Multi Variant Spatially Informed Rapid Testing for Epidemic Model

by

Allen Moncey Varghese

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2022 by the
Graduate Supervisory Committee:

Giulia Pedrielli, Chair
K. Selçuk Candan
Teresa Wu

ARIZONA STATE UNIVERSITY

August 2022

ABSTRACT

The COVID-19 outbreak that started in 2020, brought the world to its knees and is still a menace after three years. Over eighty-five million cases and over a million deaths have occurred due to COVID-19 during that time in the United States alone.

A great deal of research has gone into making epidemic models to show the impact of the virus by plotting the cases, deaths, and hospitalization due to COVID-19. However, there is very less research that has anything to do with mapping different variants of COVID-19. SARS-CoV-2, the virus that causes COVID-19, constantly mutates and multiple variants have emerged over time. The major variants include Beta, Gamma, Delta and the recent one, Omicron.

The purpose of the research done in this thesis is to modify one of the epidemic models i.e., the Spatially Informed Rapid Testing for Epidemic Model (SIRTEM), in such a way that various variants of the virus will be modelled at the same time. The model will be assessed by adding the Omicron and the Delta variants and in doing so, the effects of different variants can be studied by looking at the positive cases, hospitalizations, and deaths from both the variants for the Arizona Population. The focus will be to find the best infection rate and testing rate by using Random numbers so that the published positive cases and the positive cases derived from the model have the least mean square error.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my Committee Chair and Advisor, Dr Giulia Pedrielli, for her complete support to make me deliver my best. She offered me many opportunities to present my ideas and research which facilitated my academic growth. It has been a privilege to work under her supervision.

In addition, I would like to extend my thanks to other professors and students who were directly or indirectly involved in our research: Dr K. Selçuk Candan, Ms Fahim Tasneema Azad, Mr Robert W. Dodge and Mr Jaejin Lee. Their constant guidance and willingness to share practical knowledge made me understand the practical and theoretical working of the research.

I sincerely thank all of my committee members, Dr K. Selçuk Candan and Dr Teresa Wu, for their time, suggestions, and encouragement.

Finally, I would like to thank my friends and family, especially my mother and brother, for always motivating me throughout this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Background and Motivation

COVID -19 which started on December 31st, 2019 has changed the world drastically. Over the past three years, COVID-19 has had more than eighty-five million cases and about one million deaths. Not only did it cause one of the biggest pandemics in recent history, it had severe effect on the world economy.

To curb this, researchers created epidemic models to show the impact of the virus by forecasting future cases and deaths that can occur due to COVID-19. SEIR, which stands for susceptible, exposed, infected and recovered, is one of the most common epidemic models researchers use. SIRTEM is a modified version of SEIR that includes multiple compartments: testing, isolation, quarantine, hospitalization and immunization combined with the four compartments of SEIR. However, these models do not incorporate various variants. SARS-CoV-2, the virus that causes COVID-19, constantly mutates, and multiple variants have emerged like Beta, Gamma, Delta and the recent one, Omicron.

Our research will focus on formulating the problem of multiple variants and understanding the impact of different variants in the population. Our research goal is to redesign the SIRTEM, which will encompass various variants. The focus is to assess the effects of different variants on the number of cases, deaths and hospitalizations.

1.2 Challenges and Outcomes

The revised model presents the following new challenges:

1. **Amend the underlying equations:** The equations in the SIRTEM model should be altered to add multiple variants. There is limited knowledge about epidemic models, which have several variants. There are different ways to do the implementation, and selection of the best one is required.

2. **Calibration for different variants:** Each variant has unique features, including testing, hospitalization and death rates. This poses a challenge for the mathematical analysis, as modeling the variant's characteristics requires detailed research. Different predictive models need to be considered for this.

3. **Algorithm for simulation parameters:** A method for tuning the simulation parameters like the infection and testing rates needs to be created. The technique used in SIRTEM can be used but needs to be modified to simultaneously tune for various variants. Other parameters that need to be checked are the hospitalization and death rates.

1.3   Thesis Structure

The content of thesis is organized as follows :

1. **Chapter 2:** Presents the epidemic modeling and simulation that is present right now. It also shows the gaps in the literature.

2. **Chapter 3:** Describes the methodology and design for the modified SIRTEM which includes multiple variants. Fine tuning of the parameters will be done in this chapter as well.

3. **Chapter 4:** Presents the numerical analysis from the model.

4. **Chapter 5:** Conclusions and directions for future work are presented.

Chapter 2

LITERATURE REVIEW

In this chapter, the relevant literature is reviewed to understand the different epidemic models. For the same, a variety of topics from epidemic modeling in general and its different examples is covered.

2.1 Epidemiological Modeling

Epidemiological modeling is playing a critical role right now because of COVID-19. Policies are being carried out from the information gained from these models. Epidemiological modeling can be dated back to the 1700s, when Daniel Bernoulli devised one such model for the smallpox disease. Since then, epidemiological modelling research has advanced a lot, especially during the COVID-19 pandemic. The basic types of epidemiological modeling are (i) Stochastic (Random), (ii) Deterministic [17].

Stochastic (Random) Models

These models are essential when the infectious population is tiny. Stochastic models always converge to a state where the disease is eradicated from the people. The REED-Frost model is one of the first stochastic models created but was published later on [7; 16]. The Reed Frost model is one of the simplest stochastic epidemic models. Each infected individual at a time t independently infects individuals in the susceptible population with a probability p. The population that become infected by the individuals then makes up the infected population at time t +1 and the infected individuals at time t are removed from the epidemic process. The Greenwood Model was also developed during the same time [20]. Both of them are bi-variate Markov

models, but they differ in assuming the probability of infection. The Greenwood method assumes $p_i = p$ where p is constant, and the Reed Frost method assumes $p_i = p^i$ which changes after each time step.

Another set of stochastic models is the branching epidemic processes [9]. One of the simplest examples of the branching epidemic process is Galton–Watson [29]. This method assumes R_0 individuals from the susceptible population get infected by each infected individual. It also assumes that each infected individual is independent of all infected individuals. The probability generating function for the number of new infections is $f(t) = \sum_{k=0}^{\infty} p_k \cdot t^k$ where $f'(1) = R_0$.

Stochastic models can be extended to deterministic models like SIR, SIS, SEIR,etc [25; 23; 9; 30; 28; 26; 11; 15; 27; 19; 18; 21; 12]. These deterministic models are reviewed in the next subsection.

Deterministic Models

In deterministic models, the number of infected can be set as real numbers, which are less than one and greater than zero (so that the virus can spread). Kermack and McKendrick constructed a straightforward model called the Susceptible-Infected-Recovered (SIR) model that is still used for developing more complex epidemic models [22] which is shown in Figure 2.1. In SIR, an individual who is susceptible gets infected and moves to the infected compartment. After a few days, the individual can recover and move onto the recovery compartment. Another model is the Susceptible-Infected-Susceptible (SIS), where after an individual recovers, they go back to being susceptible, which is shown in Figure 2.2.

The major deterministic model that is used is the Susceptible-Exposed-Infected-Recovered (SEIR) model which is shown in Figure 2.3[5; 10; 24; 8]. The individual moves from susceptible to exposed compartment after being exposed to an infected

individual. There is a chance that even if exposed, an individual cannot be infected, and they go back to being susceptible. Spatially Informed Rapid Testing for Epidemic Model (SIRTEM) is a modified SEIR model, which is shown in Figure 2.4 ([13]) which is what we will use to make the revised model for multiple variants for this thesis.
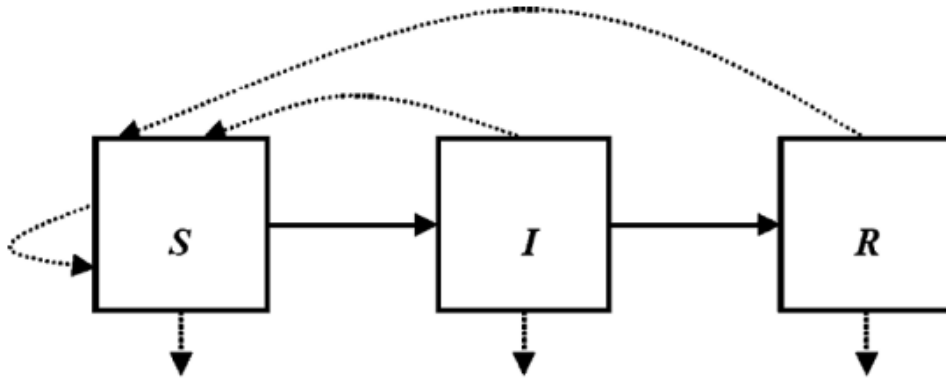
Figure 2.1: SIR Compartmental Diagram. Source - [9]
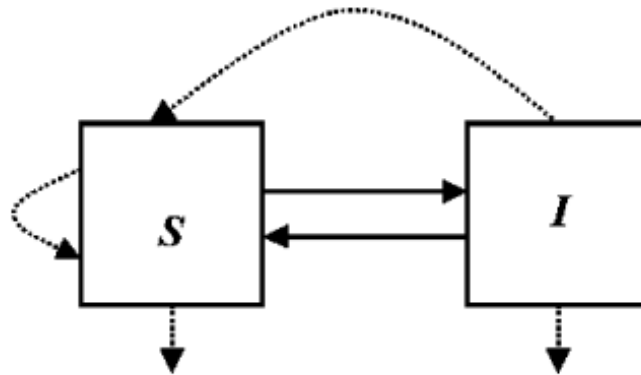
Figure 2.2: SIS Compartmental Diagram. Source - [9]
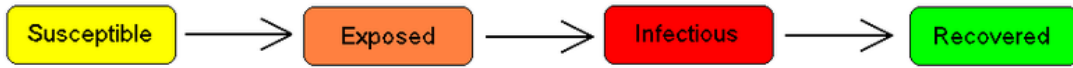
Figure 2.3: SEIR Compartmental Diagram. Source - [1]

2.2 Literature Gap

We found abundant research work on epidemic models that can forecast and predict the positive cases, hospitalizations and deaths of a disease as a whole. However, there is not much research done in finding and predicting the positive cases, hospitalizations and deaths for each disease variant.

We propose a modified version of SIRTEM which will include multiple variants simultaneously. We believe that mapping different variants can make better policies to prevent hospitalizations and deaths.
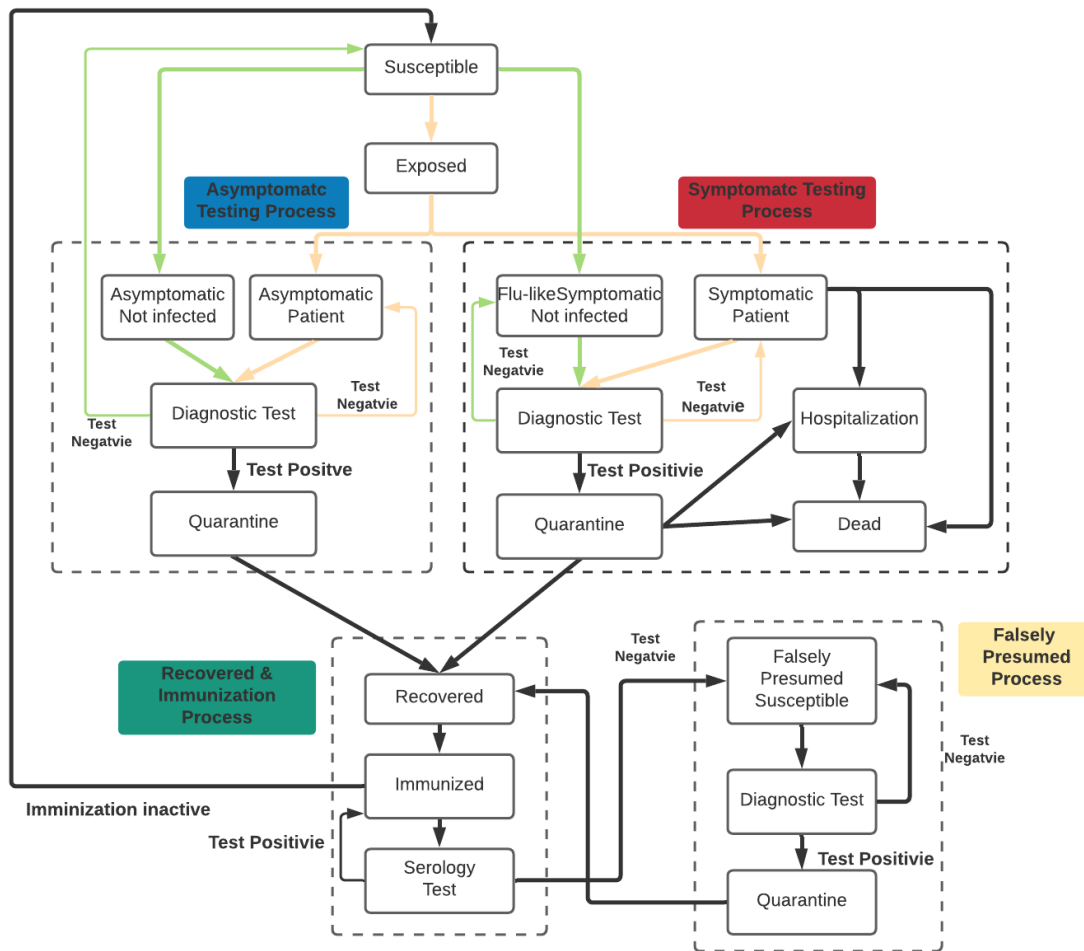
Figure 2.4: Overview of the SIRTEM Coupled Epidemic/Testing Model (Components, Sub-components, and Transitions). Source - [13]

Chapter 3

METHODOLOGY

The main goal of this research is to modify the existing SIRTEM for single city so that it can model n different variants at the same time. We will start by checking where the model needs to be split so that it goes to n separate sets of compartments. Another goal is make a custom algorithm to calibrate the simulation parameters of the model which is similar to the one used by produced in the SIRTEM model.

3.1 Multiple Variant SIRTEM (MVSIRTEM)
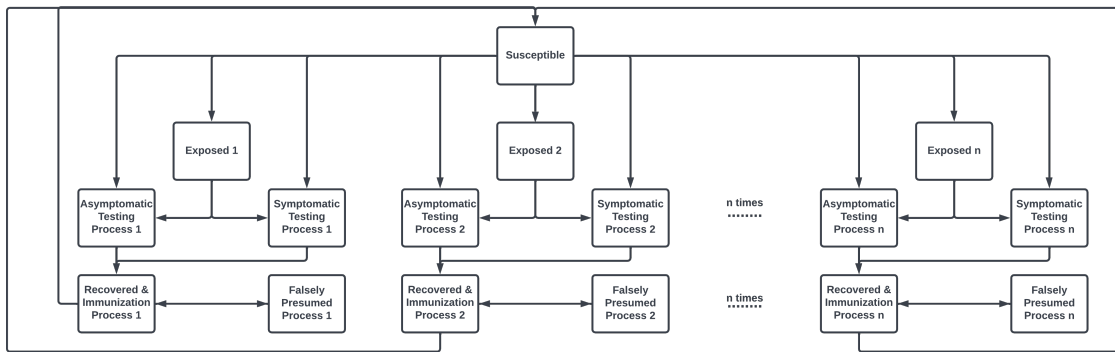
Figure 3.1: Overview of the Modified SIRTEM Coupled Epidemic/Testing Model for Simplicity, the Figure Ignores the Sub Components

Figure 3.1 shows the modified SIRTEM model with the addition of separate sets of compartments for n different variants. Let n be the number of variants of COVID-19. The model starts with a common susceptible population where no one is infected. An individual from this component can proceed to n different exposed components, one for each of the variants based. Then the individual goes onto the testing phase

9

which is different for each variant. After that the individual moves onto the recovered and immunization components. From there the individual can proceed to the falsely presumed components or back to the common susceptible population. An assumption is made that the testing rate for each of the variants will be different depending on the severity of the variant. Another assumption that is made is that testing allows us to determine the variant the patient is affected by.

3.2 Compartment Details

In the following subsections, we provide detailed description about each of the sets of compartments. The set of compartments in the model are Susceptible, Exposed, Asymptomatic Testing Process, Symptomatic Testing Process, Recovered and Immunization Process, and Falsely Presumed Process.

Susceptible

Susceptible is the compartment that houses the population that is non infected and does not have any symptoms. The susceptible population can be exposed to n different types of COVID-19 variants and they move onto the respective exposed compartment.

Exposed

The exposed compartment for each variant is controlled by the following equations:

$$\frac{dE_v}{dt} = \left(\beta_v \cdot S(t) + \beta'_v \cdot FPI_v(t) \right) \cdot \frac{Infected_v(t)}{N} - \left(per_a + per_s \right) \cdot E_v(t) \qquad (3.1)$$

where the total number of infected individuals at time t is,

$$Infected_v(t) = r \cdot \left(PS_v(t) + PA_v(t) + IA_v(t) + ATN_v(t) \right) + IS_v(t) + STN_v(t) \qquad (3.2)$$

Table 3.1: MVSIRTEM Model Parameters. Source - [13]

Parameters	Description
tp	Sensitivity of diagnostic test
tn	Specificity of diagnostic test
τ	Time to obtain the result for diag. test (days)
ϕ_v	Testing rate for diagnostic test the symptomatic population (ratio per day) for variant v of the virus
ϕ_{av}	Testing rate for diagnostic test the non-symptomatic population (ratio per day) for variant v of the virus
tp_{se}	Sensitivity of the serology test
tn_{se}	Specificity of the serology test
τ_{se}	Time to obtain the result for the serology test (days)
ϕ_{se}	Testing rate for the serology test (ratio of the relevant population per day)
β_v	Infection rate for the susceptible population for variant v (ratio)
β'_v	Infection rate for the population of individuals who are falsely presumed immune for variant v(ratio)
r	The ratio of the transmission rate of asymptomatic individuals to the transmission rate of symptomatic individuals
per_a	Percentage of individuals with COVID-19 who are asymptomatic
per_s	Percentage of individuals with COVID-19 who are symptomatic
η	Incubation length (days)
λ_a	Length of recovery for asymptomatic individuals (days)
λ_s	Length of recovery for symptomatic individuals (days)
λ_q	Length of quarantine (days)
λ_h	Hospitalization length (days)
h_v	Hospitalization rate for individuals with variant v(ratio of the quarantined population, per day)
κ_v	Mortality rate for symptomatic population with variant v(per day)
κ_{hv}	Mortality rate for the hospitalized individuals with variant v(per day)
g	Ratio of the susceptible individuals who has fever and cough for non-COVID infections (ratio, per day)

The term $Infected_v(t)$ here is the sum of the compartments that are having individuals with infections brought by variant v at time t. β_v, represents the number of new daily infection produced by a single infectious individual with the variant v and it is equivalent to the product of contact rate and disease transmission rate of variant v. Individuals move from Susceptible to Exposed compartment of variant v at the rate of $\beta_v \cdot Infected_v(t) \cdot \frac{S(t)}{N}$ where N is the total population. The parameter, β'_v, is a higher infection rate parameter, which is applied to the falsely presumed immune individuals (FPI) compartment (i.e., individuals that have a false positive

Table 3.2: Sub-compartments Relevant for the Asymptomatic Testing Process.
Source - [13]

Sub-Compartment		Sub-Compartment	
S	Susceptible population	E_v	Pop. exposed to the variant v of the virus
PA_v	Pre-asymptomatic population with variant v of the virus	IA_v	Infected pop. who are asymptomatic with variant v of the virus
AT_v	Asymp. pop. with variant v of the virus receiving diagnostic test	ATN_v	Asymptomatic pop. with variant v of the virus with negative test result
QAP_v	Asymp. pop. with variant v of the virus quarantined after a test	UR_v	Pop. with unknown immunity due to unknown infection
KR_v	Pop. of known recovered individuals from variant v of the virus	NT_v	Susceptible pop. with variant v of the virus receiving diagnostic test
NTN_v	Pop. of non-infected indiv. who test negative with variant v of the virus	NTQ_v	Non-infected pop. quarantined due to testing error with variant v of the virus
FPI_v	Pop. of indiv. with variant v of the virus who are falsely presumed immune	PS_v	Infected pop. who are pre-symptomatic with variant v of the virus
IS_v	Infected pop. who are symptomatic with variant v of the virus	STN_v	Symptomatic pop. with variant v of the virus who test negative (by error)

with the serology test). Finally, equation (3.2) shows that individuals can be infected and either can be asymptomatic or symptomatic.

Asymptomatic Testing Process

The asymptomatic testing process is used to randomly test individuals in susceptible and distinguish the infected asymptomatic from the susceptible population. The asymptomatic population contains the infected asymptomatic individuals with variant v and the susceptible population who are not infected. It is not easy to differentiate the individuals in both the populations. Table 3.2 shows the sub-compartments that are in the asymptomatic testing process. An assumption is made that the asymptomatic population consists individuals infected with variant v who are asymptomatic.

An individual exposed to variant v of the virus will become an infected asymptomatic individual at the rate of per_a. The virus will incubate for η days and the individual will be assumed to be pre-asymptomatic during that time. After η days, if the individual has symptoms, he/she will move onto to be part of the infected asymptomatic population:

$$\frac{dPA_v}{dt} = per_a \cdot E_v(t) - PA_v(t - \eta) \tag{3.3}$$

The Infected asymptomatic population of variant v can either increase or decrease depending on how many pre-asymptomatic individuals become asymptomatic and when they recover from the virus. The random testing for the asymptomatic population is done at the rate of $\phi_{av} \cdot IA_v(t)$.

$$\frac{dIA_v}{dt} = PA_v(t - \eta) + ATN_v(t) - \phi_{av} \cdot IA_v(t) - \lambda \cdot IA_v(t) \tag{3.4}$$

λ is assumed as the rate at which individuals recover naturally or without care after being infected with variant v and gains immunity.

$$\frac{dUR_v}{dt} = \lambda \cdot IA_v(t) + \lambda \cdot IS_v(t) - UR_v(t) \tag{3.5}$$

Individuals who are asymptomatic can get randomly tested. The following equation shows this:

$$\frac{dAT_v}{dt} = \phi_{av} \cdot IA_v(t) - AT_v(t - \tau) \tag{3.6}$$

Differentiating the susceptible population with the infected asymptomatic population with variant v can't be done. Therefore random testing is done also to the susceptible population:

$$\frac{dNT_v}{dt} = \phi_{av} \cdot S(t) - NT_v(t - \tau) \tag{3.7}$$

The specificity, tn, is the ability to correctly identify individual without the virus:

$$\frac{dNTN_v}{dt} = tn \cdot NT_v(t - \tau) - NTN_v(t) \tag{3.8}$$

From the asymptomatic compartment, individuals leave after testing negative:

$$\frac{dATN_v}{dt} = (1 - tp) \cdot AT_v(t - \tau) - ATN_v(t) \tag{3.9}$$

If an individual from asymptomatic tests positive, the asymptomatic individual is quarantined for λ_q days:

$$\frac{dQAP_v}{dt} = tp \cdot AT_v(t - \tau) - QAP_v(t - \lambda_q) \tag{3.10}$$

If a test result is falsely positive, a non infected susceptible individual will be, falsely, quarantined for λ_q days:

$$\frac{dNTQ_v}{dt} = (1 - tn) \cdot NT_v(t - \tau) - NTQ_v(t - \lambda_q) \tag{3.11}$$

After the end of the quarantine period, the individual can recover:

$$\frac{dKR_v}{dt} = QAP_v(t - \lambda_q) + FSQ_v(t - \lambda_q) + QSP_v(t - \lambda_q) + tn_1 \cdot HT_v(t - \tau) - KR_v(t) \tag{3.12}$$

The above equation considers other populations as well who leave quarantine. In particular, FSQ_v denotes those individuals who are falsely presumed susceptible and thus wrongly quarantined for variant v (Table 3.5), QSP_v denotes symptomatic individuals who are quarantined with positive test results from variant v (Table 3.3), and HT_v denotes portion of the population who received testing while being hospitalized (Table 3.3).

Table 3.3: Sub-compartments Relevant for the Symptomatic Testing Process.

Source - [13]

	Sub-Compartment		Sub-Compartment
E_v	Pop. exposed to the variant v of the virus	PS_v	Pre-Symptomatic population with variant v
IS_v	Infected pop. who are symptomatic with variant v	ST_v	Symptomatic pop. receiving test with variant v
STN_v	Symptomatic pop. with negative test result of variant v	QSP_v	Symp. pop. quarantined after a test with variant v
UR_v	Pop. with unknown immunity due to unknown infection	KR_v	Pop. of known recovered individuals from variant v
HBQ_v	Pop. need hospitalization from variant v before a quarantine	HDQ_v	Pop. with variant v need hospitalization during quarantine
HBT_v	Portion of HBQ tested with variant v while hospitalized	HDT_v	Portion of HDQ tested with variant v while hospitalized
FS	Pop. showing flu symptoms	FT_v	Pop. with flu symptom receiving test
FTN	Pop. with flu symptom tested negative for COVID-19	FTQ	Pop. with flu symptom quarantined due to false positive
GS_v	Pop. with other COVID-like symptoms	GT_v	Pop. with other COVID-like symptoms receiving test
GTN	Pop. with other symptoms tested negative for COVID-19	GTQ	Pop. with other symptoms quarantined due to false positive
D_v	Pop. who have not recovered from the infection with variant v (dead)		

Symptomatic Testing Process

The symptomatic testing process is designed to model the testing process for individuals who show COVID-19 like symptoms. Symptomatic individuals is separated into three populations: (a) COVID-infected symptomatic with variant v, (b) general sick (fever, coughing), and (c) flu symptomatic. They are differentiated using diagnostic testing.

Much like the asymptomatic process discussed before, the process consists of testing, isolation, and quarantine sub-processes. But also includes hospitalization and death for severe cases. The compartments presented in Table 3.3 along with the following differential equations define the transitions between relevant states in the symptomatic process.

An individual exposed to variant v of the virus will become an infected symptomatic individual at the rate of per_s rate. The virus will incubate for η days and the individual will be assumed to be pre-symptomatic during that time. After η days, if the individual has symptoms, he/she will move onto to be part of the infected

symptomatic population:

$$\frac{dPS_v}{dt} = per_s \cdot E_v(t) - PS_v(t - \eta) \tag{3.13}$$

Similar to the asymptomatic population (Equation 3.4), individuals of infected symptomatic $(IS_v(t))$ get tested at the rate ϕ_{sv}. At the rate of λ, individuals recover naturally from $IS_v(t)$. Another way an individual can leave is either at the rate of κ to death and at the rate of h rates to becoming individuals hospitalized with the severe cases, respectively:

$$\frac{dIS_v}{dt} = PS_v(t-\eta) + STN_v(t) - \phi_{sv} \cdot IS_v(t) - \lambda \cdot IS_v(t) - \kappa_v \cdot IS_v(t) - h_v \cdot IS_v(t) \tag{3.14}$$

As before, τ, represents response time of the testing:

$$\frac{dST_v}{dt} = \phi_{sv} \cdot IS_v(t) - ST_v(t - \tau). \tag{3.15}$$

If an individual is testing positive, he/she is quarantined for λ_q days until recovery:

$$\frac{dQSP_v}{dt} = tp \cdot ST_v(t - \tau) - QSP_v(t - \lambda_q) \tag{3.16}$$

Recovery process is governed by Equation 3.12 listed earlier, replicated below for quick reference:

$$\frac{dKR_v}{dt} = QAP_v(t - \lambda_q) + FSQ_v(t - \lambda_q) + QSP_v(t - \lambda_q) + tn \cdot HT(t - \tau)$$

If a symptomatic individual falsely test negative, the individual continues to spread the variant v of the virus:

$$\frac{dSTN_v}{dt} = (1 - tp) \cdot ST_v(t - \tau) - STN_v(t) \tag{3.17}$$

Equation 3.5, shown earlier for asymptomatic individuals and listed below for quick reference, also captures the rate, λ, with which symptomatic individuals recover from variant v of the disease and are immunized in the process.

If an individual's situation worsens, they become hospitalized at the rate of h_v. They can be admitted before or during quarantine:

$$\frac{dHBQ_v}{dt} = h_v \cdot IS(_vt - \lambda_h) + (1 - tn) \cdot HBT_v(t - \tau) - \kappa_{hv} \cdot HBQ_v(t) \qquad (3.18)$$

$$\frac{dHDQ_v}{dt} = h_v \cdot QSP_v(t - \lambda_h) + (1 - tn) \cdot HDT_v(t - \tau) - \kappa_{hv} \cdot HDQ_v(t) \qquad (3.19)$$

Individuals can leave the hospital with either a negative test result or through death. After spending λ_h days in the hospital, the hospitalized individuals will take a test to check if the virus is still active and they will continue to be hospitalized when the test result is positive:

$$\frac{dHBT_v}{dt} = HBQ_v(t - \lambda_h) - HBT_v(t - \tau) \qquad (3.20)$$

$$\frac{dHDT_v}{dt} = HDQ_v(t - \lambda_h) - HDT_v(t - \tau) \qquad (3.21)$$

They can also move to the recovered compartment as described in Equation 3.12.

The death rate for hospitalized individuals with variant v is a higher rate of κ_{hv}, while the death rate for symptomatic individuals with variant v who are not hospitalized is, κ_v:

$$\frac{dD_v}{dt} = \kappa_{hv} \cdot \left(HBQ_v(t) + HDQ_v(t) \right) + \kappa_v \cdot QSP_v(t) + \kappa_v \cdot IS_v(t) \qquad (3.22)$$

During the epidemic, part of the non-infected population can show COVID-19 like symptoms due to other illnesses and they may need to be tested. The following

17

two equations leverage the flu and g_{sick} rate parameters to describe the portion of the population which show flu-like and general sick symptoms, such as fever and coughing, indistinguishable from COVID-19.

$$\frac{dFS}{dt} = flu \cdot S(t) - FS(t) \tag{3.23}$$

$$\frac{dGS}{dt} = g_{sick} \cdot S(t) - GS(t) \tag{3.24}$$

Since they show COVID-19 like symptoms, these individuals may be subject to symptomatic testing, analogous to Equation 3.15:

$$\frac{dFT}{dt} = \sum_{v=1}^{n} phi_{sv} \cdot FS(t) - FT(t - \tau) \tag{3.25}$$

$$\frac{dGT}{dt} = \phi_{sv} \cdot GS(t) - GT(t - \tau) \tag{3.26}$$

Even though these individuals are not COVID-19 infected, in the case of a false positive, they will be quarantined for λ_q days:

$$\frac{dFTQ}{dt} = (1 - tn) \cdot NT_v(t - \tau) - FTQ(t - \lambda_q) \tag{3.27}$$

$$\frac{dGTQ}{dt} = (1 - tn) \cdot GT(t - \tau) - GTQ(t - \lambda_q) \tag{3.28}$$

those non-infected patients whose test are true negative return to the susceptible population:

$$\frac{dGTN}{dt} = tn.GT(t - \tau) - GTN(t) \tag{3.29}$$

$$\frac{dFTN}{dt} = tn \cdot FT(t - \tau) - FTN(t) \tag{3.30}$$

Table 3.4: Sub-compartments Relevant for the Process Through Which Individuals Obtain Immunity. Source - [13]

Sub-Compartment		Sub-Compartment	
KR_v	Pop. of known recovered individuals from variant v	IM	Pop. of individuals with immunity
STI	Pop. of immune indiv. receiving serology (anti-body) test	SII	Pop. of not-immune indiv. receiving serology test
FPI	Pop. of individuals who are falsely presumed being immune		

Recovered and Immunization Process

The role of the immunity process is to show how recovered individuals from different variants of COVID-19 get immunized against reinfection for a certain period of time. This model is governed with sub-compartments presented in Table 3.4 along with the following equations:

Any person who recovers from the different variants of the disease $(KR_v(t))$, will have immunity for γ days. The recovered individuals can be separated into immunized individuals and individuals whose immunity fades before γ days. We can assume that the government will provide serology tests at a rate of ϕ_{se} for both groups, IM(t) and FPI(t), and since we can't tell which individual has lost their immunity, serology testing is needed. The serology test takes τ_{se} days to provide the result:

$$\frac{dIM}{dt} = KR_v(t) - IM(t - \gamma) - \phi_{se} \cdot IM(t) + tp_{se} \cdot STI(t - \tau_{se}) \qquad (3.31)$$

Some of the individuals (STI) who receive serology tests are immune, while others may be falsely presumed immune (FPI):

$$\frac{dSTI}{dt} = \phi_{se} \cdot IM(t) - STI(t - \tau_{se}) \qquad (3.32)$$

$$\frac{dSII}{dt} = \phi_{se} \cdot FPI(t) - FPI(t - \tau_{se}) \qquad (3.33)$$

Table 3.5: Sub-compartments Relevant for Individuals Who Are Falsely Presumed Susceptible. Source - [13]

	Sub-Compartment		Sub-Compartment
UR_v	Pop. with unknown immunity due to unknown infection	FPS	Pop. falsely presumed susceptible
FST_i	Pop. falsely presumed susceptible receiving diag. test i	$FSTP$	Pop. falsely presumed susceptible who test negative
FSQ	Pop. falsely presumed susceptible (wrongly) quarantined		

Individuals reach Falsely Presumed Immune (FPI(t)) for various reasons: (a) the individual may lose immunity (IM(t-γ)), (b,c) some individuals may be quarantined after false positive result due to flu symptoms (FSQ(t-λ_q),or after a general sickness (GSQ(t-λ_q)),and (d) after a testing error ((NTQ(t-λ_q)):

$$\frac{dFPI}{dt} = IM(t-\gamma) + GSQ(t-\lambda_q) + FSQ(t-\lambda_q) + NTQ(t-\lambda_q) - \beta'_v \cdot FPI(t) \quad (3.34)$$

For these individuals, we can apply a higher infection rate ($\beta'_v > \beta_v$) in Equation 3.1, because these individuals are likely to socialize more than pure susceptible ones.

Finally, like the diagnostic tests, the serology tests are also imperfect. We use accuracy of tp_{se} and tn_{se} for sensitivity and specificity, respectively.

Falsely Presumed Susceptible Process

The falsely presumed susceptible process is made to study the individuals that recovered naturally from different variants of COVID-19. These individuals have immunity after recovering but they falsely presume themselves as susceptible. The sub-compartments in the falsely presumed susceptible process are listed in Table 3.5.

As formalized in Equation 3.5, unknown recovered individuals, who had variant v of COVID-19 but recovered naturally, have immunity against the different variants of COVID-19. These individuals move to the falsely presumed susceptible (FPS) since they regard themselves susceptible. Other individuals who are falsely presumed

susceptible include individuals who are already falsely presumed negative and test negative for a diagnostic test and others who are immune, but receive a false negative for a serology test:

$$\frac{dFPS}{dt} = \sum_{v=1}^{n} UR_v(t) + FSTN(t) + (1 - tp_{se}) \cdot STI(t - \tau_{se}) - FPS(t - \gamma) - \phi_{av} \cdot FPS(t)$$
(3.35)

Individuals in FPS(t) will lose immunity after γ days and move back to the main susceptible compartment. The above equation also captures that, while they are falsely presumed susceptible, they may be subject to regular random testing, with rate, ϕ_{av}. These tested individuals move to the corresponding FST compartment:

$$\frac{dFST}{dt} = \phi_{av} \cdot FPS(t) - FST(t - \tau)$$
(3.36)

Some of these falsely presumed susceptible individuals will test negative:

$$\frac{FSTN}{dt} = tn \cdot FST(t - \tau) - FSTN(t)$$
(3.37)

Others, however, may test positive, and get quarantined for λ_q days, due to testing errors:

$$\frac{dFSQ}{dt} = (1 - tn) \cdot FST(t - \tau) - FSQ(t - \lambda_q)$$
(3.38)

3.3 Model Validation

Parameter Calibration

For a given geographic location (this could be county, state, city), we can calibrate the SIRTEM model by estimating the infection rate (β_v) and testing rate (ϕ_{sv}) for each of the variants using published data of confirmed positive and negative cases (Table 3.6).

We obtain estimation of these parameters by solving an equation. Given the complexity of the simulator, we cannot optimize the likelihood in closed form. Therefore,

an algorithm where the discrepancy between the daily positive and negative cases predicted against the real data is improved after every iteration.

$$(P): \min_{\omega \in \Omega} Z(\omega) = \frac{1}{2} \left[\frac{\frac{1}{d}\sum_{t=1}^{d}(\hat{y}_{tv}^{(+)}|\omega) - y_{tv}^{(+)})^2)}{y_{tv}^{(+)}} + \frac{\frac{1}{d}\sum_{t=1}^{d}(\hat{y}_{tv}^{(-)}|\omega) - y_{tv}^{(-)})^2}{y_{tv}^{(-)}} \right] \quad (3.39)$$

$$MAPE = \frac{1}{d}\left| \sum_{t=1}^{d} \frac{(Actual_t - Forecast_t)}{Actual_t} \right| \quad (3.40)$$

In equation (3.39), $\hat{y}_{tv}^{(+)}$, $\hat{y}_{tv}^{(-)}$ are the SIRTEM predictions for positive and negative cases at time t (day) for the variant v, respectively. $y_{tv}^{(+)}$, $y_{tv}^{(-)}$ denotes the confirmed positive and negative cases for variant v which is obtained from public sources [2; 6]. Each error term is then normalized by using the average confirmed cases for each variant v, $y_{tv}^{(+)}$ and $y_{tv}^{(-)}$, respectively.MAPE in (3.40) is Mean absolute percentage error.It is used to find MAPE for positive cases, negative cases, hospitalization and deaths.

Observing that the infection rate β_v and the testing rate ϕ_v of the v-th variant are time dependent random process, we assume both to be auto-regressive processes. In particular, following the approach in [13], we use an order 2 autoregressive model (AR(2)) model. This is because the parameter β_v is the product of transmissibility of the disease and the likelihood of interactions among individuals within the population.These two probabilities are non memoryless. Similarly, in practice, the diagnostic testing rate cannot be assumed to be memoryless. Since auto-regressive processes are used to model situations where an observation at a given time depends on the past, they are suitable to model the evaluation of the β_v and ϕ_v parameters over time. As a result, for each week k, the equations of β_v and ϕ_v for a variant are:

$$\beta_{vk} = a_{0v} + a_{1v} \cdot \beta_{vk-1} + a_{2v} \cdot \beta_{vk-2} + e_{vk}^b \quad (3.41)$$

$$\phi_{vk} = b_{0v} + b_{1v} \cdot \phi_{vk-1} + b_{2v} \cdot \phi_{vk-2} + e_{vk}^s \quad (3.42)$$

Here e_{vk}^b and e_{vk}^s are not learnable parameters, but noise terms. Under these assumptions, the decision vector is $\omega = [\mathbf{av}, \mathbf{bv}]$. $\mathbf{av}$ and $\mathbf{bv}$ are 3-dimensional vectors which needs to be estimated. This results in a 6-dimensional decision problem. We use K weeks of data to calibrate. Specifically, we designed a procedure for the finding the optimum auto-regressive parameters in (3.41)-(3.42). The approach is summarized below:

Algorithm 1: Calibration Algorithm for $\mathbf{a_v}, \mathbf{b_v}$.

Input: : Initialize a set of uniform random numbers for $\mathbf{A_v}(\forall_v), \mathbf{B_v}(\forall_v)$ and

$\tilde{Z}_v^{\,*} = \infty(\forall_v)$

for $\mathbf{a_v} \subset \mathbf{A_v}$ **do**

    for $\mathbf{b_v} \subset \mathbf{B_v}$ **do**

        Step1 : Use AR(2) eqn. (3.41)-(3.42) to determine the MVSIRTEM

        parameters $\boldsymbol{\beta_v}, \boldsymbol{\phi_v}$ from $\mathbf{a_v}, \mathbf{b_v}$;

        Step2 : Run MVSIRTEM for K weeks and obtain the daily estimates

        $\left(\hat{y_{tv}}^{(+)}, \hat{y_{tv}}^{(-)}\right)_{t=1}^{K \times 7}$, use training data $\left(y_{tv}^{(+)}, y_{tv}^{(-)}\right)_{t=1}^{K \times 7}$ and calculate

        $Z(\omega)$ using eqn. (3.39) for each variant v;

        Step 3: Update the incumbent;

        if $Z_v(\omega) < \tilde{Z}_v^{\,*}$ **then**

            $\tilde{\omega}^* \leftarrow \omega$;

            $\tilde{Z}_v^{\,*} \leftarrow Z_v(\omega)$;

        end

    end

end

Step 4: Estimate an AR(2) model using $\tilde{\omega}^*$ to predict the next β_v and ϕ_v for

1 weeks and run MVSIRTEM for $k + 1$ weeks and obtain the daily estimates

and calculate MAPE for positive cases, hospitalization and deaths using

eqn. (3.40)

Table 3.6: Default Values for Various MVSIRTEM Model Parameters. Source - [13]

Parameter	Description	Values
tp	Sensitivity of diagnostic test	0.75
tn	Specificity of the diagnostic test	0.95
tp_{se}	Sensitivity of the serology test	0.84
tn_{se}	Specificity of the serology test	0.97
τ	Time to result for diagnostic test	3days
τ_{se}	Time to result for the serology test	5days
ϕ_v	Diagnostic testing rate for symptomatic individuals for variant v	Estimated
ϕ_a	Diagnostic testing rate for non-symptomatic individuals	0
ϕ_{se}	Serology Test rate	0.01
β_v	Infection rate for the susceptible pop. for variant v	Estimated
β'_v	Inf. rate for falsely presumed immune pop. (ratio) for variant v	$1.2 \cdot \beta_v$
r	Ratio of transmission rates for asympt. population against sympt. population	0.51
per_a	Percentage of ind. with COVID-19 who are asymptomatic	0.16
per_s	Percentage of ind. with COVID-19 who are symptomatic	0.84
η	Incubation length (days)	3.2 days
λ_a	Length of recovery for asymptomatic ind. (days)	3.5 days
λ_s	Length of recovery for symptomatic ind. (days)	7 days
λ_q	Length of quarantine (days)	14 days
h_v	hospitalization rate (ratio of quarantined pop, per day) for variant v	Estimated
λ_h	Hospitalization length (days)	6 days
κ_v	Mortality rate for symptomatic pop. (per day) for variant v	Estimated
κ_{hv}	Mortality rate for hospitalized individuals (per day) for variant v	Estimated
g	Ratio of susc. who have fever for non COVID infections (ratio, per day)	0.04

Chapter 4

EXPERIMENTS AND RESULTS

4.1   Calibration of $\mathbf{a_v}, \mathbf{b_v}$ for Omicron and Delta

The experiment that has been carried out is for two variants Delta and Omicron. The simulation model was developed in MATLAB where the differential equations were formulated and solved using the `dde23.m` library [4]. The experiment is calibrated using the state of Arizona's published positive and negative cases [2]. Since there is no accurate published data for positive and negative cases for Arizona, we use the Nowcast [6]. Nowcast is a data set that shows the proportions of the different variants for every week.

The simulation was run from September 6th 2021 to January 23rd 2022 for the calibration part. For simplification, the sub variants of Omicron i.e., BA1.1 and BA.2 were considered one variant. The Nowcast ratios for Omicron and Delta during that time are shown in Figure 4.1. Using the Nowcast data, we can find the ratio of Positive and Negative cases for each variant from the published data.

For finding the auto-regressive parameters for both the variants, we use the range in Table 4.1. The values are obtained by using the rand function in MATLAB. 500 values were found for each of the parameters. The next section shows how we calibrate the hospitalization and mortality rates for both variants.
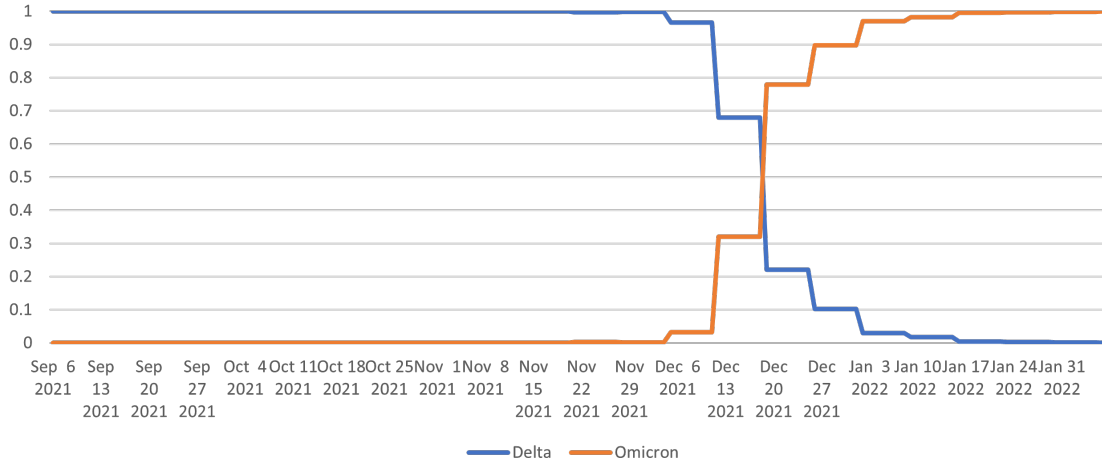
Figure 4.1: Ratio Between Omicron and Delta Using the Nowcast Data

Table 4.1: Range of Auto-regressive Parameters for Variant $v \in [Delta, Omicron]$

AR Parameter	Range
a_{0v}	(0,2)
a_{1v}	(-2,2)
a_{2v}	(-2,2)
b_{0v}	(0,2)
b_{1v}	(-2,2)
b_{2v}	(-20,20)

4.2 Other Parameters

Calibration of parameters like mortality rate and hospitalization rate needs to be done for each of the variants. The deaths and hospitalizations for each of the variant is found. The mortality rate and hospitalization rate for week i is formulated using

the below equations:

$$\kappa_{vi} = \frac{1}{9.5} \cdot \frac{d_{vi}}{y_{vi}^{(+)}} \tag{4.1}$$

$$\kappa_{hvi} = \frac{8}{9.5} \cdot \frac{d_{vi}}{y_{vi}^{(+)}} \tag{4.2}$$

$$h_{vi} = \frac{H_{vi}}{y_i^{(+)}} \tag{4.3}$$

In the above equations, $y_{vi}^{(+)}$ is the positive cases of variant v during week i. In equation (4.1) and (4.2), d_{vi} is the amount out deaths that occurred due to variant v during week i. The ratio $\frac{1}{9.5}$ and $\frac{8}{9.5}$ are assumed as hospitalised patients have a higher death rate [14; 3]. In equation (4.3), H_{vi} is the amount of hospitalised patients from variant v during week i.

4.3 Results

We first ran the simulation for 10 weeks (September 6th to November 14th 2021) where only the Delta variant was prevalent. Then when Omicron started on November 15th 2021, we ran the training simulation for n weeks where $n \in [4, 5, 6, 7, 8, 9]$ and then a forecast is done for $n + 1^{th}$ week. The parameters change over each iteration. In the following subsections, we show the results for positive cases, hospitalisation and number of deaths. As we are focusing on the multi variant results, the results for Delta variant between September 6th 2021 to November 14th 2021 is omitted.

Positive cases

Figure 4.2a and Figure 4.2c shows the trained model results during November 15th to December 12th 2021 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is very high. This due to the fact that the training is done for a short time horizon of four weeks. Figure 4.2b and Figure 4.2d shows the forecast model results during December 13th - December 19th 2021 for positive cases due to the Delta and Omicron variants respectively. Due to the high MAPE in the training data, the forecast MAPE is also high as seen in Tables 4.3 and 4.5.

Figure 4.3a and Figure 4.3c shows the trained model results during November 15th to December 19th 2021 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is decreased by 0.4% and 0.2% for Delta and Omicron respectively from the previous training time horizon. This negligible increase in training accuracy is because the time horizon for training increased to five weeks.

Figure 4.3b and Figure 4.3d shows the forecast model results during December 20th - December 26th 2021 for positive cases due to the Delta and Omicron variants respectively. The Forecast MAPE has decreased by 0.5% and 7.9% for Delta and Omicron respectively from the previous time horizon as seen in Tables 4.3 and 4.5. The longer time horizon used for training has increased forecasting accuracy of the model.

Figure 4.4a and Figure 4.4c shows the trained model results during November 15th to December 26th 2021 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is decreased by 0.7% and 0.8% for Delta and Omicron respectively from

the previous training time horizon. This negligible increase in training accuracy is because the time horizon for training increased to six weeks.

Figure 4.4b and Figure 4.4d shows the forecast model results during December 27th 2021 - January 2nd 2022 for positive cases due to the Delta and Omicron variants respectively. The Forecast MAPE has increased by 0.7% and 1.5% for Delta and Omicron respectively from the previous time horizon as seen in Tables 4.3 and 4.5. This increase in MAPE can be because, the positive case due to Delta variant started decreasing very rapidly and was replaced by positive cases due to Omicron variant.

Figure 4.5a and Figure 4.5c shows the trained model results during November 15th 2021 to January 2nd 2022 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is decreased by 0.1% and 1.1% for Delta and Omicron respectively from the previous training time horizon. This increase in training accuracy is because the time horizon for training increased to seven weeks.

Figure 4.5b and Figure 4.5d shows the forecast model results during January 3rd - January 9th 2022 for positive cases due to the Delta and Omicron variants respectively. The Forecast MAPE has decreased by 2.1% and 5.4% for Delta and Omicron respectively from the previous time horizon as seen in Tables 4.3 and 4.5. The longer time horizon used for training has increased forecasting accuracy of the model.

Figure 4.6a and Figure 4.6c shows the trained model results during November 15th 2021 to January 9th 2022 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is increased by 0.2% and decreased by 5.4% for Delta and Omicron respectively from the previous training time horizon. The increase in training accuracy for Omicron is because the time horizon for training increased to eight weeks and

30

the negligible decrease in training accuracy for Delta is because the amount exposed individuals for delta decreased rapidly.

Figure 4.6b and Figure 4.6d shows the forecast model results during January 10th - January 16th 2022 for positive cases due to the Delta and Omicron variants respectively. The Forecast MAPE has increased by 0.7% and decreased by 10% for Delta and Omicron respectively from the previous time horizon as seen in Tables 4.3 and 4.5. The forecast results is emulating the training model results for both variants.

Figure 4.7a and Figure 4.7c shows the trained model results during November 15th 2021 to January 16th 2022 for positive cases due to the Delta and Omicron variants respectively. As we can see in Tables 4.2 and 4.4, the trained MAPE during that time period is decreased by 1.5% and increased by 0.3% for Delta and Omicron respectively from the previous training time horizon. The decrease in training accuracy for Omicron is because the amount of positive cases started decreasing rapidly after reaching a peak. The negligible increase in training accuracy for Delta is because the time horizon for training increased to nine weeks.

Figure 4.7b and Figure 4.7d shows the forecast model results during January 17th - January 23rd 2021 for positive cases due to the Delta and Omicron variants respectively. The Forecast MAPE has decreased by 3.1% and 1.4% for Delta and Omicron respectively from the previous time horizon as seen in Tables 4.3 and 4.5. The longer time horizon used for training has increased forecasting accuracy of the model.
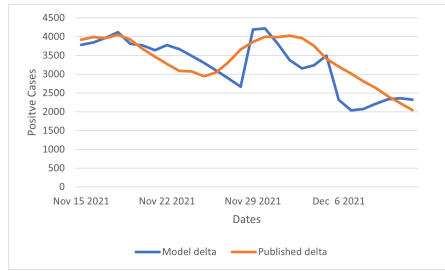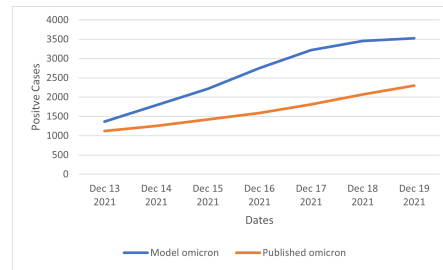
(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Dec 12th 2021
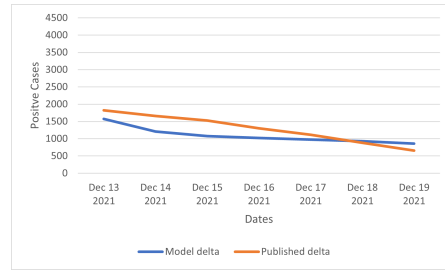
(b) Forecasted Positive Cases Due to Delta Variant Between Dec 13th 2021 - Dec 19th 2021
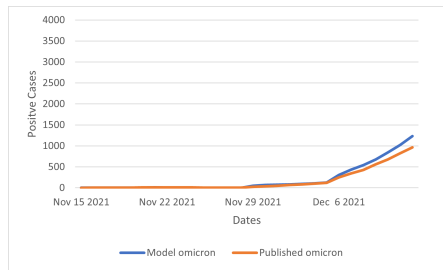
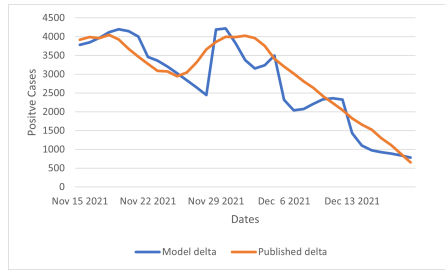(c) Trained Positive Cases Due to Omicron Variant Between Nov 15th 2021 - Dec 12th 2021

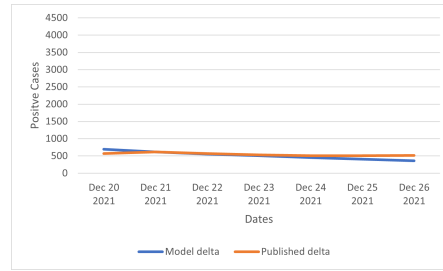(d) Forecasted Positive Cases Due to Omicron Variant Between Dec 13th 2021 - Dec 19th 2021

Figure 4.2: Positive Cases Between Nov 15th 2021 - Dec 19th 2021 (Model Vs Published)
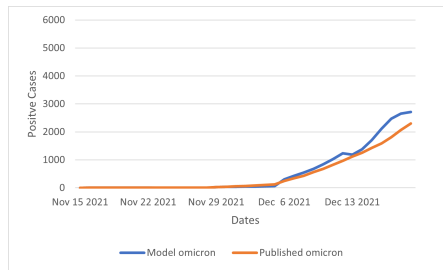
Hospitalization

Tables 4.6 and 4.8 shows the MAPE for the trained hospitalizations due to Delta and Omicron variants respectively. From this we can see that, as the training time horizon gets longer, the MAPE for the training data gradually decreases, just like we saw in the earlier subsection of positive cases. However after 4 iterations (in making the training horizon longer) in both the delta and omicron variants, it reaches a stationary state where the difference in error is almost negligible . This is because the AR(2) model is trained using the positive and negative cases and hospitalizations
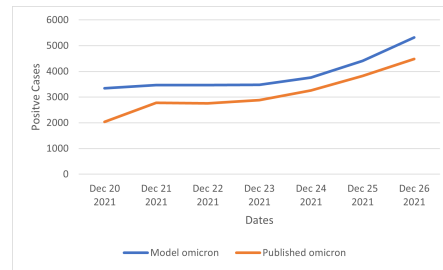
(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Dec 19th 2021

(b) Forecasted Positive Cases Due to Delta Variant Between Dec 20th 2021 - Dec 26th 2021

(c) Trained Positive Cases Due to Omicron Variant Between Nov 15th 2021 - Dec 19th 2021
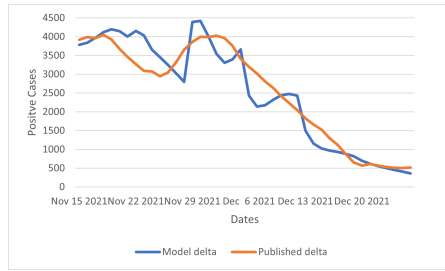
(d) Forecasted Positive Cases Due to Omicron Variant Between Dec 20th 2021 - Dec 26th 2021
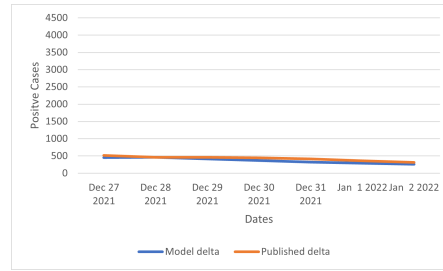
Figure 4.3: Positive Cases Between Nov 15th 2021 - Dec 26th 2021 (Model Vs Published)

rate is derived from only the positive cases.

Tables 4.7 and 4.9 shows the MAPE for the forecasted hospitalizations due to Delta and Omicron variants respectively. The MAPE for the forecasted hospitalizations has a downward trend as the training time horizon gets longer but like the training MAPE, it also reaches a stationary state where the difference in error is almost negligible also.
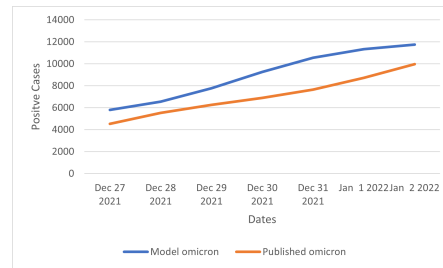
(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Dec 26th 2021

(b) Forecasted Positive Cases Due to Delta Variant Between Dec 27th 2021 - Jan 2nd 2022
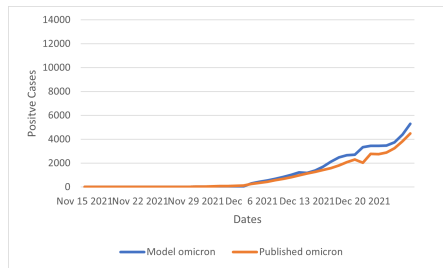
(c) Trained Positive Cases Due to Omicron Variant Between Nov 15th 2021 - Dec 26th 2021
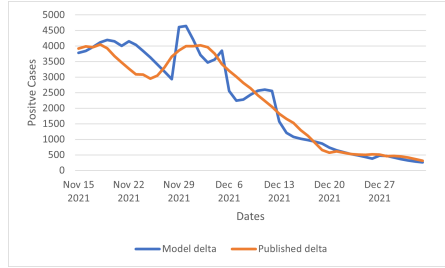
(d) Forecasted Positive Cases Due to Omicron Variant Between Dec 27th 2021 - Jan 2nd 2022
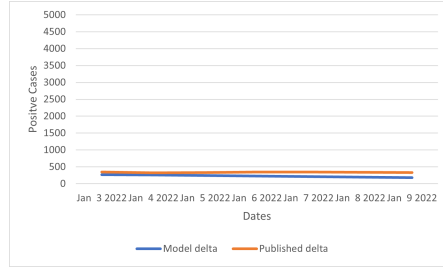
Figure 4.4: Positive Cases Between Nov 15th 2021 - Jan 2nd 2022 (Model Vs Published)
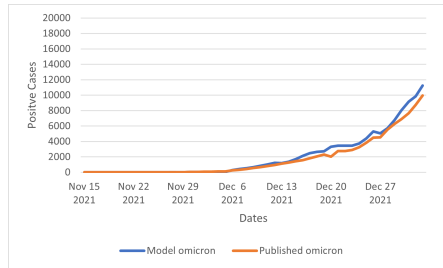
Deaths

Tables 4.6 and 4.8 shows the MAPE for the trained deaths due to Delta and Omicron variants respectively. From this we can see that, as the training time horizon gets longer, the MAPE for the training data gradually decreases, just like we saw in the earlier subsection of positive cases. However after four iterations and three iterations (in making the training horizon longer) in the delta and omicron variant respectively, it reaches a stationary state where the difference in error is almost negligible . This is because just like the hospitalization rates, the AR(2) model is trained using the
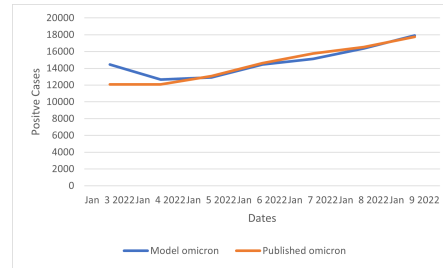
(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Jan 2nd 2022

(b) Forecasted Positive Cases Due to Delta Variant Between Jan 3rd 2022 - Jan 9th 2022

(c) Trained Positive Cases due to Omicron variant between Nov 15th 2021 - Jan 2nd 2022
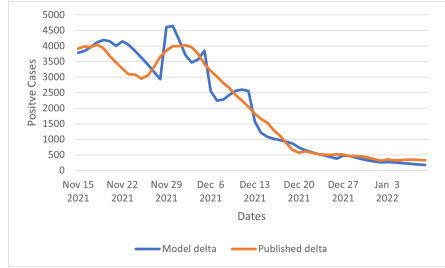

(d) Trained Positive Cases Due to Omicron Variant Between Jan 3rd 2022 - Jan 9th 2022

Figure 4.5: Positive Cases Between Nov 15th 2021 - Jan 9th 2022 (Model Vs Published)

positive and negative cases and death rates are derived from only the positive cases.

Tables 4.7 and 4.9 shows the MAPE for the forecasted deaths due to Delta and Omicron variants respectively. The MAPE for the forecasted deaths has a downward trend as the training time horizon gets longer but like the training MAPE, it also reaches a stationary state where the difference in error is almost negligible also.

In the next section we summarize the results that we obtained in the previous section.

(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Jan 9th 2022

(b) Forecasted Positive Cases Due to Delta Variant Between Jan 10th 2022 - Jan 16th 2022

(c) Trained Positive Cases Due to Omicron Variant Between Nov 15th 2021 - Jan 9th 2022

(d) Forecasted Positive Cases Due to Omicron Variant Between Jan 10th 2022 - Jan 16th 2022

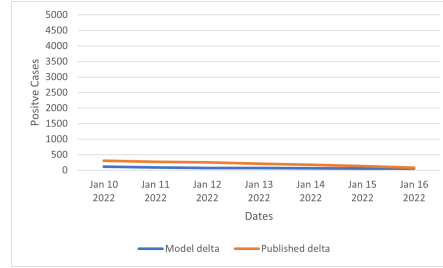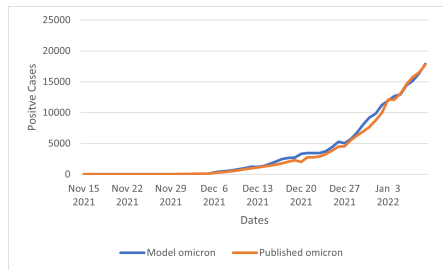Figure 4.6: Positive Cases Between Nov 15th 2021 - Jan 16th 2022 (Model Vs Published)

4.4   Summary

The summary of the results are mentioned below:

1. For positive cases the MAPE has a downward trend for both the training results and the forecast results. However, there are times where the MAPE increased because of rapidly decreasing positive cases.

2. For hospitalizations the MAPE has a downward trend for both the training results and the forecast results. However, after a few iterations where the training
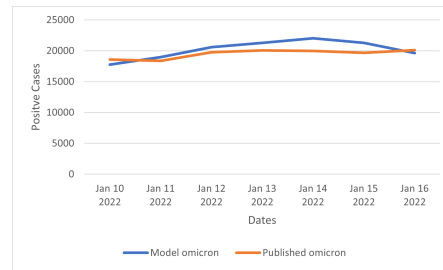
36

(a) Trained Positive Cases Due to Delta Variant Between Nov 15th 2021 - Jan 16th 2022

(b) Forecasted Positive Cases Due to Delta Variant Between Jan 17th 2022 - Jan 23rd 2022

(c) Trained Positive Cases Due to Omicron Variant Between Nov 15th 2021 - Jan 16th 2022
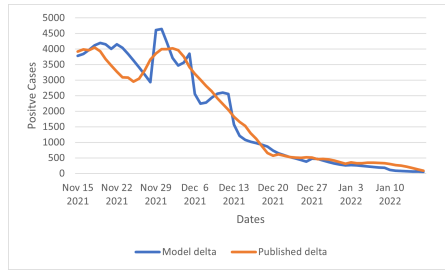
(d) Forecasted Positive Cases Due to Omicron Variant Between Jan 17th 2022 - Jan 23rd 2022

Figure 4.7: Positive Cases Between Nov 15th 2021 - Jan 23rd 2022 (Model Vs Published)
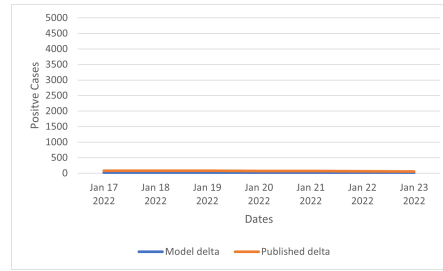
time horizon gets longer after each iteration, the difference in error is almost negligible.
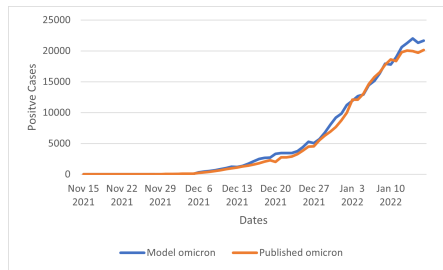
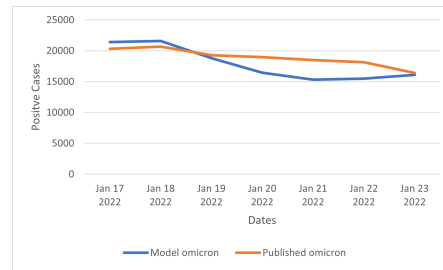3. For deaths the MAPE has a downward trend for both the training results and the forecast results. However, after a few iterations where the training time horizon gets longer after each iteration, the difference in error is almost negligible.

Now we utilize all the discussed results in this chapter to draw conclusions which is discussed in the next chapter.

Table 4.2: MAPE for the Trained Positive Cases of Delta Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.13712
Nov 15th 2021 - Dec 19th 2021	0.13391
Nov 15th 2021 - Dec 26th 2021	0.12685
Nov 15th 2021 - Jan 2nd 2022	0.12499
Nov 15th 2021 - Jan 9th 2022	0.12715
Nov 15th 2021 - Jan 16th 2022	0.11262

Table 4.3: MAPE for Forecasted Positive Cases of Delta Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.20184
Dec 20th 2021 - Dec 26th 2021	0.19597
Dec 27th 2021 - Jan 2nd 2022	0.20750
Jan 3rd 2022 - Jan 9th 2022	0.18631
Jan 10th 2022 - Jan 16th 2022	0.19341
Jan 17th 2022 - Jan 23rd 2022	0.16429

Table 4.4: MAPE for the Trained Positive Cases of Omicron Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.21609
Nov 15th 2021 - Dec 19th 2021	0.21488
Nov 15th 2021 - Dec 26th 2021	0.20689
Nov 15th 2021 - Jan 2nd 2022	0.18963
Nov 15th 2021 - Jan 9th 2022	0.1786
Nov 15th 2021 - Jan 16th 2022	0.18147

Table 4.5: MAPE for Forecasted Positive Cases of Omicron Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.42007
Dec 20th 2021 - Dec 26th 2021	0.34123
Dec 27th 2021 - Jan 2nd 2022	0.35609
Jan 3rd 2022 - Jan 9th 2022	0.30261
Jan 10th 2022 - Jan 16th 2022	0.20091
Jan 17th 2022 - Jan 23rd 2022	0.18647

Table 4.6: MAPE for the Trained Hospitalization of Delta Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.27383
Nov 15th 2021 - Dec 19th 2021	0.24936
Nov 15th 2021 - Dec 26th 2021	0.22752
Nov 15th 2021 - Jan 2nd 2022	0.22029
Nov 15th 2021 - Jan 9th 2022	0.22436
Nov 15th 2021 - Jan 16th 2022	0.22388

Table 4.7: MAPE for Forecasted Hospitalization of Delta Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.30295
Dec 20th 2021 - Dec 26th 2021	0.28273
Dec 27th 2021 - Jan 2nd 2022	0.28031
Jan 3rd 2022 - Jan 9th 2022	0.27646
Jan 10th 2022 - Jan 16th 2022	0.27753
Jan 17th 2022 - Jan 23rd 2022	0.27336

Table 4.8: MAPE for the Trained Hospitalization of Omicron Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.34863
Nov 15th 2021 - Dec 19th 2021	0.31519
Nov 15th 2021 - Dec 26th 2021	0.29324
Nov 15th 2021 - Jan 2nd 2022	0.26851
Nov 15th 2021 - Jan 9th 2022	0.25043
Nov 15th 2021 - Jan 16th 2022	0.25553

Table 4.9: MAPE for Forecasted Hospitalization of Omicron Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.36286
Dec 20th 2021 - Dec 26th 2021	0.34273
Dec 27th 2021 - Jan 2nd 2022	0.35315
Jan 3rd 2022 - Jan 9th 2022	0.33816
Jan 10th 2022 - Jan 16th 2022	0.32227
Jan 17th 2022 - Jan 23rd 2022	0.32571

Table 4.10: MAPE for the Trained Deaths of Delta Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.30738
Nov 15th 2021 - Dec 19th 2021	0.28154
Nov 15th 2021 - Dec 26th 2021	0.25647
Nov 15th 2021 - Jan 2nd 2022	0.23921
Nov 15th 2021 - Jan 9th 2022	0.23817
Nov 15th 2021 - Jan 16th 2022	0.23633

Table 4.11: MAPE for Forecasted Deaths of Delta Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.35636
Dec 20th 2021 - Dec 26th 2021	0.31236
Dec 27th 2021 - Jan 2nd 2022	0.32629
Jan 3rd 2022 - Jan 9th 2022	0.32034
Jan 10th 2022 - Jan 16th 2022	0.32974
Jan 17th 2022 - Jan 23rd 2022	0.32821

Table 4.12: MAPE for the Trained Deaths of Omicron Variant

Training Period	MAPE for Training
Nov 15th 2021 - Dec 12th 2021	0.38683
Nov 15th 2021 - Dec 19th 2021	0.35432
Nov 15th 2021 - Dec 26th 2021	0.32187
Nov 15th 2021 - Jan 2nd 2022	0.29934
Nov 15th 2021 - Jan 9th 2022	0.29979
Nov 15th 2021 - Jan 16th 2022	0.29775

Table 4.13: MAPE for Forecasted Deaths of Omicron Variant

Forecast Period	MAPE for Forecast
Dec 13th 2021 - Dec 19th 2021	0.44863
Dec 20th 2021 - Dec 26th 2021	0.38399
Dec 27th 2021 - Jan 2nd 2022	0.3251
Jan 3rd 2022 - Jan 9th 2022	0.3298
Jan 10th 2022 - Jan 16th 2022	0.33204
Jan 17th 2022 - Jan 23rd 2022	0.33311

Chapter 5

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

Epidemic models are designed to show how likely outcome of a pandemic and inform the public about the severity and spread of the disease. MVSIRTEM takes this to another level and shows how each disease variant can increase in the population. MVSIRTEM shows how Omicron became the primary variant of COVID-19 and replaced the Delta variant. The objectives of this thesis were to make changes to the underlying equations to add multiple variants and calibrate different parameters.

We summarise the mentioned points below:

1. The grid search used in finding the auto-regressive parameters of the AR(2) model can be used to accurately find the parameters, however, is computationally expensive and time-consuming. Other types of searches like jump search, interpolation search, etc, may give similar results but with less usage of computation power and time but further experimentation is required.

2. MVSIRTEM can be used to model multiple variants. The accuracy of the model increases as the training set used becomes larger.

3. MVSIRTEM has a stationary MAPE for Hospitalizations and Deaths after a few iterations of making the training set larger. This may be rectified by training the hospitalization and death rates of the model but this may also cause overfitting. Additional experimentation is required to ascertain this.

5.2 Future Work

There are immense opportunities in the presented research work for advancements. Below are a few of the mentions:

1. **Adding multiple populations:** MVSIRTEM has only one susceptible population. But in realistic terms, there can be multiple susceptible populations and can vary from a small town to different countries. The model can be extended to work with multiple populations.

2. **Adding vaccination protocols:** MVSIRTEM doesn't account for the vaccinations and how it controls the spread of the disease and reduces the hospitalization rates and death rates. The model can be extended to add vaccination protocols.

3. **Adding multiple age groups:** At present, MVSIRTEM doesn't take into account different age groups and how they have they interact with the other age groups. The model can be extended to have different age groups.

REFERENCES

[1] "Compartmental models in epidemiology", URL `https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology`.

[2] "Covid-19 data tracker", URL `https://covid.cdc.gov/covid-data-tracker/#datatracker-home`.

[3] "Covid-19 positive negative cases", URL `https://covidtracking.com/`.

[4] "Dde 23 library", URL `https://www.mathworks.com/help/matlab/ref/dde23.html`.

[5] "Lecture notes in mathematical epidemiology", URL `https://ms.mcmaster.ca/earn.old/pdfs/Earn2008_LightIntro.pdf`.

[6] "Variant proportions", URL `https://covid.cdc.gov/covid-data-tracker/#variant-proportions`.

[7] Abbey, H., "An examination of the reed–frost theory of epidemics", Hum. Biol **24**, 201–233.

[8] Aguilar, J. B. and J. B. Gutierrez, "An epidemiological model of malaria accounting for asymptomatic carriers", Bulletin of Mathematical Biology **82**, 3, 1–55.

[9] Allen, L. J. S., "An introduction to stochastic epidemic models", Lecture Notes in Mathematics **1945**, 81–130.

[10] Anderson, R. M., B. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control* (Oxford university press, na, 1992).

[11] Annas, S., M. I. Pratama, M. Rifandi, W. Sanusi and S. Side, "Stability analysis and numerical simulation of seir model for pandemic covid-19 spread in indonesia", Chaos, Solitons & Fractals **139**, 110072.

[12] Arcede, J. P., R. L. Caga-Anan, C. Q. Mentuda and Y. Mammeri, "Accounting for symptomatic and asymptomatic in a seir-type model of covid-19", Mathematical Modelling of Natural Phenomena **15**, 34.

[13] Azad, F. T., R. W. Dodge, A. M. Varghese, J. Lee, G. Pedrielli, K. S. Candan and G. Chowell, "Sirtem: Spatially informed rapid testing for epidemic modeling and response to covid-19", Association for Computing Machinery (Under Review) .

[14] Buckner, J. H., G. Chowell and M. R. Springborn, "Dynamic prioritization of covid-19 vaccines when social distancing is limited for essential workers", Proceedings of the National Academy of Sciences **118**, 16, na.

[15] Carcione, J. M., J. E. Santos, C. Bagaini and J. Ba, "A simulation of a covid-19 epidemic based on a deterministic seir model", Frontiers in public health **8**, 230.

[16] Daley, D. J. and J. Gani, "Epidemic modelling: An introduction", Cambridge Studies in Mathematical Biology **15**.

[17] Davey, R., "What is epidemiologic modeling?", URL `https://www.news-medical.net/health/What-is-Epidemiologic-Modeling.aspx#`.

[18] Engbert, R., M. M. Rabe, R. Kliegl and S. Reich, "Sequential data assimilation of the stochastic seir epidemic model for regional covid-19 dynamics", Bulletin of mathematical biology **83**, 1, 1–16.

[19] Fan, R. G., Y. B. Wang, M. Luo, Y. Q. Zhang and C. P. Zhu, "Seir-based covid-19 transmission model and inflection point prediction analysis", Dianzi Keji Daxue Xuebao/Journal of the University of Electronic Science and Technology of China **49**, 3.

[20] Greenwood, M., "On the statistical measure of infectiousness", J. Hyg Cambridge **31**, 336–351.

[21] Hamzah, F. B., C. Lau, H. Nazri, D. Ligot, G. Lee, C. Tan, M. Shaib, U. Zaidon, A. Abdullah, M. Chung *et al.*, "Coronatracker: worldwide covid-19 outbreak data analysis and prediction", Bull World Health Organ **1**, 32, na.

[22] Kermack, W. O. and A. G. McKendrick, "A contribution to the mathematical theory of epidemics", I. Proc. R. Soc. Lond. A **115**, 700–721.

[23] Lekone, P. E. and B. F. Finkenstädt, "Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study", Biometrics **62**, 4, 1170–1177.

[24] Li, M. Y. and J. S. Muldowney, "Global stability for the seir model in epidemiology", Mathematical biosciences **125**, 2, 155–164.

[25] Liu, Q., D. Jiang, T. Hayat and A. Alsaedi, "Dynamical behavior of a stochastic epidemic model for cholera", Journal of the Franklin Institute **356**, 13, 7486–7514.

[26] López, L. and X. Rodo, "A modified seir model to predict the covid-19 outbreak in spain and italy: simulating control scenarios and multi-scale epidemics", Results in Physics **21**, 103746.

[27] Mwalili, S., M. Kimathi, V. Ojiambo, D. Gathungu and R. Mbogo, "Seir model for covid-19 dynamics incorporating the environment and social distancing", BMC Research Notes **13**, 1, 1–5.

[28] Pandey, G., P. Chaudhary, R. Gupta and S. Pal, "Seir and regression model based covid-19 outbreak predictions in india", arXiv preprint arXiv:2004.00958 .

[29] Watson, H. and F. Galton, "On the probability of the extinction of families", The Journal of the Anthropological Institute of Great Britain and Ireland **4**, 138–144.

[30] Yang, Z., Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai *et al.*, "Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions", Journal of thoracic disease **12**, 3, 165.