, 1)$  from [117]

The alpha rates estimates are the same for the districts, but differ from wave to wave. Interestingly, the alpha rate for the third wave is remarkably higher than the rest of the alpha rate estimates. This indicates that the dominant variant of concern (the delta variant) during the third wave was notably more virulent than the variants of concern that were dominant in other COVID-19 infection waves.

**Time dependent parameters:** The parameters  $\kappa$  and  $g$  do not vary for district  $i$ , but vary for  $t$ . The lockdown scaling factor  $\kappa$ , from [66] is summarised in Table 4.3.

Level	Date	$\kappa$
Hard lockdown - level	26 March - 30 April 2020	0.6
Level 4 lockdown (L4)	1 May - 30 May 2020	0.7
Level 3 lockdown (L3)	1 June - 17 August 2020	0.75
Level 2 lockdown (L2)	18 August - 20 September 2020	0.8
Level 1 lockdown (L1)	21 September - 29 December 2020	0.85
Lifting of National State of Disaster	4 April 2022	1

Table 4.3: Government intervention dates with the associated scaling factor  $\kappa$  [66].

The other time dependent parameter is the vaccine efficiency  $g$ . To estimate  $g$ , the average of published vaccination effectiveness of Pfizer and Johnson & Johnson was calculated. The vaccine efficiency value is a simplified representation of the actual vaccine efficiency. Equation 4.2 describes  $g$  for wave 1.

$$g = \frac{0.95 + 0.67}{2} = 0.81 \quad (4.2)$$

For wave 2 to 5,  $g$  was adjusted to Equation 4.2

$$g = \frac{0.65 + 0.67}{2} = 0.66 \quad (4.3)$$

The next paragraph discusses some parameters in SEIRDV that vary over  $t$  and not  $i$ .

#### 4.3.1.3 Spatial parameters

The SEIRDV model assumes that the following parameters are dependent on  $i$  but are independent of time  $t$ :

- $N_i$ , for spatial area  $i$ , which are reported in Table 4.1
- $\Psi_i$ , which are reported in Equation 4.4
- $\beta_{si}$ , which are estimated by Equation 3.6
- $\beta_{vi}$ , which are estimated by Equation 3.6
- $\alpha_{23i}$  and  $\alpha_{24i}$  from<sup>9</sup>.

The vulnerability index is based on the weighted average of the normalised vulnerability indices described in [81] district and reported in the matrix in Equation 4.4.

$$\Psi_i \begin{bmatrix} Tshwane & Johannesburg & Ekurhuleni & Sedibeng & WestRand \\ 19.67188339 & 15.36825768 & 16.46564581 & 13.49495407 & 16.23844228 \end{bmatrix} \quad (4.4)$$

From the matrix in 4.4, it can be seen that the vulnerability rate is the highest in Tshwane, whereas Sedibeng has the lowest vulnerability rate. The vulnerability rates of Johannesburg, Ekurhuleni and West Rand are roughly similar to each other. Overall, it can be concluded, that the vulnerability rate  $\Psi_i$  are heterogeneous. Taking this into account is important as individuals that are more economically unequal and lack capacity in terms of social capital are more likely to die of COVID-19 [36].

The remaining parameters in Table 4.4 were estimated using data. We used box plots to display summary statistics of the estimated parameters. We discovered that the parameters are spatial temporal, since these estimates depend on  $t$  and  $i$ .

#### 4.3.1.4 Spatial temporal parameters

Recall the transition rates of Figure 3.1. Given the data at hand, we can calculate the following parameters transition rates. Table 4.4 reports the parameters and their associated equations.

The data available<sup>10</sup> is the number of weekly administered dosages ( $D_i^w$ ) for spatial area  $i$ . Using the data at hand, the vaccination rate for a district,  $v_i$ , is calculated using Equation 3.4. The vaccination rate is only

<sup>9</sup>Details and reports at [www.sacmcepidemicexplorer.co.za](http://www.sacmcepidemicexplorer.co.za) (Accessed October 2022)

<sup>10</sup><https://sacoronavirus.co.za/latest-vaccine-statistics> (Accessed October 2022)



Parameter	Equation
$\gamma_{i3}$	3.9
$\gamma_{i4}$	3.10
$\delta_{i3}$	3.12
$\delta_{i4}$	3.13
$\nu_i$	3.4 and 4.5

Table 4.4: Parameters that can be estimated using their associated equations and data.

calculated from wave 2 on-wards, as vaccinations were not yet available to individuals in South Africa in the first wave. Furthermore, the data available does not specify how many of the weekly administered dosages were Pfizer and how many Johnson & Johnson dosages. Therefore, we simplify, by using the overall numerical proportion of vaccine type given in South Africa, which is shown in Figure 4.4.

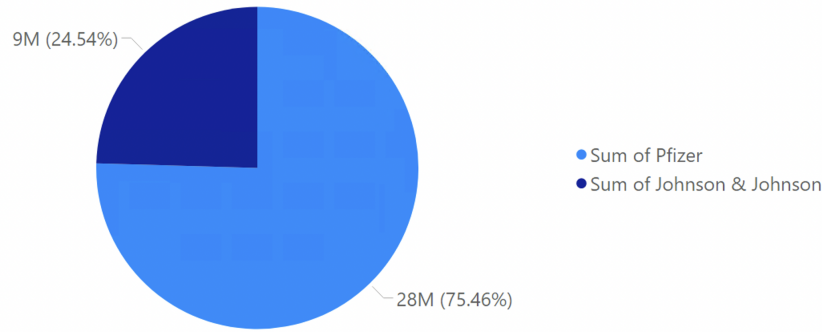


Figure 4.4: A pie chart illustrating the numerical proportion of Pfizer and Johnson & Johnson dosages in South Africa.

The pie chart reports the proportion of Pfizer dosages as approximately 0.75 and Johnson & Johnson as approximately 0.25. Note that it takes 2 Pfizer dosages to fully vaccinate an individual and 1 for Johnson & Johnson. It follows that an individual on average receives about  $E[d]$  doses. We use Equation 3.1, to obtain  $E[d]$ .

$$E[d] = 0.2454 + 2(0.7546) = 1.755 \quad (4.5)$$

The parameters listed in Table 4.4 above all depend on  $i$  and  $t$ . Hence, the parameters should differ, when comparing districts to one another for a given wave. To investigate and compare the transition rates for each district and for each wave, box-plots are used.

**Heterogeneity of daily observed vaccination rates** Previously, we have seen that the total proportion of vaccination coverage differs across districts. This however, might directly mean that the vaccination transition rates are different too.

Figure 4.5 shows a difference in  $\nu_i$  between waves, since vaccinations were not available during the first wave. The model needs to take  $\nu_i = 0$  into account, when no vaccinations were available. The values for  $\nu_i$

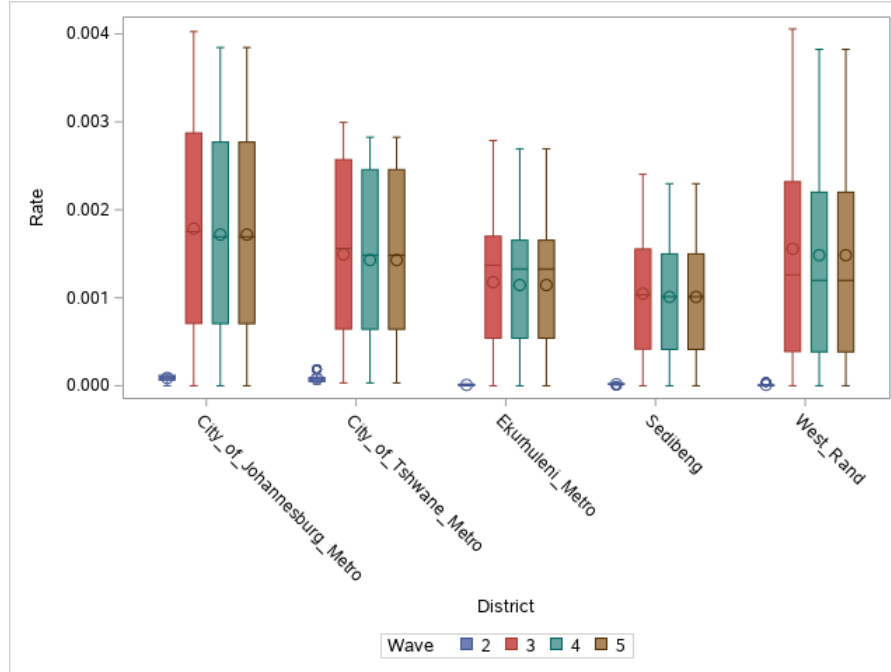


Figure 4.5: Boxplots of the vaccination rates  $v_i$  grouped by  $i$  districts and coloured in 5 waves

during wave 2 are close to zero and are displayed to illustrate that the vaccination availability was scarce. This is expected, because the vaccination roll-out started at the end of wave 2. According to the South African COVID-19 vaccination programme, front line health care workers received vaccinations first. Thus it follows that the  $v_i$  for wave 2 stem from the vaccinations given to front line health care workers.

Note how the vaccination rate are on average lower in the City of Tshwane, Ekurhuleni and Sedibeng during wave 3 to 5 compared to the average of City of Johannesburg and West Rand. The range of  $v_i$  in the City of Johannesburg and West Rand are greater for wave 3 to 5 than the rest of the districts. This is indicated by the y-axis of Figure 4.5, where City of Johannesburg and West Rand had vaccination rates up to 0.004, whereas City of Tshwane, Ekurhuleni and Sedibeng only up to 0.003. A higher vaccination rate goes hand in hand with the reported vaccination coverage in Table 4.1 for Johannesburg and West Rand that are higher than the vaccination coverage City of Tshwane, Ekurhuleni and Sedibeng. Higher vaccine coverage is good for reducing severe and critical infections against the disease, since vaccines intend to provide immunity against COVID-19. The more that others are vaccinated, the less likely individuals who are unable to be protected by vaccines are at risk of even being exposed to the harmful pathogens. Thus transition rates to general ward and ICU may be affected by the different vaccine coverages of the districts.

**Heterogeneity of daily observed general ward recovery rates** The daily general ward recovery rates are the daily rate of general ward patients that were discharged alive. This transition rate is denoted as  $\gamma_{i3}$  for district  $i$ .

Figure 4.6 shows box-plots grouped by districts and categorised by waves in colour. Any box-plots marked in the same colour differ significantly from one district to another, indicating heterogeneity between the districts.

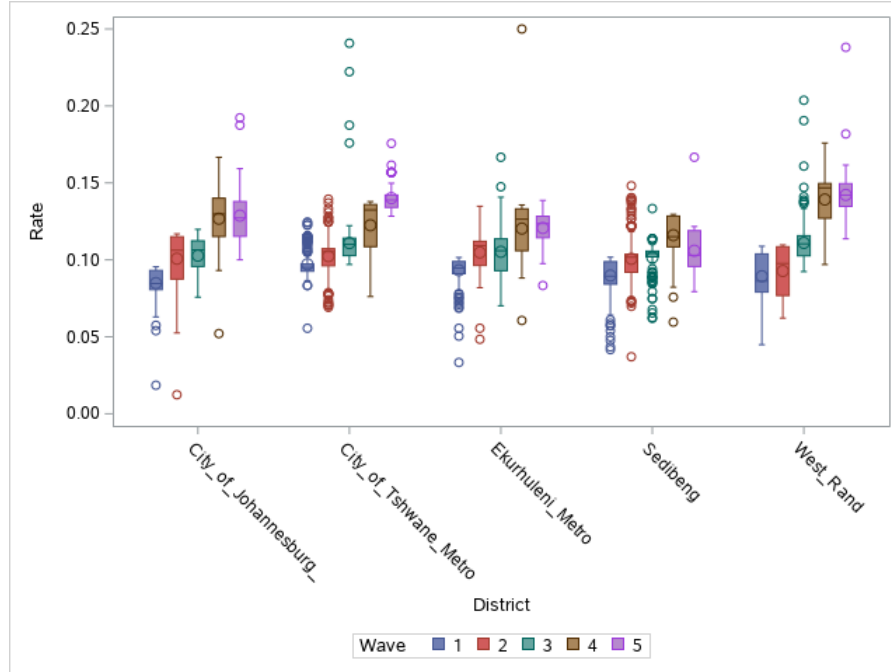


Figure 4.6: Boxplots of the General Ward Recovery Rate  $\gamma_{3i}$  grouped by districts and categorised in 5 waves

However, note that the observations of recovery rates within each district increase from wave 1 to 5. This implies that the recovery rates improve over time for all districts, because the observed recovery rates show an ascend from wave 1 – 5. Especially in waves 4 and 5, the most recovery rates are higher than those from waves 1 and 2. This implies that the recovery rate not only differs amongst districts, but also across waves. Outliers are expected, because some patients differ in recovery rate due to comorbidity or age and hence are observed and can be read from the box-plots.

**Heterogeneity of daily observed ICU recovery rates** Similarly to the daily observed general ward recovery rates, the boxplots in Figure 4.7 differ across all districts when considering waves one at a time. The major difference is that ICU patients are in critical condition, require a higher level of care and often brought into an ICU for breathing support.

Dissimilar to the daily observed general ward recovery rates is that the ICU recovery rate does not increase from wave 1 to 5. This difference could have a range of reasons, including the amount of ICU beds and ventilators available in a district and the total number of patients in critical condition. Further dissimilarities are that little to no ICU cases are observed with waves 4 and 5, giving evidence that ICU recovery rates do not only vary across districts but also across waves.

**Heterogeneity of daily observed general ward death rates** The descriptive statistics of the daily observed general ward death rates are visualised as boxplots in Figure 4.8. The observations are the opposite of daily general ward recovery rates, since patients either recover or die.

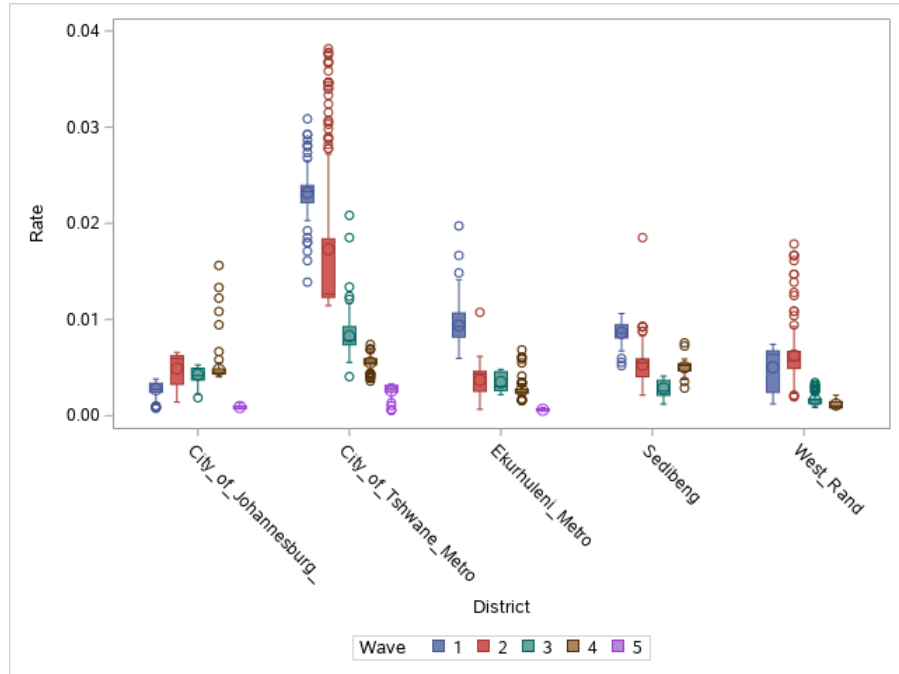


Figure 4.7: Boxplots of the ICU Recovery Rate  $\gamma_{4i}$  grouped by  $i$  districts and coloured in 5 waves

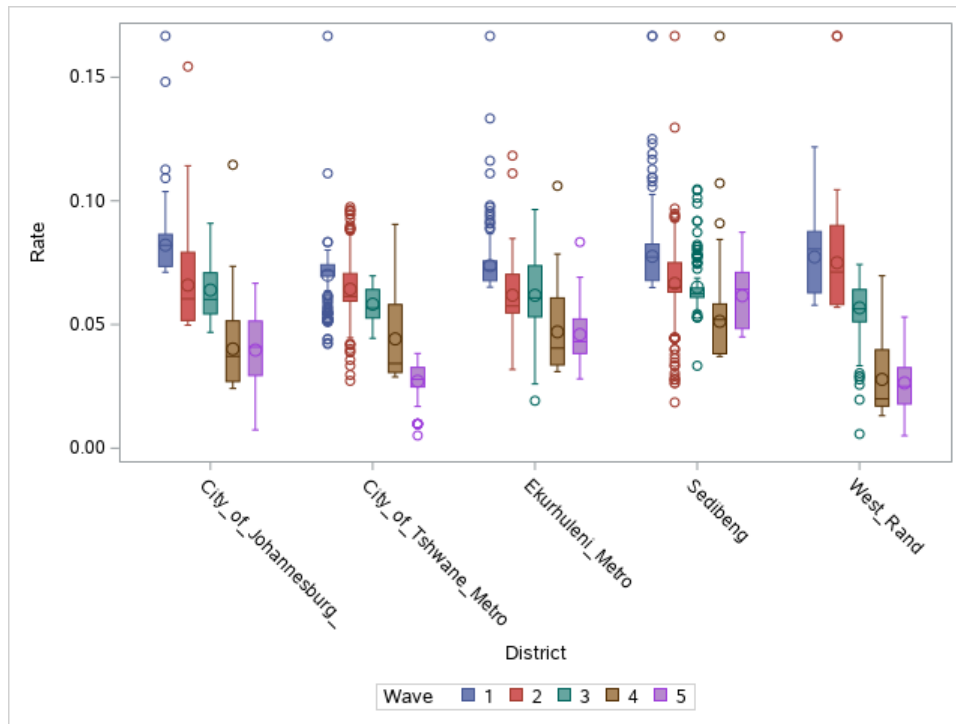


Figure 4.8: Boxplots of the GW death rate  $\delta_{3i}$  grouped by  $i$  districts and coloured in 5 waves

Like the general ward recovery rates, general ward death rates differ amongst districts and across waves.

**Heterogeneity of daily observed ICU death rates** The daily observed ICU ward death rates are the opposite of daily ICU recovery rates, since patients either recover or die of COVID-19. The descriptive statistics of ICU death rates are visualised in Figure 4.9. Figure 4.9 shows that the ICU death rates differ amongst districts, and across waves.

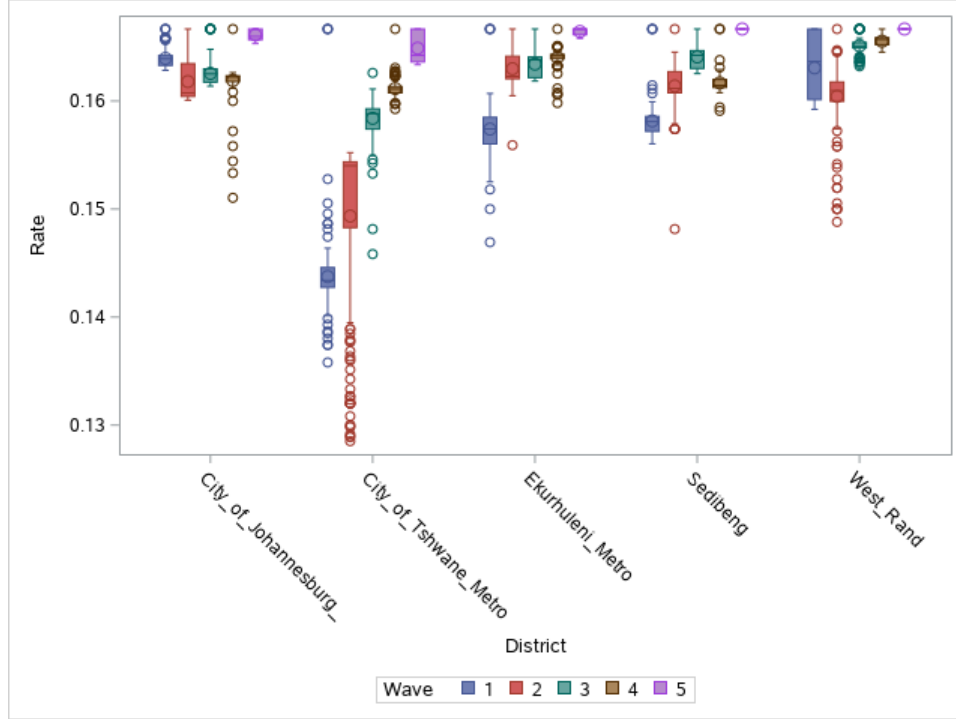


Figure 4.9: Boxplots of the ICU death rate  $\delta_{4i}$  grouped by  $i$  districts and coloured in 5 waves

**Heterogeneity of daily observed mobility rates** This study makes use of the "Movement range maps" data set<sup>11</sup>. The data indicates the change in mobility,  $F_i(t) \in (-1, 1)$  (which may be interpreted as a percentage change  $(-100\%, 100\%)$ ).  $F_i(t)$  is the percentage change of mobility for a district  $i$  on a given day  $t$  relative to a pre-covid mobility baseline calculated in February 2020.

Figure 4.10 visualises  $F_i^{(t)}$ . Figure 4.10 signifies a remarkable drop in the average mobility in late March. This corresponds to the first hard lockdown South Africa had on the March 27, 2020 2.1. The hard lockdown enforced severe travel restrictions and constituted a strict stay at home directive [120]. Overall, the  $F_i^{(t)}$  is primarily negative over the entire study period, implying that mobility patterns remain less than before South Africa's first lockdown.

We substitute  $F_i^{(t)}$  and  $d_{ij}$  in the Equation in 3.8. The 5 by 5 matrix consists of the estimated  $w_{ij}^{(t)}$ 's, which represent mobility between the 5 districts at a given day  $t$ . The daily mobility matrices are included in the exposed compartment, to simulate the spread of the virus [66].

**Fitting distributions** The data produced by estimating daily spatial temporal parameters were put through

<sup>11</sup><https://data.humdata.org/dataset/movement-range-maps> (Accessed October 2022)

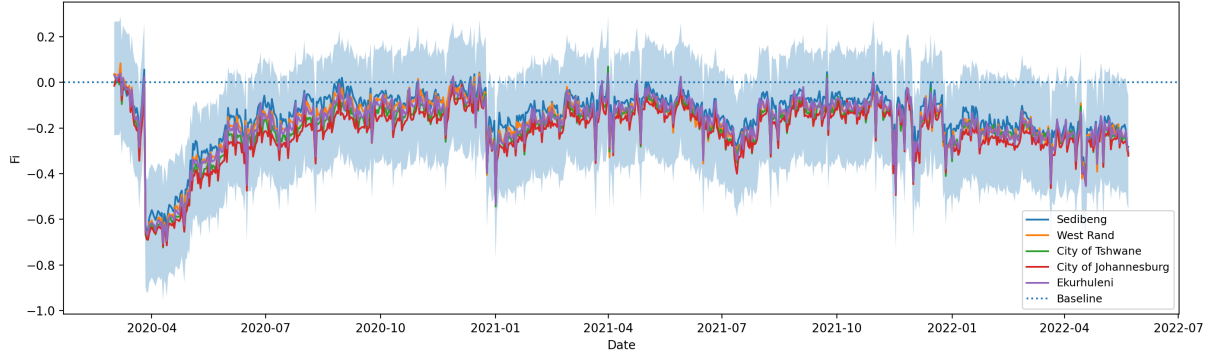


Figure 4.10: The movement range maps data (March 1, 2020 to May 31, 2022) relative to a baseline calculated in the month of February 2020.

the SAS proc univariate procedure, to fit a distribution to each parameter. These best fitted distribution are Gamma summarised in Table 4.2, following the logic in [66].

**Properties of the Gamma distributions** The statistical properties of the Gamma distribution are

$$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}} \quad (4.6)$$

It has a shape parameter  $k > 0$  and a scale parameter  $\theta > 0$ . The mean is  $k\theta$  and the variance  $k\theta^2$ . The gamma distribution is a flexible distribution that can be bell shaped or skewed shaped, depending on  $k$  and  $\theta$  [141].

### 4.3.2 Results of the SEIRDV model simulation study

In this application, the detection of waves was carried out for each of the 5 district municipalities in Gauteng using the proposed method by Fabris-Rotelli et al. [66] with a vaccination compartment. The differential equations of the SEIRDV model is used with input parameters displayed in Table 4.2.

Figure 4.11 shows the simulation study for the mild cases. The model generally over-predicts the mild cases. This is as expected, since the COVID-29 cases are under-reported. Figure 4.12 shows the simulation study for the hospitalised cases. The hospitalised cases is displayed as the sum of general ward cases and ICU cases. Hospitalised cases are predicted accurately, since the true observations fall within the confidence intervals for each wave and each district.

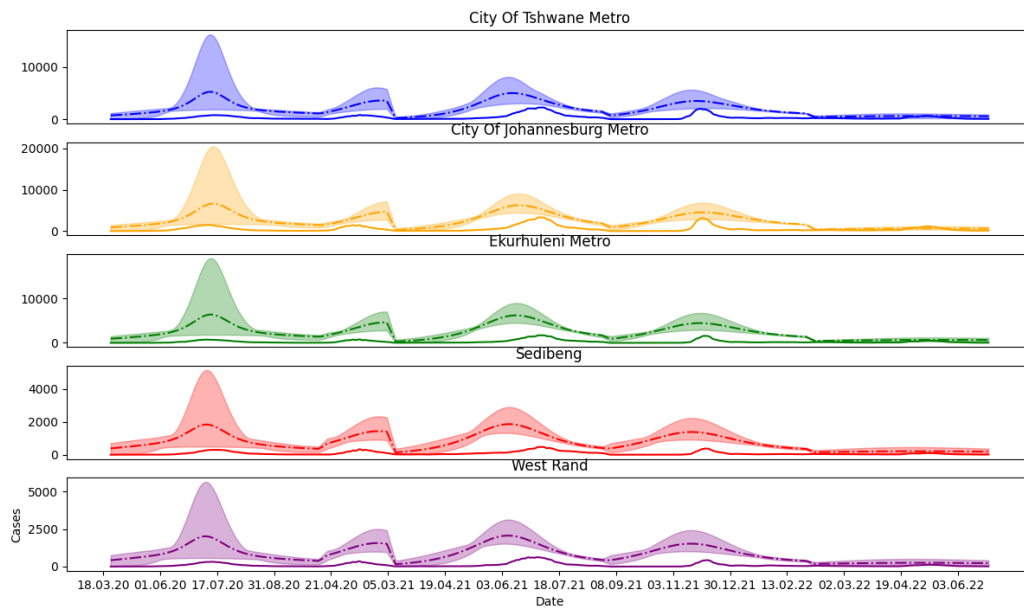


Figure 4.11: Actual mild cases (solid line) vs the estimated mild cases in a (dashed line), and its confidence bounds.

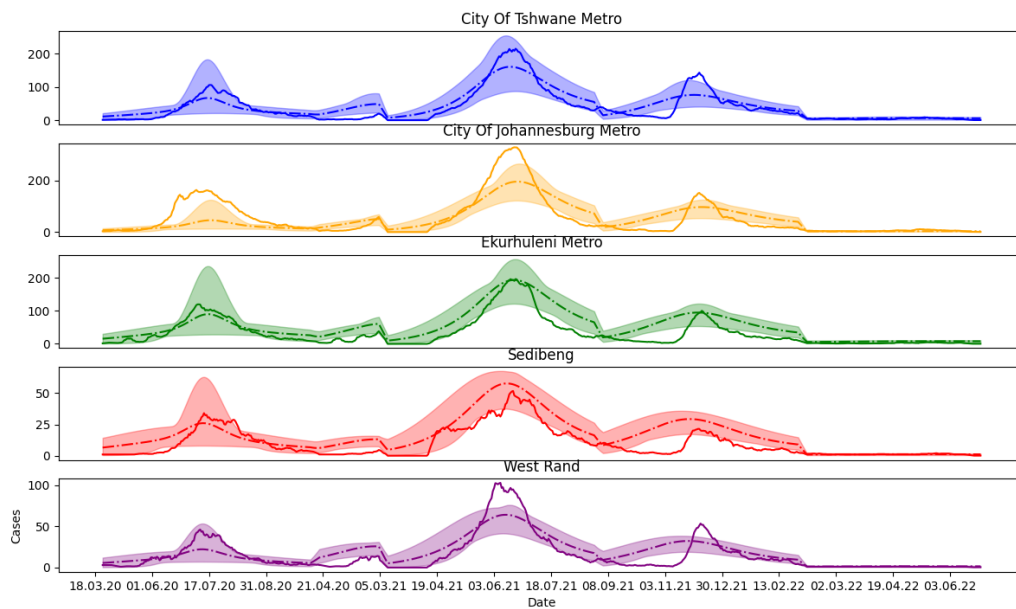


Figure 4.12: Actual hospitalised cases (solid line) vs the estimated hospitalised cases in a dashed line, with confidence bounds. The hospitalised cases is a sum of general ward cases and ICU cases.

## 4.4 Concluding remarks

This application focused on the districts of Gauteng. The district data was gathered from ministerial dashboards, institutes and social media. Using data and literature, each parameter of each wave was estimated and reported in Table 4.2. The SEIRDV model has a mixture of homogeneous, spatial or spatial temporal parameters. The simulation study of the spatial SEIRDV model on Gauteng data showed promising results in predicting the number of cases as well as the peak point and longevity of the wave.



## Chapter 5

# Conclusion and further research

This study addressed the problem of improving current mathematical models by adding a geospatial perspective. Factors that govern spatial areas may lead towards more geospatially oriented governmental interventions such as a vaccination campaign, are considered within the context of a spatially extended SEIRDV model.

The main objective was to assess if a COVID prevention mechanism such as a geospatial sensitive vaccination campaign responds better to the severe burden of the COVID-19 pandemic, given a socially heterogeneous and very mobile population. The objective of incorporating the spatial factors of this study has been achieved. The main drivers of a spatial spread are immunity and lockdown levels of a given area, infectiousness, severity and virulence of the COVID-19 virus.

Further goals of using the model to guide governments response was met, as the model proved to fall under the responsible artificial intelligence framework (RAI) and could be used to guide decision-making authorities. The results show that model accurately predicts the number of hospitalised cases, the longevity and the peaking points of the waves, which earns trust of health care and government decision-making authorities.

Under-reporting of tested positive cases was the leading challenge, as the model would mostly over-predict mild cases. Under-reporting of tested positive cases under-estimates the infectiousness and is strong arguing point of why more attention should be given to test and tracing coverage in future.

Other challenges when fitting the model were the COVID-19 variants that changed the transition rates over time. This gives reason to extend this study and focus on estimating transition rates as functions of time. The sensitivity analysis brought to light that is important to report how many tested cases were admitted to hospital. It is crucial to monitor this transition rate ( $\alpha$ ), because it is the pivoting point of hospitals becoming overloaded. Aside from transition rates, improvements of this study is to add age and comorbidity as dimensions to the model, as the virulence of COVID-19 also proved to be comorbidity and age-dependant [4]. Comorbidity and age risks are also highlighted by the vaccination roll-out plan, that chose to vaccinate the elderly and people with comorbidities first. In future, stigma and vaccine hesitancy could also be implemented in the model. It is also important to highlight that in paper [66], that this research is based on, reported on ward and not district level.

The methodology for estimating the number of susceptible and vaccinated individuals during a pandemic is open to refinement and improvement. Our understanding of pandemic models has grown with the lessons learned from the COVID-19 experience. The methods used in this study are limited by the data that we had available to us. Given the limitations of our data, the methods described in Section 3.1.1 were built to make use of the data presented in Section 4.1.

This study was limited by the spatial resolution of district levels of Gauteng. Access to data at ward level will provide data at a higher spatial resolution, to record the heterogeneity of each spatial area better. This could pinpoint local breakouts accurately and strengthen control of the spread of the disease, a future study should data be available.

This study also comes to the same conclusion as the paper on population mobility [120], where the daily mobility were mostly identical, giving further grounds to improve the method, such that it captures mobility changes at a higher spatial resolution in South Africa.

Overall, the study proved that immunity is the best defence weapon to have during a pandemic and that vaccinations gave over-whelming evidence, that it protected many South Africans during the 4th (and 5th wave if any) of the pandemic. Hence, we conclude:

1. The study extended the SEIRD model in [66] with a spatial vaccination compartment, discussed in section 3.1.1.
2. The result in Figure 3.2 showed reliability and construct validity of the SEIRDV model. The simulation study in 4.3 validated that the SEIRDV model can reproduce the mild and hospitalised case numbers observed in Gauteng. The Figures in 4.3.2 showed promising results in predicting the number of cases as well as the peak point and longevity of the wave.

# Bibliography

- [1] D. Adam. Special report: The simulations driving the world's response to COVID-19. *Nature*, 580(7802):316–319, 2020.
- [2] R.M. Anderson and R.M. May. Vaccination and herd immunity to infectious diseases. *Nature*, 318(6044):323–329, 1985.
- [3] J. Anguelov, R. and Banasiak, C. Bright, J. Lubuma, and R. Ouifki. The big unknown: The asymptomatic spread of COVID-19. *Biomath*, 9(1):ID–2005103, 2020.
- [4] F. Aràndiga, A. Baeza, I. Cordero-Carrión, R. Donat, M.C. Martí, P. Mulet, and D.F. Yáñez. A Spatial-Temporal Model for the Evolution of the COVID-19 Pandemic in Spain Including Mobility. *Mathematics*, 8(10):1677, 2020.
- [5] A.H. Auchincloss, S.Y. Gebreab, C. Mair, and A. V Diez Roux. A review of spatial methods in epidemiology, 2000–2010. *Annual Review of Public Health*, 33:107–122, 2012.
- [6] N. Bacaër. Ross and malaria (1911). In *A Short History of Mathematical Population Dynamics*, pages 65–69. Springer, 2011.
- [7] S.J. Baigent and J.W. McCauley. Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. *BioEssays*, 25(7):657–671, 2003.
- [8] Bangalee V. Bangalee A. Fake news and fallacies: Exploring vaccine hesitancy in South Africa. *South African Family Practice*, 2021.
- [9] F. Bavaud. Models for Spatial Weights: A Systematic Look. *Geographical Analysis*, 30(2):153–171, 1998.
- [10] D. Bernoulli. Réflexions sur les avantages de l'inoculation. *Mém. Paris*, pages 439–482, 1758.
- [11] D. Bernoulli. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Histoire de l'Acad., Roy. Sci.(Paris) avec Mem*, pages 1–45, 1760.
- [12] L. Bologna, K.V. Stamidis, S. Paige, R. Solomon, F. Bisrat, A. Kisanga, S. Usman, and A. Arale. Why Communities Should Be the Focus to Reduce Stigma Attached to COVID-19. *The American Journal of Tropical Medicine and Hygiene*, 104(1):39, 2021.

- [13] Fred Brauer. Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling*, 2(2):113–127, 2017.
- [14] D. Brockmann, V. David, and A.M. Gallardo. *Human Mobility and Spatial Disease Dynamics*, volume 2. Wiley Online Library, 2009.
- [15] R. Brookmeyer. Incubation Period of Infectious Diseases. *Wiley StatsRef: Statistics Reference Online*, pages 1–8, 2014.
- [16] W. Budd. *Typhoid Fever: its Nature, Mode of Spreading and Prevention*. Longmans, Green, 1873.
- [17] Miles W. Carroll, D.A. Matthews, J.A. Hiscox, M.J. Elmore, G. Pollakis, A. Rambaut, R. Hewson, I. García-Dorival, J.A Bore, R. Koundouno, S. Abdellati, B. Afrough, J. Aiyepada, P. Akhilomen, D. Asogun, B. Atkinson, M. Badusche, A. Bah, S. Bate, J. Baumann, D. Becker, B. Becker-Ziaja, A. Bocquin, B. Borremans, A. Bosworth, J.P. Boettcher, A. Cannas, F. Carletti, C. Castilletti, S. Clark, F. Colavita, S. Diederich, A. Donatus, S. Duraffour, D. Ehichioya, H. Ellerbrok, M.D Fernandez-Garcia, A. Fizet, E. Fleischmann, S. Gryseels, A. Hermelink, J. Hinzmann, U. Hopf-Guevara, Y. Ighodalo, L. Jameson, A. Kelterbaum, Z. Kis, S. Kloth, C. Kohl, M. Korva, A. Kraus, E. Kuisma, A. Kurth, B. Liedigk, C.H. Logue, A. Lüdtke, P. Maes, J. McCowen, S. Mély, M. Mertens, S. Meschi, B. Meyer, J. Michel, P. Molkenthin, C. Muñoz-Fontela, D. Muth, E.N.C. Newman, D. Ngabo, L. Oestereich, J. Okosun, T. Olorok, R. Omiunu, E. Omomoh, E. Pallasch, B. Pályi, J. Portmann, T. Pottage, C. Pratt, S. Priesnitz, S. Quartu, J. Rappe, J. Repits, M. Richter, M. Rudolf, A. Sachse, K.M. Schmidt, G. Schudt, T. Strecker, R. Thom, S. Thomas, E. Tobin, H. Tolley, J. Trautner, T. Vermoesen, I. Vitoriano, M. Wagner, S. Wolff, C. Yue, M.R. Capobianchi, B. Kretschmer, Y. Hall, J.G. Kenny, N.Y. Rickett, G. Dudas, C.E.M. Coltart, R. Kerber, D. Steer, C. Wright, F. Senyah, S. Keita, P. Drury, B. Diallo, H. de Clerck, M. Van Herp, A. Sprecher, A. Traore, M. Diakite, M.K. Konde, L. Koivogui, N. Magassouba, T. Avšič-Županc, A. Nitsche, M. Strasser, G. Ippolito, S. Becker, K. Stoecker, M. Gabriel, H. Raoul, A. Di Caro, R. Wölfel, P. Formenty, and S. Günther. Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. *Nature*, 524(7563):97–101, 2015.
- [18] D.T. Casto and P.A. Brunell. Safe Handling of Vaccines. *Pediatrics*, 87(1):108–112, 1991.
- [19] M. Cevik, N.D. Grubaugh, A. Iwasaki, and P. Openshaw. COVID-19 vaccines: Keeping pace with SARS-CoV-2 variants. *Cell*, 184(20):5077–5081, 2021.
- [20] C. Chen, Y. Chen, J. Liou, and M. Wu. From germ theory to germ therapy. *The Kaohsiung Journal of Medical Sciences*, 35(2):73–82, 2019.
- [21] Y. Cheng, N.B. Tjaden, A. Jaeschke, S.M. Thomas, and C. Beierkuhnlein. Deriving risk maps from epidemiological models of vector borne diseases: State-of-the-art and suggestions for best practice. *Epidemics*, page 100411, 2020.

- [22] G. Chowell, L. Sattenspiel, S. Bansal, and C. Viboud. Mathematical models to characterize early epidemic growth: A review. *Physics of Life Reviews*, 18:66–97, 2016.
- [23] A.K. Clift, C.A.C. Coupland, R.H. Keogh, K. Diaz-Ordaz, E. Williamson, E.M. Harrison, A. Hayward, H. Hemingway, P. Horby, N. Mehta, J. Bengler, K. Khunti, D. Spiegelhalter, A. Sheikh, J. Valabhji, R.A. Lyons, J. Robson, M.G. Semple, F. Kee, P. Johnson, S. Jebb, T. Williams, and J. Hippisley-Cox. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from Coronavirus 19 in adults: national derivation and validation cohort study. *British Medical Journal*, 371, 2020.
- [24] C. Cohen. The Initial and Daily COVID-19 Effective Reproduction Number in South Africa. Technical report, The National Institute For Communicable Diseases, May 2020.
- [25] C. Cohen. The Daily COVID-19 Effective Reproductive Number (R) in South Africa. Technical report, The National Institute For Communicable Diseases, September 2022.
- [26] C. Colombo and M. Diamanti. The smallpox vaccine: the dispute between Bernoulli and d’Alembert and the calculus of probabilities. *Lettera Matematica*, 2(4):185–192, 2015.
- [27] H. Connor. John graunt frs (1620-74): The founding father of human demography, epidemiology and vital statistics. *Journal of Medical Biography*, page 09677720221079826, 2022.
- [28] G.C. Critchfield, K.E. Willard, and D.P. Connelly. Probabilistic sensitivity analysis methods for general decision models. *Computers and Biomedical Research*, 19(3):254–265, 1986.
- [29] R.P. Curiel and H. G. Ramírez. Vaccination strategies against COVID-19 and the diffusion of anti-vaccination views. *Scientific Reports*, 11(1):1–13, 2021.
- [30] D.J. Daley and J. Gani. *Epidemic Modelling: An Introduction*. Number 15. Cambridge University Press, 2001.
- [31] T. Dhirasakdanon and H.R. Thieme. Persistence of Vertically Transmitted Parasite Strains which Protect against More Virulent Horizontally Transmitted Strains. In *Modeling and Dynamics of Infectious Diseases*, pages 187–215. World Scientific, 2009.
- [32] O. Diekmann and J.A.P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, volume 5. John Wiley & Sons, 2000.
- [33] K. Dietz and J.A.P. Heesterbeek. Daniel Bernoulli’s epidemiological model revisited. *Mathematical Biosciences*, 180(1-2):1–21, 2002.
- [34] P.J. Diggle. Spatio-Temporal Point Processes: Methods and Applications. *Monographs on Statistics and Applied Probability*, 107:1, 2006.

- [35] T. Dzinamarira, B. Nachipo, B. Phiri, and G. Musuka. COVID-19 Vaccine Roll-Out in South Africa and Zimbabwe: Urgent Need to Address Community Preparedness, Fears and Hesitancy. *Vaccines*, 9(3):250, 2021.
- [36] E.J. Elgar, A. Stefaniak, and M.J.A. Wohl. The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries. *Social Science & Medicine*, 263:113365, 2020.
- [37] L. Engelmann. A box, a trough and marbles: How the Reed-Frost epidemic theory shaped epidemiological reasoning in the 20th century. *History and Philosophy of the Life Sciences*, 43(3):1–24, 2021.
- [38] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae*, 6:290–297, 1959.
- [39] P. Erdős and A. Rényi. On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60, 1960.
- [40] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1):261–267, 1961.
- [41] W. Farr. Progress of Epidemics. *Second report of the Registrar General of England and Wales*, pages 16–20, 1840.
- [42] L. Faust, Y. Schreiber, and N. Bocking. A systematic review of BCG vaccination policies among high-risk groups in low TB-burden countries: implications for vaccination strategy in Canadian indigenous communities. *BMC Public Health*, 19(1):1–32, 2019.
- [43] Z. Feng, S. Towers, and Y. Yang. Modeling the Effects of Vaccination and Treatment on Pandemic Influenza. *The AAPS Journal*, 13(3):427–437, 2011.
- [44] G. Forni and A. Mantovani. COVID-19 vaccines: where we stand and challenges ahead. *Cell Death & Differentiation*, 28(2):626–639, 2021.
- [45] R. Fry, J. Hollinghurst, H.R. Stagg, D.A. Thompson, C. Fronterre, C. Orton, R.A. Lyons, D.V. Ford, A. Sheikh, and P.J. Diggle. Real-time spatial health surveillance: Mapping the UK COVID-19 epidemic. *International Journal of Medical Informatics*, 149:104400, 2021.
- [46] M.G. Garner and S.A. Hamilton. Principles of epidemiological modelling. *Revue Scientifique et Technique-OIE*, 30(2):407, 2011.
- [47] R. Gasparini, P.L. Amicizia, D. and Lai, and D. Panatto. Clinical and socioeconomic impact of seasonal and pandemic influenza in adults and the elderly. *Human Vaccines & Immunotherapeutics*, 8(1):21–28, 2012.
- [48] W.M. Getz, R. Salter, and W. Mgbara. Adequacy of SEIR models when epidemics have spatial structure: Ebola in Sierra Leone. *Philosophical Transactions of the Royal Society B*, 374(1775):20180282, 2019.

- [49] W.M. Getz, R. Salter, O. Muellerklein, H.S. Yoon, and K. Tallam. Modeling Epidemics: A Primer and Numerus Software Implementation. *bioRxiv*, page 191601, 2017.
- [50] I. Ghosh, P.K. Tiwari, S. Samanta, I.M. Elmojtaba, N. Al-Salti, and J. Chattopadhyay. A simple SI-type model for HIV/AIDS with media and self-imposed psychological fear. *Mathematical Biosciences*, 306:160–169, 2018.
- [51] J.E. Gnanvi, K.V. Salako, G.B. Kotanmi, and R.G. Kakaï. On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques. *Infectious Disease Modelling*, 6:258–272, 2021.
- [52] E. González and M.J. Villena. On the spatial dynamics of vaccination: A spatial SIRS-V model. *Computers & Mathematics with Applications*, 80(5):733–743, 2020.
- [53] J. Graunt. Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality. In *Mathematical Demography*, pages 11–20. Springer, 1977.
- [54] M.S. Green, E. Swartz, T. and Mayshar, B. Lev, A. Leventhal, P.E. Slater, and J. Shemer. When is an epidemic an epidemic? *The Israel Medical Association journal: IMAJ*, 4(1):3–6, 2002.
- [55] M. Greenwood. On the Statistical Measure of Infectiousness. *Epidemiology & Infection*, 31(3):336–351, 1931.
- [56] B.T. Grenfell, A. Kleczkowski, C.A. Gilligan, and B.M. Bolker. Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases. *Statistical Methods in Medical Research*, 4(2):160–183, 1995.
- [57] A. Hafeez, S. Ahmad, S.A. Siddqui, M. Ahmad, and S. Mishra. A review of COVID-19 (Coronavirus Disease-2019) diagnosis, treatments and prevention. *Eurasian Journal of Medicine and Oncology*, 4(2):116–125, 2020.
- [58] S.F. Harmer. The British Association. Section D. Zoology. *Nature*, 78(2029):488–503, 1908.
- [59] H.W. Hethcote. Qualitative analyses of communicable disease models. *Mathematical Biosciences*, 28(3-4):335–356, 1976.
- [60] H.W. Hethcote. An immunization model for a heterogeneous population. *Theoretical Population Biology*, 14(3):338–349, 1978.
- [61] H.W. Hethcote. The Mathematics of Infectious Diseases. *SIAM Review*, 42(4):599–653, 2000.
- [62] A.C. Hindmarsh. ODEPACK, a systematized collection of ode solvers. *IMACS Transactions on Scientific Computation*, 1:55–64, 1982.

- [63] S.H. Hodgson, K. Mansatta, G. Mallett, V. Harris, K.R.W. Emary, and A.J. Pollard. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. *The Lancet Infectious Diseases*, 21(2):e26–e35, 2021.
- [64] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, and X. Gu. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 395(10223):497–506, 2020.
- [65] D. Huremović. Brief History of Pandemics (Pandemics Throughout History). In *Psychiatry of Pandemics*, pages 7–35. Springer, 2019.
- [66] Z. Kimmie-S. Archibald P. Debba R. Manjoo-Docrat A. le Roux N. Dudeni-Tlhone C. Janse van Rensburg R. Thiede N. Abdelatif A. Potgieter S. Makhanya I. Fabris-Rotelli, J. Holloway. A spatial SEIR Model for COVID-19 in South Africa. *Journal of Data Science, Statistics and Visualisation*, 2022.
- [67] S.J. Ismail, L. Zhao, M.C. Tunis, S.L. Deeks, and C. Quach. Key populations for early COVID-19 immunization: preliminary guidance for policy. *Canadian Medical Association Journal*, 192(48):E1620–E1632, 2020.
- [68] S. Jalilisadrabad and E. Zabetian Targhi. Investigating the Location of Organizational Housing on the Outskirts of Cities (Study Sample: Accommodation of Bandar Abbas gas Condensate Refinery Staff). *Naqshejahan-Basic Studies and New Technologies of Architecture and Planning*, 10(1):19–31, 2020.
- [69] N.P. Jewell, J.A. Lewnard, and B.L. Jewell. Predictive Mathematical Models of the COVID-19 Pandemic. *Jama*, 323(19):1893–1894, 2020.
- [70] S. Johnson. *The Ghost Map: The Story of London's Most Terrifying Epidemic—and How It Changed Science, Cities, and the Modern World*. Penguin, 2006.
- [71] E.K. Vraga C.A. Miller P. B. Perrin C.W. Burton-M. Ryan B.F. Fuemmeler K.E. Carlyle. J.P.D. Guidry, L.I. Laesadius. Willingness to get the COVID-19 vaccine with and without emergency use authorization. *American Journal of Infection Control*, 49(2):137–142, 2021.
- [72] J. Jung. Preparing for the Coronavirus Disease (COVID-19) Vaccination: Evidence, Plans, and Implications. *Journal of Korean Medical Science*, 36(7), 2021.
- [73] A. Katzourakis. COVID-19: endemic doesn't mean harmless. *Nature*, pages 485–485, 2022.
- [74] W.O. Kermack and A.G. McKendrick. 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, 1927.
- [75] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics. II. – The problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 138(834):55–83, 1932.



- [76] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics. III. – further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 141(1933):94–112, 1933.
- [77] G. Kim, M. Kim, S.H. Ra, S. Lee, J. and Bae, J. Jung, and S. Kim. Clinical characteristics of asymptomatic and symptomatic patients with mild COVID-19. *Clinical Microbiology and Infection*, 26(7):948–e1, 2020.
- [78] L. Kong, M. Duan, J. Shi, J. Hong, Z. Chang, and Z. Zhang. Compartmental structures used in modeling COVID-19: a scoping review. *Infectious Diseases of Poverty*, 11(1):1–9, 2022.
- [79] A. Kumar, T. Choi, S. F. Wamba, S. Gupta, and K.H. Tan. Infection vulnerability stratification risk modelling of COVID-19 data: a deterministic SEIR epidemic model analysis. *Annals of Operations Research*, pages 1–27, 2021.
- [80] N. Lane. The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1666):20140344, 2015.
- [81] A. Le Roux, A.K. Cooper, C. Ludick, K.A. Arnold, and G. Mans. Creating a Set of High-Resolution Vulnerability Indicators to Support the Disaster Management Response to the COVID-19 pandemic in South Africa. In *COVID-19 Pandemic, Geospatial Information, and Community Resilience*, pages 291–304. CRC Press, 2021.
- [82] P. Lehohla. *Use of Health Facilities and Levels of Selected Health Conditions in South Africa: Findings from the General Household Survey, 2011*. Statistics South Africa Pretoria, 2013.
- [83] N. LePan. Visualizing the history of pandemics. *Visual Capitalist*, 14, 2020.
- [84] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490):489–493, 2020.
- [85] Ch. Lin, P. Tu, and L.M. Beitsch. Confidence and receptivity for COVID-19 vaccines: a rapid systematic review. *Vaccines*, 9(1):16, 2021.
- [86] I. Locatelli, B. Trächsel, and V. Rousson. Estimating the basic reproduction number for COVID-19 in Western Europe. *PLOS One*, 16(3):e0248731, 2021.
- [87] S.J. Lycett, F. Duchatel, and P. Digard. A brief history of bird flu. *Philosophical Transactions of the Royal Society B*, 374(1775):20180257, 2019.
- [88] J. Tao Y. L. Sun B.L. Dickens. M. Park, A.R. Cook. A systematic review of COVID-19 epidemiology based on current evidence. *Journal of Clinical Medicine*, 9(4):967, 2020.
- [89] N.A. Maidana and H.M. Yang. Spatial spreading of West Nile Virus described by traveling waves. *Journal of Theoretical Biology*, 258(3):403–417, 2009.

- [90] Mwelecele N Malecela and Camilla Ducker. A Road Map for Neglected Tropical Diseases 2021–2030. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 115(2):121–123, 2021.
- [91] S. Mandal, R.R. Sarkar, and S. Sinha. Mathematical models of malaria - a review. *Malaria Journal*, 10(1):1–19, 2011.
- [92] Y. Mardian, K. Shaw-Shaliba, M. Karyana, and C. Lau. Sharia (Islamic Law) perspectives of COVID-19 vaccines. *Frontier in Tropical Diseases*, 2, 2021.
- [93] G. Massonis, J.R. Banga, and A.F. Villaverde. Structural identifiability and observability of compartmental models of the COVID-19 pandemic. *Annual Reviews in Control*, 51:441–459, 2021.
- [94] Gallo M.B., G. Aghagoli, K. Lavine, L. Yang, E.J. Siff, S.S. Chiang, T.P. Salazar-Mather, L. Dumenco, M.C. Savaria, S.N. Aung, T. Flanagan, and I.C. Michelow. Predictors of COVID-19 severity: A literature review. *Reviews in Medical Virology*, 31(1):1–10, 2021.
- [95] X. Meng, Z. Cai, S. Si, and D. Duan. Analysis of epidemic vaccination strategies on heterogeneous networks: Based on SEIRV model and evolutionary game. *Applied Mathematics and Computation*, 403:126172, 2021.
- [96] T.U. Metcalf, R.A. Cubas, K. Ghneim, M.J. Cartwright, J. Van Grevenynghe, J.M. Richner, D.P. Olagnier, P.A. Wilkinson, M.J. Cameron, B.S. Park, J.B. Hiscott, M.S. Diamond, A.M. Wertheimer, J. Nikolich-Zugich, and E.K. Haddad. Global analyses revealed age-related alterations in innate immune responses after stimulation of pathogen recognition receptors. *Aging Cell*, 14(3):421–432, 2015.
- [97] M. Mhango, I. Chitungo, and T. Dzinamarira. COVID-19 Lockdowns: Impact on Facility-Based HIV Testing and the Case for the Scaling up of Home-Based Testing Services in Sub-Saharan Africa. *AIDS and Behavior*, 24:3014–3016, 2020.
- [98] A.K. Mitra and A.R. Mawson. Neglected Tropical Diseases: Epidemiology and Global Burden. *Tropical Medicine And Infectious Disease*, 2(3):36, 2017.
- [99] P. Montagnon. A stochastic SIR model on a graph with epidemiological and population dynamics occurring over the same time scale. *Journal of Mathematical Biology*, 79(1):31–62, 2019.
- [100] J.L. Moore, S. Liang, . Akullian, and J.V. Remais. Cautioning the use of degree-day models for climate change projections in the presence of parametric uncertainty. *Ecological Applications*, 22(8):2237–2247, 2012.
- [101] Z. Mukandavire, F. Nyabadza, N.J. Malunguza, D.F. Cuadros, T. Shiri, and G. Musuka. Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios. *PLOS One*, 15(7):e0236003, 2020.

- [102] S. Munzert, P. Selb, A. Gohdes, L.F. Stoetzer, and W. Lowe. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour*, 5(2):247–255, 2021.
- [103] F. Kirsebom S. Toffa T. Rikeard E. Gallagher C. Gower M Kall N. Groves M.Sc.-A. O’Connell D. Simons P.B. Blomquist A. Zaidi S. Nash N.I. Binti A. Aziz S. Thelwall G. Dabrera R. Myers G. Amirthalingam S. Gharbia J.C. Barrett R. Elson S.N. Ladhani N. Ferguson M. Zambon C.N.J. Campbell K. Brown S. Hopkins M. Chand M. Ramsay J. Lopez Bernal N. Andrews, J. Stowe. Covid-19 vaccine effectiveness against the Omicron (B. 1.1. 529) variant. *New England Journal of Medicine*, 386(16):1532–1546, 2022.
- [104] G. Nedjati-Gilani N. Imai K. Ainslie M. Baguelin S. Bhatia A. Boonyasiri Z. Cucunubá G. Cuomo-Dannenburg A. Dighe I. Dorigatti H. Fu K. Gaythorpe W. Green A. Hamlet W. Hinsley L. C. Okell S. van Elsland H. Thompson R. Verity E. Volz H. Wang Y. Wang P.G.T. Walker C. Walters P. Winskill C. Whittaker C. A Donnelly S. Riley A.C. Ghani N. M Ferguson, D. Laydon. Impact of non-pharmaceutical interventions (NPI’s) to reduce COVID-19 mortality and healthcare demand. 2020.
- [105] M.Z. Ndi and A.K. Supriatna. Stochastic Mathematical Models in Epidemiology. *Information*, 20:6185–6196, 2017.
- [106] D. Ndwandwe and C.S. Wiysonge. COVID-19 vaccines. *Current Opinion in Immunology*, 71:111–116, 2021.
- [107] J.P. Newhouse and A.M. Garber. Geographic variation in medicare services. *New England Journal of Medicine*, 2013.
- [108] C.N. Ngonghala, E. Iboi, S. Eikenberry, M. Scotch, C.R. MacIntyre, M.H. Bonds, and A.B. Gumel. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Mathematical Biosciences*, 325:108364, 2020.
- [109] V.K.M. Niyas and R. Arjun. Breakthrough COVID-19 infections among health care workers after two doses of Chadox1 nCoV-19 vaccine. *QJM: An International Journal of Medicine*, 114(10):757–758, 2021.
- [110] B. Nogrady. What the data say about asymptomatic COVID infections. *Nature*, 587(7835):534–536, 2020.
- [111] F. Nyabadza, F. Chirove, C.W. Chukwu, and M.V. Visaya. Modelling the Potential Impact of Social Distancing on the COVID-19 Epidemic in South Africa. *Computational and Mathematical Methods in Medicine*, 2020, 2020.
- [112] D.A. O’Brien and C.F. Clements. Early warning signal reliability varies with COVID-19 waves. *Biology Letters*, 17(12):20210487, 2021.
- [113] C. Okaïs, S. Roche, M. Kürzinger, B. Riche, H. Bricout, T. Derrough, F. Simondon, and R. Ecochard. Methodology of the sensitivity analysis used for modeling an infectious disease. *Vaccine*, 28(51):8132–8140, 2010.

- [114] P. Olliaro. What does 95% COVID-19 vaccine efficacy really mean? *The Lancet Infectious Diseases*, 21(6):769, 2021.
- [115] World Health Organisation. Estimating cases for COVID-19 in South Africa. Assessment of alternative scenarios. <https://www.who.int/publications/m/item/draft-landscape-of-COVID-19-candidate-vaccines>.
- [116] F.B. Osei and A. Stein. Temporal trend and spatial clustering of cholera epidemic in Kumasi-Ghana. *Scientific Reports*, 8(1):1–11, 2018.
- [117] T. Pillay. Clinical management of suspected or confirmed COVID-19 disease. Technical report, The National Institute For Communicable Diseases, 2022.
- [118] J. Piret and G. Boivin. Pandemics Throughout History. *Frontiers in Microbiology*, 11:631736, 2021.
- [119] C. M Pooley, A.B. Doeschl-Wilson, and G. Marion. Estimation of age-stratified contact rates during the COVID-19 pandemic using a novel inference algorithm. *medRxiv*, 2022.
- [120] A. Potgieter, I. Fabris-Rotelli, Z. Kimmie, N. Dudeni-Tlhone, J. Holloway, C. Janse Van Rensburg, R. Thiede, P. Debba, N. Docrat, R. and Abdelatif, and S. Khuluse-Makhanyam. Modelling Representative Population Mobility for COVID-19 Spatial Transmission in South Africa. *Frontiers in Big Data*, 4:718351, 2021.
- [121] A. Putrino, M. Raso, C. Magazzino, and G. Galluccio. Coronavirus (covid-19) in Italy: knowledge, management of patients and clinical experience of Italian dentists during the spread of contagion. *BMC Oral Health*, 20(1):1–15, 2020.
- [122] W. Zhang L. Wang M. Li J. Shi Y. Zhai D. Sun L. Wang B. Chen S. Jiang Q. Ma, J. Gao and J. Zhao. Spatio-temporal distribution characteristics of COVID-19 in China: a city-level modeling study. *BMC Infectious Diseases*, 21(1):1–14, 2021.
- [123] M Ramsay. COVID-19: The Green Book, chapter 14a. Technical report, 2020. Immunisation against infectious diseases: Public Health England.
- [124] A. Rosenberg, D.E. Keene, P. Schlesinger, A.K. Groves, and K.M. Blankenship. COVID-19 and hidden housing vulnerabilities: implications for health equity, New Haven, Connecticut. *AIDS and Behavior*, 24:2007–2008, 2020.
- [125] M. Roskosky, M. Ali, S.R. Upreti, and D. Sack. Spatial clustering of Cholera cases in the Kathmandu Valley: implications for a ring vaccination strategy. *International Health*, 13(2):170–177, 2021.
- [126] R. Ross. *The Prevention of Malaria*. John Murray, 1911.
- [127] G. Röst, F.A. Bartha, N. Bogya, P. Boldog, A. Dénes, T. Ferenci, K.J. Horváth, A. Juhász, C. Nagy, T. Tekeli, and B. Vizi, Z. and Oroszi. Early Phase of the COVID-19 Outbreak in Hungary and Post-Lockdown scenarios. *Viruses*, 12(7):708, 2020.

- [128] A.B. Sadavare and R.V. Kulkarni. A Review of Application of Graph Theory for Network. *International Journal of Computer Science and Information Technologies*, 3(6):5296–5300, 2012.
- [129] A. Safarishahrbijari, T. Lawrence, R. Lomotey, J. Liu, C. Waldner, and N. Osgood. Particle filtering in a SEIRV simulation model of H1N1 influenza. In *2015 Winter Simulation Conference (WSC)*, pages 1240–1251. IEEE, 2015.
- [130] B.D. Schoub. Dial down the rhetoric over COVID-19 vaccines. *South African Medical Journal*, 111(6):522–523, 2021.
- [131] S.K. Seaholm, E. Ackerman, and S. Wu. Latin hypercube sampling and the sensitivity analysis of a Monte Carlo epidemic model. *International Journal of Bio-medical Computing*, 23(1-2):97–112, 1988.
- [132] S. Shankar, S. Sourabh Mohakuda, K.A., P.S. Nazneen, A.K. Yadav, and K. Chatterjee. Systematic review of predictive mathematical models of COVID-19 epidemic. *Medical Journal Armed Forces India*, 77:S385–S392, 2021.
- [133] S. Silal, J. Pulliam, G. Meyer-Rath, B. Nichols, L. Jamieson, Z. Kimmie, and H. Moultrie. Estimating cases for COVID-19 in South Africa Update: 19 May 2020. "[https://www.nicd.ac.za/wp-content/uploads/2020/05/SACMC\\_19052020\\_slides-for-MoH-media-briefing.pdf](https://www.nicd.ac.za/wp-content/uploads/2020/05/SACMC_19052020_slides-for-MoH-media-briefing.pdf)", 2020.
- [134] W. Slauter. Write Up Your Dead: The Bills of Mortality and the London Plague of 1665. *Media History*, 17(1):1–15, 2011.
- [135] A.J. Smit, J.M. Fitchett, F.A. Engelbrecht, R.J. Scholes, G. Dzhivhuho, and N.A. Sweijd. Winter Is Coming: A Southern Hemisphere Perspective of the Environmental Drivers of SARS-CoV-2 and the Potential Seasonality of COVID-19. *International Journal of Environmental Research and Public Health*, 17(16):5634, 2020.
- [136] J. Snow. *On the Mode of Communication of Cholera*. John Churchill, 1855.
- [137] M.H. Suleman and D.E. Lucero-Prisno. South Africa's COVID-19 vaccine rollout amid the emergence of Omicron. *Population Medicine*, 4(January):1–2, 2022.
- [138] B. Tang, N. L. Bragazzi, Q. Li, S. Tang, Y. Xiao, and J. Wu. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infectious Disease Modelling*, 5:248–255, 2020.
- [139] M.W. Tenforde, W.H. Self, K. Adams, M. Gaglani, A.A. Ginde, T. McNeal, S. Ghamande, D.J. Douin, H.K. Talbot, J.D. Casey, N.M. Mohr, A. Zepeski, N.I. Shapiro, K.W. Gibbs, D.C. Files, D.N. Hager, A. Shehu, M.E. Prekker, H.L. Erickson, M.C. Exline, M.N. Gong, A. Mohamed, D.J. Henning, J.S. Steingrub, I.D. Peltan, S.M. Brown, E.T. Martin, A.S. Monto, A. Khan, C.L. Hough, L.W. Busse, C.C. ten Lohuis, A. Duggal, J.G. Wilson, A.J. Gordon, N. Qadir, S.Y. Chang, C. Mallow, C. Rivas, H. M. Babcock, J.H. Kwon, N. Halasa, J.D.

- Chappell, A.S. Luring, C.G. Grijalva, T.W. Rice, I.D. Jones, W.B. Stubblefield, A. Baughman, K.N. Womack, J.P. Rhoads, C.J. Lindsell, K.W. Hart, Y. Zhu, S.M. Olson, M. Kobayashi, J.R. Verani, and M.M. Patel. Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. *Jama*, 326(20):2043–2054, 2021.
- [140] R. Thiede, N. Abdelatif, I. Fabris-Rotelli, R. Manjoo-Docrat, J. Holloway, C. Janse van Rensburg, P. Debba, N. Dudeni-Tlhone, Z. Kimmie, and A. le Roux. Spatial variation in the basic reproduction number of COVID-19: A systematic review. *arXiv preprint arXiv:2012.06301*, 2020.
- [141] Herbert CS Thom. A Note On The Gamma Distribution. *Monthly weather review*, 86(4):117–122, 1958.
- [142] M.J. Tildesley, T.A. House, M.C. Bruhn, R.J. Curry, M. OâNeil, J.L.E. Allpress, G. Smith, and M.J. Keeling. Impact of spatial clustering on disease transmission and optimal control. *Proceedings of the National Academy of Sciences*, 107(3):1041–1046, 2010.
- [143] Ivan Turok, Justin Visagie, and Andreas Scheba. Social Inequality and Spatial Segregation in Cape Town. In *Urban Socio-Economic Segregation and Income Inequality*, pages 71–90. Springer, Cham, 2021.
- [144] D.A.J. Tyrrell and S.H. Myint. Coronaviruses. *Medical Microbiology. 4th edition*, 1996.
- [145] E.O. Voit, H.A. Martens, and S.W. Omholt. 150 Years of the Mass Action Law. *PLOS Computational Biology*, 11(1):e1004012, 2015.
- [146] J. Wang, R. Jing, H. Lai, X. and Zhang, Y. Lyu, M.D. Knoll, and H. Fang. Acceptance of COVID-19 Vaccination during the COVID-19 Pandemic in China. *Vaccines*, 8(3):482, 2020.
- [147] J.S. Weitz and J. Dushoff. Modeling Post-death Transmission of Ebola: Challenges for Inference and Opportunities for Control. *Scientific Reports*, 5(1):1–7, 2015.
- [148] C.J.M. Whitty. What makes an academic paper useful for health policy? *BMC Medicine*, 13(1):1–5, 2015.
- [149] R.R. Wilkinson, F.G. Ball, and K.J. Sharkey. The deterministic Kermack-McKendrick model bounds the general stochastic epidemic. *Journal of Applied Probability*, 53(4):1031–1040, 2016.
- [150] S. Wood. mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation. 2012.
- [151] S.E. Woolf and J.W. Joubert. A people-centred view on paratransit in South Africa. *Cities*, 35:284–293, 2013.
- [152] S.E. Woolf and J.W. Joubert. A look at paratransit in South Africa. Technical report, University of Pretoria, South Africa, 2014.
- [153] J. Wu, R. Dhingra, M. Gambhir, and J.V. Remais. Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of The Royal Society Interface*, 10(86):20121018, 2013.

- [154] J. Yang, W. Zheng, H. Shi, X. Yan, K. Dong, Q. You, G. Zhong, H. Gong, Z. Chen, M. Jit, C. Viboud, M. Ajelli, and H. Yu. Who should be prioritized for COVID-19 vaccination in China? A descriptive study. *BMC Medicine*, 19(1):1–13, 2021.
- [155] J. Zheng, Y. Zhou, H. Wang, and X. Liang. The role of the China Experts Advisory Committee on Immunization Program. *Vaccine*, 28:A84–A87, 2010.
- [156] S. Zhou, S. Zhou, Z. Zheng, and J. Lu. Optimizing Spatial Allocation of COVID-19 Vaccine by Agent-Based Spatiotemporal Simulations. *GeoHealth*, page e2021GH000427.