Pandemic Prediction and Control with Integrated Dynamic Modeling of Disease Transmission and Healthcare Resource Optimization

by

Mingdong Lyu

A Dissertation Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the Requirements for the Degree
DOCTOR OF PHILOSOPHY
(INDUSTRIAL AND SYSTEMS ENGINEERING)

May 2023

Copyright 2023 Mingdong Lyu

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Significance and Importance of Pandemic Research

Pandemics are large-scale outbreaks of infectious diseases that can significantly increase morbidity and mortality over a wide geographic area, leading to substantial economic, social, and political disruption. When a new virus emerges, people typically have little or no immunity, allowing diseases like influenza pandemics to impact a considerable proportion of the global population and place immense strain on healthcare systems. Regardless of their severity, pandemics affect large swaths of the population, necessitating a multisectoral response that can last several months or even years. Besides influenza, novel viral diseases such as COVID-19 have the potential to impact all segments of the population, potentially resulting in millions of deaths, causing a global economic recession, and exacerbating political stress. These pandemics are particularly harmful to vulnerable groups, including people living in poverty, older individuals, and persons with disabilities.

As global travel and integration have increased, so has the risk of small local epidemics transforming into global pandemics. In the 21st century, prior to the outbreak of COVID-19, the world experienced five major epidemics:1) In 2002, Severe Acute Respiratory Syndrome (SARS) emerged in Guangdong, China, killed 774 people. 2) In 2003, Avian flu broke out in Southeast Asia, causing more than 400 deaths. 3) In 2009, 18,500 people died of "Swine flu" or H1N1, which originated in Mexico and the U.S.. 4) In 2012, Middle East Respiratory Syndrome (MERS) emerged in Saudi Arabia, resulting in nearly 1,000 deaths. 5) In 2013, the Ebola Virus Disease emerged in West Africa and lasted more than two years, killing more than 11,300 people. A continuous process of surveillance, reporting, intervention, and resource allocation are critical to

effective response to such contagions. Countries should therefore prepare for pandemics considering their recurring nature.

Generally, individuals will experience five periods for a viral disease, including incubation, prodromal, illness, decline, and convalescence. Pandemic controls are designed to target these periods, with the aim of minimizing the impact of the disease on individuals and communities. Comprehensive surveillance, effective intervention, and optimal healthcare resource allocation can help control the spread of diseases and mitigate the economic and health consequences to human life. These measures may include contact tracing, quarantine, public education campaigns, and the development and distribution of vaccines and treatments. However, in spite of the efforts devoted to pandemic control from all over the world, there's still a knowledge gap when a new infectious disease is transmitted among the public, due to incomplete understanding of transmission methods, public health inventions, severity and vaccination. Pandemic research could help to monitor a new outbreak, evaluate the likelihood of transmission, and inform policy intervention and resource allocation, which could avoid unnecessary casualties and reduce economic hardship.

1.2 Objectives of Proposed Thesis Research

To mitigate the impact of future pandemics on social and economic life, our research will focus on enhancing surveillance, developing targeted interventions, and optimizing vaccine distribution. In the initial stages of a novel infectious disease, the lack of biological and medical knowledge makes it difficult to estimate the extent and rapidity of its spread. A primary job is ascertaining the case and fatality patterns of the disease based on which we could evaluate the transmissibility and severity of the diseases.

Moreover, the spread of an epidemic hinges on the probability of infection and the nature of individual interactions. Mobility networks are instrumental in shaping the temporal and spatial

dynamics of disease transmission within populations. The rising global mobility of humans and increased trade volumes facilitate the introduction of infectious diseases to new regions. Consequently, our research will employ mathematical modeling that incorporates time-varying transmission and fatality rates, as well as spatial interactions, to create a comprehensive representation of transmission dynamics. This modeling approach will generate quantitative predictions to inform health policies, enabling evaluation of epidemiological outcomes and the effectiveness of intervention strategies.

Uncertainties during the early stages of infectious diseases, such as insufficient testing and delayed reporting, can hinder the understanding of a disease's severity and potentially postpone effective measures to curb large-scale infections. Therefore, another crucial aspect is to evaluate the likelihood of a disease evolving from a local outbreak to a global pandemic. In addition, it is essential to account for uncertainties stemming from model parameter estimation and variations in global events and intervention policies when analyzing disease dynamics. To address these issues, our research will apply probabilistic uncertainty analysis to a dynamic model of viral disease transmission, assessing the uncertainty in the healthcare outcomes.

Global pandemics can place immense strain on healthcare systems, necessitating efficient allocation of limited resources. To enable an effective medical response, healthcare systems must be well-prepared and organized, considering their capabilities and capacities. Our research will investigate the optimal distribution of health resources, leveraging reliable historical data from uncertainty analyses. Ultimately, our research aims to develop a comprehensive surveillance, evaluation, and guidance platform to better prepare for and respond to future pandemics.

To summarize, the specific aims of the research will be:

Aim 1: To comprehensively understand the transmission dynamics of a new infectious disease, we aim to accurately predict time-varying transmission and fatality rates with a variation of the Susceptible-Exposed-Infectious-Recovered-Death (SEIRD) model. We will propose a continuous representation of the transmission rate and case fatality with a small number of parameters. The model will be validated by checking the goodness of fit to historical data.

Aim 2: To enhance the credibility of the model from Aim 1, we aim to investigate and quantify the effect of transportation and vaccine allocation. To accomplish this aim, we will incorporate mobility data and vaccination data into the dynamic model from Aim 1 through introduction of variables estimating interregional movement and vaccination rate. Based on this comprehensive model, we will analyze the relationship of transmission dynamics with respect to the mobility pattern and vaccination progress.

Aim 3: To find the optimal vaccine allocation strategy across different phases of the epidemic, we aim to dynamically optimize vaccine allocation to maximize the health benefit with equity. To achieve this goal, we will define a multi-objective function considering both the total infections and deaths over the whole period of epidemic with the constraints of limited vaccine capacity. Under reasonable assumptions, the optimal allocation strategy will be solved adaptively.

Aim 4: To inform the choice of intervention policies for decision makers, we aim to evaluate the healthcare outcomes under different policy scenarios. Based on the model from Aim 3, we will evaluate the effect of interventions on the spread of disease with uncertainty. According to the evaluation results, we will assess effectiveness and uncertainty of intervention methods at different phases of an epidemic.

In summary, by accomplishing the four objectives outlined above, we will establish a comprehensive modeling framework for analyzing transmission dynamics and optimizing vaccine

allocation. This framework, combined with uncertainty analysis, will enable us to evaluate the scale of an epidemic and the effectiveness of intervention strategies. By accurately describing historical trends under various scenarios, our approach to vaccine optimization will offer decision-makers the most effective vaccine allocation plan in the face of uncertainty. Ultimately, this research will provide a systematic methodology for monitoring, predicting, and responding to future epidemics, enhancing preparedness, and promoting effective public health responses.

Chapter 2. Background and Prior Research

2.1 Transmission Modeling for Infectious Diseases

2.1.1 Models of Community Spread of Disease

Modeling, analyzing, and predicting the impact of infectious diseases is critical to preventing, controlling, and managing the spread of infectious agents. It is a common perception that infectious diseases can be transmitted from person to person through a pattern of space and time. In addition to vaccination and medication, a great amount of effort has been put into the modeling and analysis of the temporal and spatial dynamics of disease spreading in order to better understand the complex patterns of infectious diseases and devise effective intervention strategies.

The history of modern epidemiological analysis and modeling can be traced back to the late nineteenth and early twentieth centuries, when early mathematical models of disease transmission were developed. One of the earliest examples is John Snow's work in 1849, when he plotted the cholera epidemic cases on a map and concluded that contaminated water was the predominant contributor to cholera transmission in London [1]. In a similar vein, Arthur Ransome developed a discrete-time epidemic model for cholera transmission in 1906 [2]. These early mathematical models of disease transmission have been helpful in gaining insights into the transmission dynamics of infectious diseases and the potential role of different intervention policies. Over the past few decades, various models have been developed and applied to modeling the transmission dynamics of infectious diseases, including metapopulation models, statistical models, and agent-based modeling, all of which have contributed to a more comprehensive understanding of the spread of diseases.

Compartmental models are widely used in epidemiology to simulate the spread of infectious diseases and understand their dynamics. These models divide the population into distinct compartments based on their health status, such as susceptible, infected, and recovered individuals, and describe the flow of individuals between these compartments using differential equations. In this section, we will discuss common compartmental models, their advantages and disadvantages, and their applications in various epidemiological contexts.

One of the most successful models is Susceptible-Infectious-Recovered (SIR) Model, described by Kermack and McKendrick in 1927 [3]. Originally, This model was proposed to explain the rapid rise and fall in the number of infected patients of plague [4]. This type of model is based on our intuitive understanding of the epidemic transmission and compromises three categories of individuals: those who are infectious (I) mix among those who are susceptible to disease (S) and transmit the disease to the susceptible population in a stochastic process, and those who are infected will recover from the infection and acquire immunity to the disease. People may progress between compartments, the dynamic of which is illustrated by a series of ordinary differential equations in Equation (1):

$$\begin{aligned} \frac{dS}{dt} &= \beta \frac{SI}{N} \\ \frac{dI}{dt} &= \beta \frac{SI}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{1}$$

where S, I and R denote the susceptible, infectious and recovered population respectively; N is the total population; β is the transmission rate, indicating the average effective infection transmitted by one infectious person; γ is the recovery rate. By adjusting the β and γ, the differential equations 1 could capture the initial rapid increment of infectious population (I) of a given epidemic and decrement of the number of individuals after the epidemic peak. The compartmental model could be used with a stochastic framework to better describe the reality of the transmission. The SIR

model is particularly useful for studying diseases that confer long-lasting immunity, such as measles or chickenpox. It has been widely applied to understand the basic reproductive number (R_0) of various diseases, which is the average number of secondary infections caused by a single infectious individual in a fully susceptible population. This parameter is essential for understanding the potential for disease spread and evaluating the effectiveness of control measures.

The Susceptible-Exposed-Infectious-Recovered (SEIR) Model is an extension of the SIR model that incorporates an additional compartment for exposed individuals (E) [5]–[8]. Exposed individuals are those who have been infected but are not yet infectious themselves. The inclusion of this compartment accounts for the incubation period of the disease, allowing for a more accurate representation of diseases with a delay between infection and infectiousness, such as COVID-19, Ebola, or SARS.

The SEIR model's additional compartment enables it to more accurately represent the dynamics of diseases with latent periods, providing a more detailed understanding of the transmission process. It has been used in various epidemiological studies to investigate the spread of infectious diseases and evaluate the impact of control measures, such as quarantine and isolation [9].

The Susceptible-Infectious-Recovered-Death (SIRD) Model expands upon the SIR model by including a death (D) compartment to account for disease-related fatalities. This model is particularly useful for studying diseases with significant mortality rates, as it allows for the estimation of both infection and death rates within a population.

The SIRD model has been applied to various infectious diseases with high mortality rates, such as H1N1 influenza, to understand the dynamics of disease transmission and death rates [10]. By incorporating the death compartment, the model can provide insights into the severity of an

outbreak and help inform public health decision-making regarding resource allocation, healthcare capacity, and intervention strategies.

The Susceptible-Exposed-Infectious-Recovered-Death (SEIRD) Model combines the features of the SEIR and SIRD models, adding both an exposed (E) and a death (D) compartment. This model provides a more comprehensive representation of disease dynamics, accounting for the incubation period and disease-related fatalities. The SEIRD model has been used to study various infectious diseases, including COVID-19, where both latency and mortality play significant roles in disease dynamics. The model allows researchers to better understand the factors influencing the spread and severity of a disease, enabling the development of more targeted and effective intervention strategies.

In addition to the basic compartmental models discussed above, several variations and extensions have been developed to address specific epidemiological questions or account for additional factors that influence disease transmission. Some of these variations include:

Age-structured compartmental models divide the population into age groups, reflecting the fact that disease transmission and severity can vary significantly by age. For example, younger individuals may have a higher risk of infection due to increased social contacts, while older individuals may experience more severe outcomes. Age-structured models have been used to study diseases such as influenza, pertussis, and COVID-19, helping to inform age-specific vaccination strategies and public health interventions [11]–[13].

Multi-strain compartmental models are used to study diseases caused by multiple strains or subtypes of a pathogen, such as influenza or dengue fever [14], [15]. These models can help researchers understand the dynamics of strain competition and the potential for strain replacement following vaccination or other interventions. Multi-strain models can also provide insights into the

evolution of pathogens and the emergence of new strains, informing the development of vaccines and other control measures.

Network-based compartmental models incorporate social network structures to account for heterogeneous mixing patterns within a population. These models can more accurately represent the spread of diseases in populations with complex social structures, such as schools, workplaces, or communities. Network-based models have been used to study the spread of sexually transmitted infections, respiratory infections, and other diseases where contact patterns play a critical role in transmission dynamics [16], [17].

Spatially explicit compartmental models incorporate geographic information to account for the spatial distribution of individuals and the impact of local population density, environmental factors, and human mobility on disease transmission. These models can help researchers understand the role of spatial factors in disease spread and inform targeted interventions, such as travel restrictions, quarantine measures, or targeted vaccination campaigns. Spatially explicit models have been used to study a wide range of infectious diseases, including malaria, cholera, and COVID-19 [18]–[20].

Strengths of compartmental models are numerous, with one of the most significant advantages being their simplicity and flexibility. This allows researchers to easily adapt the models to a wide range of diseases, populations, and research questions by modifying the number of compartments, parameters, or assumptions. Additionally, compartmental models are mathematically tractable and computationally efficient, making them particularly suitable for large-scale simulations, real-time forecasting, and situations with limited computational resources. Furthermore, these models provide a solid foundation for more complex models, enabling researchers to systematically build upon and refine the basic compartmental structure to better represent specific aspects of a disease or population of interest. Importantly, the parameters and variables in compartmental models often

have clear biological or epidemiological interpretations, making the models easy to understand and communicate to a broad audience, including policymakers, public health officials, and other stakeholders who may not have an extensive background in mathematical modeling [21].

In spite of the strengths of the compartmental model, the limitations of the compartmental model have gained growing attention. One primary drawback is the homogeneous mixing assumption, which can be unrealistic in real-world populations with complex contact patterns due to factors such as age, social networks, or spatial distribution. This assumption can lead to inaccurate predictions of disease spread, particularly for diseases where contact patterns play a critical role in transmission dynamics. Another limitation is the lack of spatial consideration in most basic compartmental models, which can restrict their ability to accurately represent the spread of diseases with strong spatial components, such as vector-borne diseases or those influenced by environmental factors. While spatially explicit compartmental models have been developed to address this issue, they can be more complex and computationally demanding. Compartmental models can also oversimplify disease dynamics by neglecting important factors such as individual heterogeneity, which can result in an inaccurate understanding of the true dynamics of disease transmission. Lastly, parameter uncertainty is a significant concern in compartmental models, particularly for emerging diseases with limited data available. This uncertainty can limit the reliability of model predictions and may impact their usefulness in informing public health decision-making. Researchers must carefully assess parameter uncertainty and its potential impact on model results and should consider using techniques such as sensitivity analysis or Bayesian approaches to account for this uncertainty.

Besides the compartmental modeling, various statistical analysis methods have also been applied to the modeling of infectious diseases. Usually, statistical modeling formalizes relationships

among variables that may influence the spread of disease, describes how the variables are related to each other, and test the hypothesis or statements about the disease transmission. Statistical models could involve the statistical analysis and modeling of the disease observation with a broader variety of variables including the special factors, economy factors, and environmental factors.

Various complicated statistical models have been constructed. For example, time series analysis is a statistical method that focuses on analyzing data points collected over time to predict future trends. This approach is useful for understanding temporal patterns in infectious diseases, such as seasonality or long-term trends [22]. However, the main drawback of time series analysis compared to compartmental modeling is that it does not consider the underlying disease transmission mechanisms, making it challenging to identify specific factors driving the spread of the disease or evaluate the impact of control measures.

Regression models, including linear regression and logistic regression, are used to establish relationships between disease spread and various factors, such as demographic characteristics or environmental factors. They can help identify risk factors and guide public health interventions. For example, generalized linear mixed model (GLMM), an extension of linear mixed models, could incorporate both individual level and integrated level data by allowing the response variables from different distributions, such as binary responses [23].However, compared to compartmental modeling, regression models are limited by their reliance on a fixed functional form, which may not accurately capture the complex dynamics of disease transmission. Additionally, they may not account for spatial dependencies, limiting their applicability in spatially heterogeneous settings.

Bayesian hierarchical models offer a flexible framework for incorporating various sources of information and uncertainty in infectious disease modeling [24]. These models can account for

both spatial and temporal dependencies and integrate multiple data sources. However, the main drawback of Bayesian hierarchical models compared to compartmental modeling is their computational complexity. As these models often involve a high number of parameters and require Markov chain Monte Carlo (MCMC) methods for estimation, they can be computationally intensive and time-consuming, especially for large-scale applications [25].

Generally, statistical models are flexible regarding the format of the input data and the parameters selection, which makes them more powerful for analyzing disease on a detailed level. However, statistical model is a data-based method, requiring a large variety and quantity of high-quality data. Missing data, underreporting, and uncertainty are common for the disease data, especially at the early stage of the transmission. In practice, it is challenging to build a complicated statistical model incorporating different aspects of transmission without the right data. Meanwhile, some of the advanced statistical methods, like Markov Chain Monte Carlo, is usually very computationally expensive and time consuming while the model considers many factors.

With the development of computational technology and accessibility to big data, computational simulation approaches have been used increasingly in the epidemic analysis and forecasting. One of the popular beliefs of the disease transmission is that population is heterogeneous and each individual exhibit unique spread pattern. Agent-based modeling (ABM) is a powerful computational tool for simulating complex systems, including infectious diseases. This approach allows for the representation of individual agents, such as people or animals, and their interactions within a population [26]. Different from the metapopulation model, agent-based models assume: (a) individuals are characterized uniquely by their age, race, income, and so on;(b) the number of interactions varies by person; (c) individuals are spatially distributed and mobile.[27] Based on the

assumptions mentioned above, agent-based models represent each individual uniquely by their characteristics, behaviors and interaction with others.

Object-oriented modeling is one of the most popular and widely used agent-based modeling approaches.[28] This approach constructs the agents using a collection of features and organize the agents in a hierarchical way. Bian used the object-oriented GIS framework to model the individual fish movement and growth in a heterogeneous aquatic environment. As shown in object-oriented modeling, agent-based modeling could provide a clear framework to incorporated individual characteristics, individual interaction patterns, which ultimately in total determine the transmission results [28]. Agent-based model, combined with social contact networks, have been widely used to simulate influenza-type diseases like West Nile virus in the United States and Canada with the agent representing mosquitoes, avian hosts, mammalian host, and human.[29] These models simulate the interaction patterns between the agents by considering habitat location and weather conditions.

One of the primary strengths of agent-based modeling is its ability to represent heterogeneity within a population, including variations in demographics, behavior, and disease susceptibility. This flexibility can lead to more accurate predictions and better understanding of the dynamics of infectious diseases. However, a drawback of this approach compared to compartmental modeling is that it can be computationally expensive, particularly for large populations or when simulating detailed individual-level interactions.

Meanwhile, agent-based modeling can explicitly model individual behaviors and their impact on disease transmission, such as social distancing or vaccination uptake. This capability allows researchers to study the effects of behavior change interventions on disease spread. However, compared to compartmental modeling, a limitation of agent-based modeling is that it often relies

on assumptions about individual behaviors and their interactions, which may not always be accurate or generalizable to other populations.

Agent-based models are well-suited for capturing spatial dynamics in infectious disease transmission, as they can represent individuals' locations and the impact of spatial structure on disease spread. This ability can lead to more accurate predictions of disease spread in spatially heterogeneous environments. However, compared to compartmental modeling, a drawback of agent-based modeling is that they may require detailed spatial data, which can be difficult to obtain or may be subject to privacy concerns.

Agent-based methods in infectious disease modeling offer several advantages over compartmental modeling, such as capturing heterogeneity, modeling individual behaviors, and representing spatial dynamics. However, these methods also have their drawbacks, such as computational expense, reliance on assumptions about individual behaviors, and the need for detailed data. In the cases when new diseases first appear, compartmental modeling, which simplifies the disease transmission process by grouping individuals into compartments, may offer a more computationally efficient and easily interpretable approach.

2.1.2 Spatial Interaction Modeling in Epidemic Modeling

Commuting patterns and travel behavior underlies the spread of infectious diseases among locations. Spatial interaction models of human mobility have been used in the study of the epidemic transmission dynamics. While the temporal dynamics of epidemics have always been a primary focus of most models, over the years, efforts have been made to model spatial dynamic in epidemic processes. The efforts could be divided into two parts: (1) modeling of spatial interaction between different groups of research units, like livestock or human. (2) incorporating the spatial interaction into the epidemic models. In this part of review, we will first review the most common

interaction models used in epidemic modeling and then review the main categories of spatial oriented epidemic model.

2.1.2.1 Spatial Interaction models

For the consideration of the spatial dynamic of the epidemic, one of the main focuses is the modeling of the spatial interaction among modeling units. Many spatial interaction models have been proposed to uncover the relationships between spatiotemporal infectious disease patterns and environmental characteristics. These models rely mostly on two frameworks, the gravity model and radiation model. In the absence of easily available data on travel behavior, these models intuitively describe the mobility in the population.

The gravity model is the most common spatial interaction model, based on the principle of gravity as described in Issac Newton's law of gravity. The model has two basic assumptions that the interaction between two places is proportional to the product of population totals of the places and inversely proportional to the square of distance. The gravity model is generalized as follows in Equation (2) [30]:

$$I_{ij} = k\frac{p_i . p_j}{D_{ij}^b} \tag{2}$$

where p_i is the population size in region i, I_{ij} is the estimate of the volume of spatial interaction between place of origin i and place of destination j, k is a constant scaling factor, D_{ij} is the distance between i and j, b is the distance effect acting as an impedance on the interactions. Many empirical studies have been conducted on spatial interaction models. For example, different measures of distance (straight-line distance, Euclidean distance, city-block metric and travel time) have been suggested. Meanwhile, different functions for distance effects and extra components have been proposed to specify the attractiveness of the destination considering the benefits of economies scale.

The advantages of using gravity approach for modeling transition process during the epidemic are the ability to explain interregional trade patterns under the conditions of comparatively sparse data and the validity of theoretical background of the model to human interactions in a large scale.

In spite of the advantages of the gravity model, the spatial interaction model is typically based on aggregate, zonal data under the assumption of homogeneity. The model might fail to accordingly represent description for regions with high heterogeneity and uncertainty[31]. Meanwhile, the gravity model requires tedious tunning process based on the real observations. However, in reality, mobility data are difficult to complete.

To overcome limitations of gravity models, Simini [32] proposed the radiation model, which originally appeared in physics to study the process of energetic particles or wave travel through vacuums. The radiation model describes the mobility patterns without any parameter estimation. The traffic from site i to j, with populations P_i and P_j, is given by Equation (3).

$$T_{ij} = T_i \frac{P_i P_j}{(P_i + s_{ij})(P_i + s_{ij} + P_j)} \quad (3)$$

Where $T_i = \sum_{j \neq i} T_{ij}$ is the outgoing traffic from the site i. Compared to the gravity model, the radiation model has broader flexibility without the need of parameter estimation, and better prediction for long distance travel. However, as formulated in equation (3), the radiation model only considers the large-scale parameters and does not incorporate the parameters to calibrate for smaller scales. Thus, the radiation model has relatively poor predictability on short-distance travel [33]. Moreover, the radiation model requires additional information on T_i, in contrast to the gravity model.

The radiation and gravity models have been compared to each other in terms of the predictability of mobility patterns observed in various empirical data sets.[33]

In terms of the application of the spatial interaction model in the epidemic research. Li et al. [34] validated the performance of the gravity model in predicting the global spread of H1N1 by formulating the global transmission between major cities and Mexico. The variables, such as population sizes, per capita gross domestic production (GDP), and the distance between the countries and Mexico were incorporated in the gravity model, which is validated by the estimation of global transmission trend. Kraemer et al. [35] described a flexible transmission model to test the utility of generalized human movement models in estimating Ebola virus disease (EVD) cases and spatial spread over the course of the outbreak. Gravity model, radiation model and adjacent network model were applied to show the generalized movement models have the ability to improve the modeling of spread of EVD epidemic.

2.1.2.2 Spatial oriented models

Besides modeling of spatial interaction, another concern is the design of the spatially oriented epidemic models incorporating different kinds of spatial interactions. The spatially oriented models could be categorized according to the scale of the modeling unit and the mobility of the modeling units. According to the scale of the modeling unit, the models could be divided into three types: (1) Population-based models, (2) Sub-population models, and (3) Individual-based models. In each category, models could be further divided according to the level of mobility.

1) Population-based models

Population-based models divide the whole population into different segments with various assumptions of the characteristics of each segment. The simplest version of the family of population-based models is the SIR model, which consists of three exclusive segments: Susceptible(S), Infectious(I), and Recovered(R). However, the classic population-based models do not consider spatial dynamics at all. In the 1980s, geographers proposed a spatial framework

for epidemiological models that explicitly considers the spatial dispersion of infectious diseases[36]. A simple form of these spatial models is a three-ring wave model, where the first infection case occurs at the center of a space and spread towards all directions like a wave. The infection segment stays at the center of the wave, the susceptible segment surrounds the infection segment, and the recovery segment is at the outermost ring surrounding the infection segment[36], [37]. The location of the three-ring changes dynamically as the infectious wave spread through the space. The three-ring wave model differs from the classic model by how it projects the three populations into space, which also marks the beginning of spatially explicit epidemiological models. However, like the other population-based models, the inherent homogeneity assumptions impede the models' ability to describe the heterogeneity in disease transmission[38]–[40]. One application for population-based wave models is studying pandemic waves in a large space level, such as the 1918-1920 Spanish flu that spread globally [41].

2) Sub-population models

Instead of assuming the whole population as an identical model unit, sub-population models divide the population into a substantial number of subpopulations. The increased number of model units (i.e., the subpopulations) could increase the heterogeneity considered in the model in order to produce more realistic results than those produced by the classic population-based models. One branch of the sub-population models divides the population according to the spatial distribution of each group, called spatially structured models[37], [42], [43]. Spatially structured models usually divide the space into regular grid cells. Each cell, containing one subpopulation, inherits the same identical and homogeneously mixed assumption as the classic population-based model. With the increased number of spatially distributed units, the interactions between the subpopulations are also added to the model enhancing the model ability to account for the heterogeneity in certain

levels. Due to the homogeneity assumption within each cell, the sub-population models suit more for the modeling of diseases between high-density, immobile communities, such as the diseases transmitted in livestock[44], [45]. For example, Doran and Laffan implemented the SIR model in the spatially structured sub-population framework to evaluate the season impact on the transmission of foot and mouth disease between different groups of livestock [44].

To address the limitations of the classic models, a variety of revisions have been used to improve the sub-population models. One revision is adding heterogeneity within a subpopulation according to the characteristics of subpopulations (such as population size, age structure, and education level). Different from the individual-based models, these characteristics are only represented in a statistical way (e.g., statistical distribution) across each subpopulation. Another revision is the finer representation of the transmission between subpopulations by statistical representation of the interaction between subpopulations or explicitly tracking of the exchanging individuals [46]. Subpopulations could then become mobile and move to different locations. One or multiple infected subpopulations could split and then merge with other subpopulations, where the location of units are explicitly recorded. In this way, diseases are transmitted across space through time. Specifically, the homogeneity assumption within each subpopulation still holds for the revised models. Hence, these models are more effective when implemented in mobile and high-density population, such as military or refugee camps, where the subpopulations will merge and split dynamically in the transmission process. [46], [47].

3) Individual-based models

Individual-based models could be considered as a further extension of the sub-population models by dividing the whole population of the individual level [41]. In addition to treating individuals as the modeling unit, the interactions between the units are explicitly represented. Individual-based

models break the homogeneous assumption held by the population-based models or sub-population models. Each individual is unique, interacts with a limited number of other individuals, and collectively contributes to the dynamic pattern of transmission in a population level. For example, when the individuals are immobile and only interact with the adjacent individuals, Martins et al. proposed a cellular automata model to simulate the progress of the citrus variegated chlorosis disease by adjusting the parameters controlling vectors motility, plant stress, and initialization[48]. When the disease transmission depends heavily on the structure of social networks and the spatial movement of the units is not explicit represented, Eames and Keeling developed pair-wise network equations utilizing the essential characteristics of the mixing network to estimate the effectiveness of various control strategies towards sexually transmitted diseases[49]. However, individual-based models require a great amount of detailed information about individuals, which is hard to obtain at the early stage of the epidemic. To compensate for this scarcity of data, efforts have been made to investigate individualized contact behavior from mobile device data[50]. The development of the individual-based models relies on the build-up of a secured, united, real-time information tracking system.

In the section, spatial oriented models are categorized according to the scale of the modeling units and illustrated by their major implementation on the analysis of spatial dynamics in the epidemic process. The model selection should depend on the characteristics of research target, availability of the data and the scope of the analysis.

2.1.2.3 Applications of Spatial Interaction Models in Epidemiology

Spatial interaction models have demonstrated significant potential in various aspects of infectious disease epidemiology, from disease spread modeling to healthcare resource allocation and surveillance. This part will dive deeper into the applications of spatial interaction models in

epidemiology, examining their role in enhancing epidemic management and highlighting opportunities for future research.

1) Disease Spread Modeling

Spatial interaction models simulate the spread of infectious diseases over space and time by incorporating the movement of people, goods, and services [51]. These models are built on the foundation of gravity and entropy maximization principles [52], which assume that the interaction between two locations is proportional to the product of their masses (e.g., population size) and inversely proportional to the distance between them. Recent developments in spatial interaction modeling have incorporated additional variables, such as socio-economic factors, transportation infrastructure, and regional characteristics [53].

Spatial interaction models were used to study the global spread of the 2009 H1N1 influenza pandemic [54]. By incorporating air travel data and population distribution, the models successfully predicted the initial wave of the pandemic and provided valuable insights into the potential effectiveness of various intervention strategies, such as travel restrictions and vaccination campaigns. During the Ebola outbreak in West Africa (2014-2016), spatial interaction models were employed to analyze the spread of the virus and inform containment measures [55]. The models incorporated data on population movement, healthcare infrastructure, and social factors, enabling a better understanding of the epidemic's dynamics and informing the deployment of resources to affected regions. The COVID-19 pandemic has underscored the importance of spatial interaction models in predicting and managing the spread of infectious diseases [56]. These models have been used to assess the role of international travel in the early stages of the pandemic and to evaluate the potential impact of different containment measures, such as social distancing and lockdowns, on disease transmission.

2) Healthcare Resource Allocation

Spatial interaction models can help identify areas with a high demand for healthcare services, enabling more efficient allocation of resources during an outbreak [57]. By accounting for factors such as population density, disease prevalence, and transportation networks, these models can predict the spatial distribution of healthcare needs, informing decisions on resource allocation and facility planning.

In malaria-endemic regions, spatial interaction models have been employed to optimize the allocation of resources for malaria control interventions, such as insecticide-treated bed nets and indoor residual spraying [58]. By incorporating data on population distribution, environmental factors, and transportation networks, these models can help identify areas with the highest potential impact of interventions, thereby maximizing their cost-effectiveness. During the COVID-19 pandemic, spatial interaction models were used to identify areas with high potential for hospital overload, allowing policymakers to prioritize resources and implement containment measures [59]. These models incorporated data on population distribution, healthcare infrastructure, and mobility patterns, enabling a more targeted response to the crisis.

3) Surveillance and Early Detection

Spatial interaction models can be used to optimize surveillance networks, ensuring that resources are allocated efficiently to detect and monitor disease outbreaks [60]. By incorporating data on population movement, transportation networks, and disease prevalence, these models can identify high-risk areas and vulnerable populations, enabling the design of more effective surveillance systems.

Spatial interaction models can help detect the early stages of an outbreak by modeling the movement of people and the flow of information [61]. Early detection is critical for implementing

timely interventions and preventing the spread of infectious diseases. These models can also be used to identify potential hotspots for targeted interventions, such as vaccination campaigns and public health messaging.

2.1.2.4 Challenges for spatial interaction modeling in epidemiology

Spatial interaction models offer significant potential for epidemic forecasting and management. However, addressing several challenges is essential to fully harness their potential. Firstly, obtaining high-quality, real-time data on population mobility and transportation networks is critical for accurate spatial interaction modeling [62]. As data availability continues to expand through mobile devices and social media, future research should prioritize improving data quality and incorporating these novel data sources into models.

Secondly, ensuring the accuracy of predictions and the reliability of policy recommendations in epidemiological models requires thorough validation and calibration using historical data [63]. Future research should focus on devising robust validation and calibration techniques, as well as incorporating uncertainty quantification into spatial interaction models.

In conclusion, the promotion of interdisciplinary cooperation among epidemiology, transportation, and urban planning domains has the potential to significantly augment the advancement of thorough spatial interaction models. Facilitating such partnerships may ultimately result in more efficient tactics for managing epidemics [64].

2.2 Vaccine Management and Prioritization during Pandemics

In recent years, large outbreaks of new emerging epidemics, such as the SARS-CoV-2 virus in 2019, the Ebola virus in 2014, and the HINI virus in 2009, bring about deaths, health losses and economic damage. To combat the potential crisis caused by emerging epidemics, public health interventions have been limited to the non-pharmaceutical interventions at the early stage,

including travel restriction, contact tracing and lockdown [65], [66]. While critical to slowing down the viral spread, non-pharmaceutical interventions have brought about occupation misfortune and financial difficulties [67], [68]. Given the considerable political and financial expenses related with non-pharmaceutical interventions, long-term solutions such as vaccines that protect from viral infection are needed. Vaccines provide direct protective benefits to the public, including protection from infection, reduced symptom development, and lower mortality rates. Meanwhile, vaccination can also bring about indirect benefits by decreasing the risk of exposure, even for the unvaccinated susceptible individuals.

The manufacturing of the vaccine, such as influenza vaccine, follows a tight schedule to produce sufficient doses in the initial phases[69], [70]. Developing a healthy and reliable vaccine involves many steps including initial experiments, three phases of clinical trials, US Food and Drug Administration (FDA) authorization, manufacturing and distribution. And it usually takes between 6 and 36 months to get a healthy vaccine available to the public. However, once the vaccine is authorized by the FDA, the first batch of vaccine is limited. The large number of susceptible individuals result in overwhelming demand over supply. Therefore, it is critical to find out the best decision-making method on vaccine distribution.

Two popular ways to prioritize vaccine distribution: (i) directly vaccinate the high-risk individuals for severe outcomes, such as people over 65 years old, and (ii) indirectly protect them by vaccinating individuals with the highest numbers of potentially disease-causing contacts. Most of the model-based vaccination allocation research investigates trade-offs between these two strategies. Considering the high contact rate related to healthy school children and their expected remaining life, epidemiologists have suggested vaccinating school age children first to moderate the spread of disease and thereby indirectly decrease mortality [71]–[73]. Monto et al. vaccinated

85% of younger students in Tecumseh, Michigan with inactivated flu vaccine in 1968 and discovered two-third lower rate of illness than the neighbor town of Adrian during a rush of flu A [72]. Reichart et al. argue that routine influenza vaccination of school-aged children in Japan from 1962 to 1994 has prevented 10,000–12,000 deaths annually from pneumonia and influenza [74]. Longini et al. utilized stochastic epidemic simulations to investigate the effectiveness of target antiviral prophylaxis to contain influenza. They found that vaccinating 80% of all school-age children is almost as effective as vaccinating 80% of the entire population [71]. In addition, modeling studies have estimated that annual influenza epidemics could be contained if 50%–70% of children were vaccinated [75], [76]. Moreover, similar modeling analysis has found that the optimal balance between direct and indirect protection depends on the reproduction numbers of the disease and vaccination efficiency. Bansal et al. have shown that direct protection strategies outperform the other if the transmission level (reproduction number) is moderate, while the reverse is true for highly transmissible diseases [77].

Considering the limited availability of vaccine stockpiles, several studies focus on the optimal allocation policies to contain an outbreak in its early stages using the smallest amount of vaccine. One general mathematical formulation adopted by researchers is formulated in Equation (4) [78]–[81].

$$\begin{aligned} & minmize \sum_i n_i f_i \\ s.t. \quad & R_f \leq 1 \\ & 0 \leq f_i \leq 1 \end{aligned} \tag{4}$$

where n_i refers to the number of people in each subgroup, f_i denotes the vaccine coverage in subgroup i. R_f denotes the effective reproduction number, the expected number of secondary infections transmitted by an infectious individual. The epidemic will prevail if $R_f > 1$ [79]. Hence, the first constraint in (4) forces $R_f < 1$, which is non-convex or non-concave Hill and Longini

characterized the threshold surface of critical vaccine allocations in a population with interacting subgroups [78]. They proposed a solution approach based on Lagrangian multipliers applied with Maple to minimize the amount of distributed vaccine. They pointed out that R_f is equal to the spectral radius of the product matrix of vaccination fraction matrix and next generation matrix, which has been widely adopted in later research. Duijzer et al. utilized the SIR compartmental model considering the geographically separated subgroups and interactions to analyze the relation between the vaccination fraction and herd effect [80]. They conclude that to successfully alleviate a continuous episode with a restricted vaccine reserve, policymakers should focus on subgroups where the disease has not made much progress yet. However, they solved the optimization model with a nonlinear programming solver that does not guarantee an global optimal. Based on similar assumptions, Enayati and Ozaltin further improved the solution quality by discretization and reformulation of the constraints. They iteratively solve the upper and lower bound of the mixed-integer program to find the approximated global optimum.

Besides the problem of containing the epidemic with the minimal vaccine, another branch of research is seeking to find the optimal allocation policy to get the best healthcare outcomes given limited stockpile of the vaccine. The general formulation for this type of question is formulated as below in Equation (5):

$$\begin{aligned} &minmize \sum_i \sigma_i \, n_i w_i \\ &s.t. \quad \sum_i v_i \, n_i \leq V \\ &0 \leq v_i \leq 1 \end{aligned} \tag{5}$$

where σ_i measures the proportion of people get infected among subgroup i, n_i refers to the population of subgroup i, w_i denotes the weights assigned to each subgroup, v_i refers to the vaccination proportion in each subgroup and V is the total stockpile of the vaccination. Under this optimization framework, Patel et al. utilized genetic algorithms and random mutation hill climbing

to find optimal distributions with the stochastic epidemic simulations to minimize the number of infections and deaths. However, due to the non-linearity and stochasticity of the epidemic models, both of the algorithms could only approximately find the local optimum. Medlock and Galvani developed a compartmental model that incorporates both age groups and vaccination status for the outbreak of swine-origin influenza in 2009 [73]. They enumeratedly compared different vaccine allocation policies by the amount of vaccine needed to reduce the effective reproduction number less than one. Dalgic et al. takes advantage of the mesh-adaptive direct search (MADS) algorithm to numerically compare the effectiveness of vaccine policies derived from agent-base models and compartmental models for influenza pandemic. Starting from the initialization, the MADS algorithm iteratively enhances the current best solution by generating test points on the variable space around current best solution. The algorithm will decrease the step size if it fails to find a better solution until the limits of the mesh size has been reached. The MADS algorithm used in this study is derivative free, which makes it more flexible when the gradient of the epidemic models are not available.

As pointed out in previous research, vaccine allocation problems are usually non-convex (or non-concave) due to the nonlinearity, complexity, and stochasticity of the epidemic. Therefore, the derivative of the model may not be calculated analytically, which makes the gradient-based method inapplicable. Most algorithms will converge to a local optimum.

For the modeling of COVID-19 vaccination, Matrajt et al. used an age-stratified mathematical model paired with optimization algorithms to study the relationship between the effectiveness of the vaccination and health outcomes [82]. They pointed out that direct vaccination of elders brings out the best health outcomes when the vaccination efficiency is low or the stockpile of the vaccine is limited. Gallagher et al. emphasized that the indirect benefits of the SARS-CoV-2 vaccines has

the potential to quell the pandemic even the vaccine has a weaker direct protection but stronger indirect effects [83]. Under the constraints of limited availability, uncertain vaccination efficiency, and ethical equity problems, a fair and credible decision-making method with feasible optimal is still well-motivated.

Unlike what has been done before for influenza, the vaccine of COVID-19 could only be available after the epidemic is widely spread. Hence, vaccine allocation should consider the severity of the infection and the regional characteristics (such as age structure) in each region. Meanwhile, different from previous research, the manufacturing and distribution of the COVID-19 vaccine should be represented as a continuous process. The allocation of the vaccine is then a dynamic process. The optimal dynamic allocation policy for the vaccine to minimize the overall burden of disease under the constraints of limited vaccine is a serious and urgent problem to be addressed.

2.3 Uncertainties Analysis in Epidemic Modeling

Epidemiological models serve as vital tools in understanding and forecasting the dynamics of infectious diseases, which in turn enables policy makers to effectively allocate resources and implement intervention strategies. However, the credibility of these models is contingent upon the quality of data used for parameter estimation, which is influenced by factors such as monitoring, timeliness, privacy restrictions, and reporting accuracy. Thus, addressing uncertainty in model predictions is essential. By conducting parameter sensitivity and uncertainty analyses for model outcomes, confidence in result interpretation and decision-making can be bolstered. To achieve best practice in uncertainty analysis, general modeling standards recommend employing probabilistic sensitivity analyses that address both global parameter uncertainty and output uncertainty [84].

Previous uncertainty analysis applications in dynamic models can be classified according to transmission models. The deterministic SIR model [85]–[87] serves as the most prevalent approach for evaluating interventions in dynamic systems. Although deterministic SIR models provide useful intervention effect estimates, the optimal parameter estimates are often imprecise and non-unique. To address this issue, sensitivity analyses explore the relationship between model parameters and corresponding outcomes [88]. These analyses typically involve perturbing one or more parameters and investigating the resulting outcomes. The perturbation can either involve assessing the impacts of arbitrarily small changes in parameter values [89] or evaluating the effects across a range of realistic probability density functions [90]. For example, Zhang et al. applied Sobol's method to a compartmental COVID-19 model to identify key model parameters and controlling parameters, which can help policy makers explore various intervention options [91]. Similarly, Sarabaz et al. utilized the SimBiology Toolbox in MATLAB with three different techniques to analyze parameter sensitivity in a system of differential equations built for the same purpose of identifying key spreading parameters [92]. However, these methods fail to account for interaction effects in non-linear dynamic models and do not evaluate global uncertainty in parameters or outcomes since other parameter values remain constant at their best point estimations.

In contrast to deterministic dynamic model sensitivity analyses that only address variation in each parameter, global probabilistic sensitivity analyses examine each parameter's contribution to model outcomes while accounting for the uncertainty of other model parameters [93], [94]. This allows modelers to convey the robustness of their predictions to policy makers. Uncertainty in parameter values can be addressed by randomly sampling from empirical data or fitting probability density functions to empirical data using methods such as bootstrapping, Monte Carlo sampling,

and Latin hypercube sampling [88], [93]. To assess each parameter's contribution to the variance in output values, model output from parameter samples can be examined using linear, monotonic, and non-monotonic statistical tests [88], [93]–[95]. By incorporating data-driven parameter uncertainty into model outputs, probabilistic uncertainty analyses produce probabilistic distributions rather than single-value estimates of potential outcomes, which enables modelers to communicate the robustness of their predictions to policy makers. Amaku et al., for example, calculated the force of infection for six distinct viruses in a Brazilian community using seroprevalence study data and estimated parameter confidence intervals using Monte Carlo simulations. Similarly, Ciufolini and Paolozzi used an empirical Gauss error function model for COVID-19 cases in Italy, with 150 Monte Carlo simulations providing a more robust peak day prediction [96].

During the COVID-19 pandemic, numerous non-pharmaceutical interventions (NPIs) have been implemented worldwide to mitigate viral spread, and uncertainty analysis plays a crucial role in quantifying the effects of these NPIs. Chinazzi et al. applied the global epidemic and mobility model (GLEAM) to understand the impact of travel restrictions on COVID-19 transmission in Wuhan and assessed the robustness of their results through extensive sensitivity analysis [97]. In another study, Quilty et al. employed a stochastic branching process with a negative binomial offspring distribution to estimate the intervention effects of cordon sanitaire and holiday travel, simulating outbreaks generated by arrivals in four representative major cities [98]. Sensitivity analyses for the overdispersion factor and serial interval were performed to provide confidence intervals for the effective reproduction number. Their findings suggest that the cordon sanitaire alone did not significantly impact epidemic progression in major cities.

Children often serve as key transmitters in viral epidemics like influenza due to their high levels of direct contact with other children and parents [99]. Consequently, all 50 U.S. states implemented statewide school closures in March 2020. Evaluating the effect of school closures on viral spread control is therefore essential. Auger et al. conducted a population-based time series analysis of all 50 states from March 9 to May 7, 2020, which included a lag period to account for policy-related changes [100]. They examined the sensitivity of the lag period for incidence and mortality by varying its length, finding that school closures led to the largest reduction in incidence and mortality when implemented as early as possible. Burns and Gutfraind introduced a novel SEIR model incorporating student location and grade to compare the effect of two intervention policies in schools [101]. To ensure the robustness of their results, they sampled parameters such as start day, contact rate from normal distributions, and simulated the epidemic 500 times. Their findings suggest that shortening the school week could serve as an important tool for controlling COVID-19 in schools.

In summary, addressing uncertainties in epidemiological models is crucial for ensuring robust predictions and informing policy decisions. Sensitivity and uncertainty analyses allow researchers to identify key parameters and assess the impact of various interventions while accounting for uncertainties in the data. Probabilistic sensitivity analyses offer a more comprehensive approach by considering the uncertainty of multiple parameters simultaneously. Ultimately, these analyses contribute to more effective and targeted intervention strategies in the face of infectious disease outbreaks, such as the COVID-19 pandemic.

2.4 Summary of Literature

In summary, great effort has been placed on research and policies about the modeling, forecasting, and containing infectious diseases. In the literature review above, we summarized the research about transmission modeling of infectious diseases and the optimization of vaccine allocation.

For the transmission modeling of infectious diseases, compartmental model, statistical model, and agent-based model are most widely adopted. At the early stage of a new epidemic, the relevant data usually suffer from a lack of monitoring and inconsistency between datasets from different surveillance systems. Considering the sufficiency and correctness of the epidemic data, it is challenging to build a complicated statistical model incorporating various aspects of transmission. Meanwhile, it is hard to use an agent-based model due to the tedious parameter tuning in practice. The compartmental model is still a good choice to capture the transmission dynamics at a high level, though it is limited by the homogeneity assumption. However, most of the deterministic compartmental models fail to capture the time dependency of the transmission dynamics. A concise and general form of the compartmental model with the consideration of time dependency and uncertainty plays a critical role in the modeling of future epidemics.

Regarding research on vaccine allocation, most of the previous studies have only focused on the epidemic that has not been widely spread, such as influenza. For an ongoing epidemic, such as COVID-19, prioritizing the vaccine allocation with the consideration of current and future infection has not been addressed too much. Meanwhile, the vaccine manufacturing and allocation of an ongoing epidemic is a dynamic process, which is different from allocating a fixed amount of vaccine in the previous studies. Dynamic optimization of vaccine allocation is still an urgent need for decision-makers.

Chapter 3. Research Design and Methods

3.1 Application of Model to Thesis Aims

In this section, we introduce our comprehensive modeling framework for transmission dynamics and vaccine allocation, which will be utilized in the thesis to accomplish the four research aims stated in Chapter 1. The methodology is illustrated in Figure 1.

Utilizing disease tracking data, such as case numbers, hospitalizations, fatalities, and recoveries, we first create a dynamic model incorporating time-varying transmission and fatality rates based on the widely recognized Susceptible-Exposed-Infectious-Recovered-Death (SEIRD) framework. By introducing these time-varying rates, we capture the evolving nature of the disease and the effects of various interventions, enabling a more comprehensive understanding of regional transmission dynamics and facilitating more accurate predictions of the epidemic's progression. We refine our model using advanced statistical and machine learning techniques for parameter estimation and incorporate real-time data updates to ensure responsiveness to changes in the disease landscape. This dynamic approach offers valuable insights into the effectiveness of interventions and their potential to alter the course of the epidemic within a specific region, as well as allowing for the investigation of potential correlations between different variables, such as the relationship between transmission rates and socioeconomic factors or the impact of pre-existing health conditions on fatality rates. Through this detailed examination of regional transmission dynamics, we can identify trends and patterns that inform public health policies and targeted interventions to curb the spread of the disease more effectively.

After developing a dynamic model for one region, we broaden our analysis to consider interactions between different regions and the impact of vaccinations. To accomplish this, we introduce a spatial interaction model and a vaccinated compartment to the existing SEIRD framework. The

spatial interaction model captures population movements between regions, which can contribute to disease spread. The vaccinated compartment represents the portion of the population that has been vaccinated, making them less susceptible to contracting and spreading the virus. By leveraging various data sources, such as mobility data and vaccination records, we can parameterize the spatial interaction model and the vaccinated compartment. This integrated approach enables us to analyze the effects of vaccination and regional interactions on transmission dynamics and generate more precise analysis of epidemic progression.

To assess the impact of intervention policies on transmission dynamics, we will conduct a multi-phase analysis for selected regions, based on local policy implementation timelines. This analysis will allow us to evaluate the effects of measures such as lockdowns, social distancing, and vaccination campaigns. Furthermore, we will examine the uncertainty of historical transmission under conditions of limited data and varying public responses to interventions. By employing a combination of sensitivity analysis and Monte Carlo simulations, we can quantify this uncertainty and determine its implications for our predictions and policy recommendations.

With an accurate representation of historical trends, we will optimize vaccine allocation, taking into account constraints on vaccine availability. We will devise mathematical models to identify optimal allocation strategies that maximize health outcomes, such as minimizing infections, hospitalizations, and fatalities. Subsequently, we will evaluate health outcomes under different policy scenarios, including varying levels of vaccine coverage, alternative prioritization strategies, and the effects of non-pharmaceutical interventions. This assessment will offer valuable insights for decision-makers in designing effective policies to control the epidemic.

To demonstrate the practical application of our methodology for pandemic prediction and control, we will apply our comprehensive modeling framework to the COVID-19 pandemic. By using real-

world data to calibrate our models and validate our predictions, we will showcase the relevance of our approach and its potential to contribute to improved pandemic preparedness and response in the future.

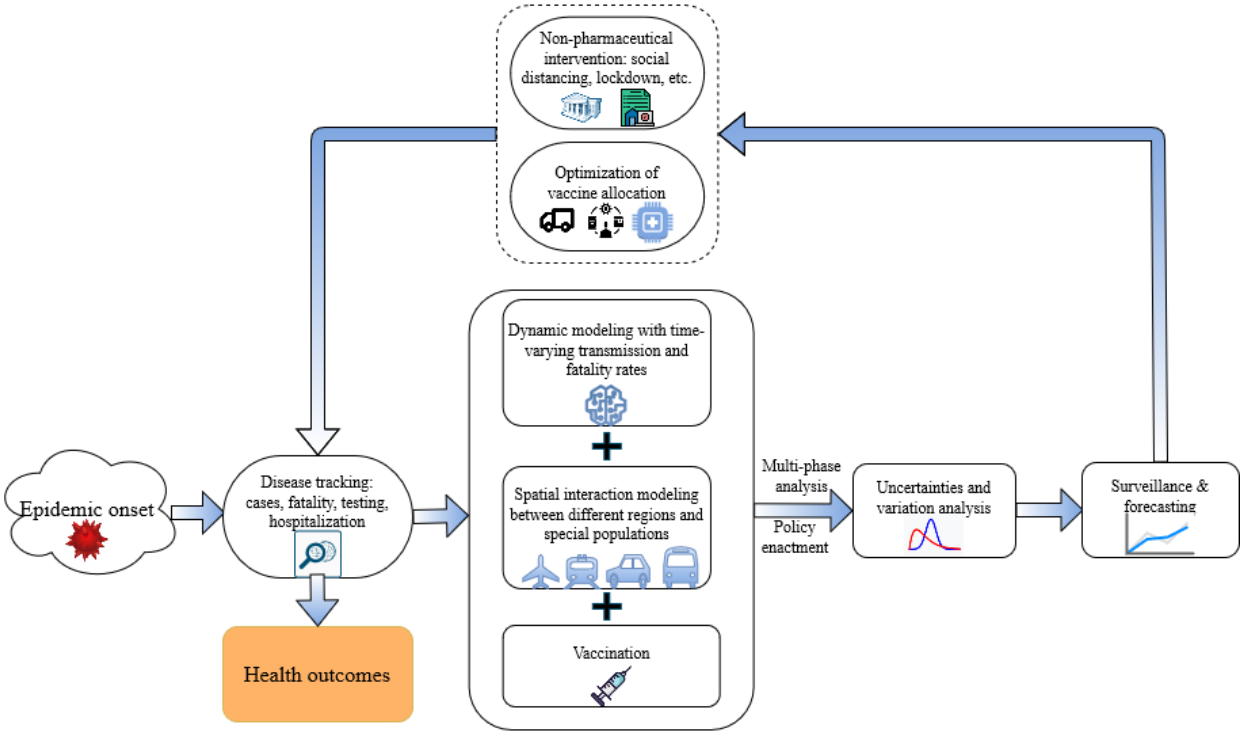

Figure 1: Research design for the pandemic prediction and control

3.2 COVID-19 Data Sources

In this section, we will introduce the datasets we have identified related to COVID-19 and systematically summarize the main categories of the data. We have identified four types of public data sets to support the research: (1) Disease tracking, including the daily infectious cases, hospitalization, fatalities and recoveries; (2) Travel and mobility data; (3) Hospitalization and utilization data; (4) Policy enactment data. All the data resources could be indexed by our website: www.covid19datasource.usc.edu. As we explained in the end of this section, not all data sources are in quality and we will focus on the disease tracking data, mobility data, and policy enactment data.

3.2.1 Disease Tracking (Cases, Hospitalization, Fatalities, Recoveries)

The basic data requirement for the COVID-19 related analysis is the disease tracking of the time-series data with respect to the daily case/fatality monitoring, PCR and antibody testing, hospitalization, vaccination and recovery. We could leverage the disease tracking data to analyze the transmission pattern, evaluate the effect of intervention, allocate medical resources, and control the immunization process. Table 1 below shows the most credible and wildly cited data sources for disease tracking.

Table 1: Data sources for COVID-19 disease tracking

Data source	Data information	Brief Description
The COVID Tracking Project [102]	Cases, testing, hospitalization, and patient outcomes, demographic, and long-term care facilities information	Current daily time-series data on cases, fatalities, tests, and hospitalizations.
Johns Hopkins University: COVID-19 Data Repository [103]	Cases, fatality, incidence rate, testing rate	Interactive dashboard with downloadable real-time data.
Our World in Data [104]	Cases, deaths, hospitalizations, testing, and vaccinations	Comprehensive COVID-19 pandemic data on cases, deaths, hospitalizations, tests, and vaccinations.
Facebook Data for Good[105]	Population density, movement range, forecasting, social connectedness index	Tools and initiatives for organizational COVID-19 response.
Google COVID-19 public dataset program [106]	Demographics, economy, epidemiology, geography, health, hospitalizations, mobility, government response, and weather	Global COVID-19 daily time-series datasets for US and EU countries.

3.2.2 Travel and Mobility

Recent research points out the significant and positive association between the frequency of air, automobile, rail and transit travel with the daily increase of infections [107], [108]. To better understand the mutual impact of the COVID-19 transmission and transportation related behavior, we identified several general public datasets. In Table 2 below, we summarize data sources from different aspects of transportation, including migration, airline, daily travel and change of mobility.

Table 2: Public travel and mobility data sources during the COVID-19 pandemic

Data source	Data type	Brief description
Airline On-Time Statistics[109]	Airline data	Departure/arrival statistics (including delays) by airport and airline, airborne time, cancellations, and diversions
Daily Travel during the COVID-19 Public Health Emergency	Airline, driving, rail and transit data	Daily percentage and count of people staying home/not staying home, trips taken across 10 distance groupings
FHV Trip Records[110]	Taxi	Trip records of For-Hire Vehicles (FHV) in New York City
Facebook movement map[111]	Mobility changes	Movement change comparisons, using a pre-social distancing baseline
Google Community Mobility Reports[112]	Mobility changes	Visits to places (e.g. grocery stores, parks) in each region, showing changes over time
Baidu China Migration Data[113]	Migration	Changes in population movement between Chinese cities, measured by Baidu Migration index

3.2.3 Hospital Utilization and Resource Availability

With the outbreak and resurgence of COVID-19, the medical system has been stressed by multiple waves of patients. Preparing, monitoring and allocating medical resources, such as ICU beds, ventilators, expendable supplies, and vaccines, affect the capability to response to the surge of patients. Table 3 below summarizes hospital resources, monitoring data and possible strategies for increasing relevant capacity.

Table 3: Hospital utilization and resource availability

Key Covid-19 Healthcare Resources			University of Southern California COVID-19 Data Source
Resource Category	Resource Types	Data Source	Strategies for Increasing Capacity
Personnel	Doctors -- certain types	STATISTA[114]	Implement overtime, postpone non-urgent care, extend staff availability, recruit temporary medical professionals
	Emergency Medical Technicians	NREMT	Implement overtime, recruit additional EMTs, enhance training programs
	Skilled Nursing	BUREAU OF LABOR STATISTICS	Hire traveling nurses, implement overtime, extend staff availability, postpone non-

			urgent care, recruit temporary nursing professionals
Medical Devices	Antibody Test	CV19 LAB TESTING DASHBOARD	Increase production, streamline distribution, prioritize high-risk populations
	Infusion Pumps		Optimize use, manage inventory efficiently, invest in additional equipment
	Oxygen Tanks		Optimize use, improve supply chain, prioritize critical cases
	PCR Tests	HEATHDATA.GOV	Enhance manufacturing and supply chain, prioritize testing for symptomatic or high-risk individuals, increase testing capacity
	Rapid Testing		Increase production, streamline distribution, expand testing sites
	Ventilators	DEFINITIVE HEALTH CARE[115]	Use alternative devices (CPAP, Noninvasive Ventilation), prioritize ventilator allocation, invest in additional equipment
Pharmaceuticals	Anticoagulation		Streamline distribution, prioritize high-risk patients, optimize treatment protocols
	Convalescent Plasma		Incentivize COVID-recovered patients to donate plasma, raise awareness, improve plasma collection infrastructure
	Dexamethasone	VIZIENTIC	Reserve for patients in critical need, establish stricter criteria for patient selection, increase production
	Monoclonal Antibody	HHS PROTECT PUBLIC DATA HUB[116]	Coordinate primary care for high-risk patients with underlying conditions, streamline distribution, enhance production
	Remedesivir		Diversify production sources, streamline distribution, prioritize high-risk patients
	Bulk Oxygen	GETUSPPE[117]	Optimize supply chain, prioritize allocation, invest in additional infrastructure
Expendable Supplies	Disinfecting wipes	GETUSPPE	Conservation, Substitution, Re-use, Limit HCP face-to-face interaction, Telemedicine, Proper use training, Velarization rooms, Physical barriers, Selective airborne infection room use, Cohosting patients
	Face Shields	GETUSPPE	Conservation, Substitution, Re-use, Limit HCP face-to-face interaction, Telemedicine, Proper use training, Velarization rooms, Physical barriers, Selective airborne infection room use, Cohosting patients
	Gloves	GETUSPPE	Conservation, Substitution, Re-use, Limit HCP face-to-face interaction, Telemedicine, Proper use training, Velarization rooms, Physical barriers,

			Selective airborne infection room use, Cohosting patients
	Gowns	GETUSPPE	
	Hand Sanitizer	GETUSPPE	
	N95 Masks	STATISTA	
	Surgical Masks	GETUSPPE	
Vaccination	Moderna	CDC M	Standardize vaccine distribution approaches, expand vaccination sites, increase production
	Pfizer-BioNTech	CDC P	Standardize vaccine distribution approaches, expand vaccination sites, increase production
Hospital Beds	ICU Beds	UNIVERSITY OF MINNESOTA[118]	Convert standard beds/PACU to ICU, deploy modular hospitals, reduce patient length of stay, focus on preventing severe disease
	Licensed Beds	COVID-19 CARE MAP[119]	Repurpose non-COVID areas (hotels, alternate spaces), utilize virtual care, deploy modular hospitals, reduce patient length of stay, focus on preventing severe disease
	Skilled Nursing, After Discharge	HEALTHDATA	Reduce post-discharge recovery stay length, prepare skilled nursing facilities for COVID patients, promote home-based recovery
Emergency Response	Ambulance	HOMELAND SECURITY	Reduce ER wait times, expand telemedicine use, prevent severe hospitalization by early intervention

3.3.4 Policy Enactment

In response to the COVID-19 pandemic, governments have implemented a wide range of intervention policies. The intervention policy, especially non-pharmaceutical interventions, plays a key role in flattening the infection number and reducing the stress on the healthcare system when the effective vaccines and medications are still under development.

During the COVID-19, government intervention policy has shown a large divergence in the response to the rapidly changing and unprecedented circumstances of COVID-19. The UK government and Swedish governments initially adopted a herd immunity approach, which implies doing little or nothing to stop the epidemic. [120]. The predictive model and the transmission results show such a strategy fails to flatten the curve and prevent overwhelming the healthcare

system. [121] Some countries applied stringent interventions, sometimes limiting residents' liberty. In China, for example, mandatory quarantine, social distancing, contact tracing and testing, have successfully mitigated the COVID-19 transmission.

In the United States, non-pharmaceutical intervention measures fall into five categories: Movement restrictions, Public health measures, Social and economic measures, Social distancing and Lockdown [122]. The subtypes of each category have been summarized below in Table 4.

Table 4: Intervention measures taxonomy

CATEGORY	MEASURES	BRIEF DESCRIPTION
Movement restrictions	Health/doc requirements	Health declarations or doctor's certifications required upon arrival to assess traveler's health and minimize the risk of spreading the virus.
	Border inspections	Travel and ID document checks at land and sea entry points to control the movement of people and curb the spread of the virus.
	Partial border closure	Borders closed to non-nationals/residents to minimize the influx of potentially infected individuals.
	Full border closure	Borders closed to everyone, including nationals, to prevent any cross-border spread of the virus.
	Internal checkpoints	Checkpoints within a country for health checks and controlling internal movement to limit the spread of the virus between regions.
	Flight suspensions	Government suspension of international/domestic flights to restrict travel and reduce the risk of importing new cases.
	Domestic movement limits	Limited movement within the country to reduce the virus's spread within communities.
	Visa limitations	Entry limitations for specific nationalities or new visa restrictions to control the flow of people from high-risk areas.
	Curfew implementation	Regional or country-wide curfews to limit the movement of people and reduce potential exposure to the virus.
	Movement monitoring	Electronic surveillance for case tracing or movement monitoring to ensure compliance with restrictions and track potential virus spread.
Public Health Measures	Public awareness campaigns	Media campaigns promoting hygiene, social distancing, and other preventive measures to raise public awareness and reduce the virus's spread.
	Quarantine/isolation policies	Self-quarantine/isolation for arrivals, symptomatic individuals, or contact cases to prevent the spread of the virus among the general population.
	Hygiene recommendations	Government-issued hygiene guidelines and precautions to educate the public on effective ways to protect themselves and others.
	Health screenings	Temperature controls and health screenings at airports/border crossings to identify potentially infected individuals before they enter the country.

Social economic measures	Non-COVID medical tests	Forced health checks unrelated to COVID-19 (e.g., HIV) to assess the overall health of the population and address other health concerns.
	Psychosocial support	Support for patients, families, and quarantined/locked down individuals to address mental health needs and alleviate stress related to the pandemic.
	Large-scale testing	Country or regional population screening to identify infected individuals, track the virus's spread, and inform public health measures.
	Public health system reinforcement	Hiring more medical personnel, expanding facilities, and increasing resources to better handle the pandemic and provide adequate healthcare.
	Infection testing	Tests to identify infected individuals, enabling targeted isolation and quarantine measures.
	Public protective gear	Mandatory masks/gloves, etc., to reduce the risk of transmitting the virus in public spaces.
	Additional health measures	Transport sanitation, additional health regulations to improve public health and reduce the potential spread of the virus.
	Burial regulation changes	Changes to burial regulations/attendance limits to minimize large gatherings and reduce the risk of virus transmission.
	Economic interventions	Measures to mitigate the economic and societal impact of the pandemic, such as financial aid, tax breaks, and subsidies.
	Emergency structures	Emergency Response committees for coordination, decision-making, and monitoring the implementation of pandemic response measures.
	Import/export restrictions	Restrictions on food/health item imports/exports to ensure adequate supply and distribution within the country.
	Emergency declaration	Declaration allowing the implementation of extraordinary measures to address the pandemic that may not be allowed under normal circumstances.
	Military assistance	Military deployment to support medical operations, enforce restrictions, and ensure compliance with public health measures.
Social distancing	Public gathering limits	Canceling public events and limiting gathering sizes to reduce close contact between individuals and minimize virus transmission.
	Business/service closures	Closure of businesses and public services, with online alternatives when possible, to reduce human interaction and curb virus spread.
	Prison policy changes	Changes to prison policies, such as early release, suspension of day-release programs, and visitation restrictions, to mitigate disease spread within prisons.
	School shutdowns	Authorities closing schools to limit close contact between students and staff, reducing potential virus transmission.

Lockdown	Partial lockdown	Limited reasons for leaving home (e.g., essential shopping, medical appointments); non-essential stores closed to minimize public movement and virus spread.
	Complete lockdown	Limited reasons for leaving home; non-essential services/production halted to enforce strict social distancing and minimize virus transmission.
	Camp/minority lockdown	Movement limitations for populations in camps or camp-like conditions, such as refugees or internally displaced persons, to reduce virus spread within vulnerable communities.

In our research, we will utilize the COVID Tracking Project [102] and Our World in Data [104] for the daily/cumulative confirmed cases, deaths, and demographic. For spatial interaction data between states, we will use the Daily Travel during the COVID-19 Public Health Emergency from BTS and daily percentage of out-of-state trips from Maryland Transportation Institute. For the vaccination data, we will use COVID-19 Vaccine Distribution Allocations by Jurisdiction from CDC, which provides the total distributed and administered vaccine of Janssen, Moderna, and Pfizer by Jurisdiction. For the analysis of policy effects, we will utilize the time series policy data from COVID-19 Government Response Tracker of Oxford University.

Chapter 4 Dynamic Modeling with Time-Varying Transmission and Fatality Rates

Dynamic modeling with time-varying transmission rates is a powerful tool in the field of epidemiology, enabling insights into complex disease outbreak dynamics. By accounting for fluctuations in transmission rates and fatality rates over time, these models offer a robust framework for anticipating the future trajectory of epidemics and informing evidence-based public health decision-making. Furthermore, dynamic modeling allows for the evaluation and comparison of various intervention strategies, such as vaccination campaigns, social distancing measures, and testing, tracing, and isolation protocols, enabling the optimization of resource allocation and the development of targeted, effective public health policies. This is particularly pertinent in the context of emerging infectious diseases, where initial information is often scarce, and the potential for rapid spread and severe consequences necessitates swift, informed action. By leveraging dynamic modeling, researchers and policymakers can better understand the interplay between time-varying transmission and fatality rates, assess the potential impacts of different interventions, and ultimately, help to mitigate the devastating effects of disease outbreaks on global health and wellbeing.

In this chapter, we present an enhanced SEIRD (Susceptible-Exposed-Infectious-Recovered-Death) model that incorporates time-varying case fatality and transmission rates for confirmed cases and deaths, aiming to provide a more comprehensive understanding of infectious disease dynamics. Our analysis demonstrates that, by representing case fatalities and transmission rates as simple Sigmoid functions, historical cases and fatalities can be accurately fit with a root-mean-squared-error accuracy on the order of 2% for the majority of American states during the period

from the initial cases up to July 20, 2020. For states experiencing multiple waves of infection, we propose an alternative multi-phase model, which allows for a nuanced understanding of the varying dynamics in these regions and the potential effects of intervention strategies throughout the epidemic [123]. The enhanced SEIRD model offers a valuable method for explaining historical reported cases and deaths using a compact set of parameters, thereby enabling the analysis of uncertainty and variations in disease progression across different regions. This approach provides crucial insights for public health officials and policymakers, supporting the development of targeted and effective strategies to control the spread of infectious diseases.

4.1 The Proposed Time-Varying Model

We draw from the SEIRD compartmental model, which divides the population into five groups: susceptible(S), exposed (E), infected (I), recovered (R) and dead (D). The SEIRD model is selected due to its simplicity and flexibility. The model can be easily adapted to capture the unique characteristics of different diseases and populations without requiring detailed individual level data, which is usually not available at beginning of an epidemic.

SEIRD utilizes differential equations to model the evolution of the number of people in these states over time. Initially, individuals are classified as susceptible, meaning they have not been infected with the disease and are at risk of becoming infected. When an individual comes into contact with an infected person, they may become exposed, meaning that the disease is in its incubation period and the individual is not yet infectious. After the incubation period, the individual becomes infectious and is classified as infected. Finally, if the individual recovers from the disease, they are classified as recovered and become immune to the disease.

The transmission of the disease is governed by the transmission rate $\beta(t)$, which represents the probability of an infected individual transmitting the disease to a susceptible individual. The

transmission rate is typically assumed to be constant in the basic SEIR model, but it is modeled as a time-varying parameter to capture changes in behavior or the impact of interventions on disease transmission. Death rate $\alpha(t)$ is also treated as a time varying function, representing the proportion of infectious individuals who eventually die from the disease, by date. Those who eventually die transfer from the infected to the died state at a rate of ρ, representing the inverse of the time from becoming infectious until time of death. In our model, ρ is assumed to be constant over time. Those who eventually recover do so at the γ, representing the inverse of the time from becoming infectious until recovery. We will also later derive the effective reproduction number $Rep(t)$, representing the average number of persons who are exposed to the disease by each infectious person, as a function of time.

Taking these factors into account, the system of equations of the proposed SEIRD model is given by Equation (6):

$$\begin{aligned}
\frac{\partial S(t)}{\partial t} &= -\beta(t) \cdot I(t) \cdot \frac{S(t)}{N} \\
\frac{\partial E(t)}{\partial t} &= \beta(t) \cdot I(t) \cdot \frac{S(t)}{N} - \sigma \cdot E(t) \\
\frac{\partial I(t)}{\partial t} &= \sigma \cdot E(t) - \big(1 - \alpha(t)\big) \cdot \gamma I(t) - \alpha(t) \cdot \rho \cdot I(t) \\
\frac{\partial R(t)}{\partial t} &= \big(1 - \alpha(t)\big) \cdot \gamma \cdot I(t) \\
\frac{\partial D(t)}{\partial t} &= \alpha(t) \cdot \rho \cdot I(t)
\end{aligned} \tag{6}$$

where:

$S(t) =$ number of people in susceptible state at time t

$E(t) =$ number of people in exposed, but uninfected at time t

$I(t) =$ number of people in infectious state at time t

$D(t) =$number of people who have died at time t

$R(t) =$number of people who have recovered at time t

N = total number of people

$\beta(t) =$ transmission rate at time t

$\sigma =$ transformation rate from exposed to infectious, which is the reciprocal of the incubation period

$\alpha(t) =$ likelihood of eventual death of a person who is infected at time t

$\gamma =$ transformation rate from infectious to recovered, which is the reciprocal of the recovery time

$\rho =$ transformation rate from infectious to death

Transmission rates in epidemics are influenced by several factors, including pathogen characteristics, population density, mobility, and social behaviors. The virulence and transmissibility of a pathogen directly impact its spread, with more infectious pathogens resulting in higher transmission rates. High population densities, particularly in urban centers, and increased movement of people due to travel and globalization facilitate transmission by enabling more frequent contact between individuals. Social behaviors, such as handshaking or attending large gatherings, can increase the likelihood of transmission, while cultural practices, like funeral rituals, can also contribute to disease spread if they involve close contact with infected individuals or their bodily fluids. Public awareness and adherence to hygiene practices, such as handwashing and sanitizing, can further influence transmission rates.

Death rates during epidemics can be affected by factors such as age distribution, healthcare infrastructure, and the prevalence of comorbidities. Older individuals and those with underlying health conditions, like diabetes or heart disease, are more susceptible to severe outcomes, leading to higher death rates in populations with a greater proportion of vulnerable individuals. The capacity and quality of healthcare systems also play a crucial role in determining death rates.

Access to timely diagnosis, adequate hospital capacity, and effective treatments can mitigate the severity of the disease, reducing the likelihood of fatalities. Furthermore, socioeconomic factors, such as access to healthcare, nutrition, and living conditions, can influence the overall health and resilience of a population, affecting death rates during an epidemic.

Intervention policies, such as non-pharmaceutical interventions (NPIs), vaccination programs, and testing, tracing, and isolation strategies, can have a significant impact on transmission and death rates during epidemics. NPIs like social distancing, mask-wearing, and school or workplace closures help reduce transmission by limiting contact between individuals, slowing the spread of the disease and preventing healthcare systems from becoming overwhelmed. Vaccination programs are vital for controlling infectious diseases, as they lower the number of susceptible individuals in a population, indirectly reducing death rates by protecting vulnerable individuals from infection. Effective testing and tracing programs can identify and isolate infected individuals, further limiting the spread of the disease, providing critical information for public health decision-making, and potentially saving lives. Public health communication and community engagement are also essential for the successful implementation of these intervention policies, as they help to build trust and ensure adherence to guidelines.

Recognizing the importance of these intervention policies, it is crucial to understand how changes in such policies, global events, and medical care affect transmission dynamics, represented by $\alpha(t)$ and $\beta(t)$. While these functions could potentially change erratically due to discrete events such as the introduction of new public health measures, we hypothesize that such events do not cause abrupt alterations in either function. Therefore, we explore whether a simple continuous model, with a minimal set of parameters, can accurately represent historical data. For instance, when a new intervention policy is enacted, the public may not react immediately, and the transmission

parameters do not shift instantaneously. Over time, the public adapts to the policy, and the effective reproduction number eventually stabilizes. Furthermore, the public's response is influenced not only by government policies but also by effective communication about the disease. Communication comes from many, sometimes conflicting, sources. How the public at large absorbs and responds to such often confusing messages may be gradual. The public will get used to the policy after a period of adaptation, and eventually the effective reproduction number will stabilize.

A natural function to describe this pattern of change is the Sigmoid function. Equation (7) is the general form of the Sigmoid function, where k determines the slope of the function and a determines the x value at the middle point (i.e., point of time when y=.5).

$$S(x) = \frac{1}{1 + e^{k(x-a)}} \tag{7}$$

Thus, we define the function for transmission rate and death rate as Equation (8) and (9).

$$\beta(t) = \beta_{end} + \frac{\beta_{start} - \beta_{end}}{1 + e^{m\cdot(x-a)}} \tag{8}$$

$$\alpha(t) = \alpha_{end} + \frac{\alpha_{start} - \alpha_{end}}{1 + e^{n\cdot(t-b)}} \tag{9}$$

where,

β_{start} is the starting reproduction number

β_{end} is the ending reproduction number

α_{start} is the starting death rate, ranging from 0 to 1

α_{end} is the ending death rate, ranging from 0 to 1

m, n, a, b are the shape parameters

4.2 Parameter Estimation and Model Fitting

Parameters in Eqs. 1 will be estimated with the objective of minimizing the weighted summation of squared error between cumulative predicted and measured confirmed cases and the summation

of squared error between cumulative predicted and cumulative confirmed deaths. Our analysis encompasses the period from the day of the first reported case in each state until July 28, 2020, covering all 50 American states. For each state, we selected a start date four days prior to the date of the first confirmed case, in accordance with a report by the Centers for Disease Control and Prevention (CDC), which indicates that the median incubation period is 4 days, with a range of 2~7 days.

To estimate the shape parameters m, n, a, b and the starting/ending parameters $\beta_{start}, \beta_{end}, \alpha_{start}, \alpha_{end}$, we fit Eqs. 1 to the cumulative confirmed case numbers and the cumulative confirmed death numbers with the nonlinear least square method. Other parameters were derived from prior research.

Among 305 hospitalized patients and 10,647 recorded deaths, the median time of hospitalization was 8.5 days and the median interval from illness onset to death was 10 days (IQR = 6 - 15 days). We assume the median hospitalization time is the median time for infectious people to stop being contagious. Hence, we set these parameters as the inverse of these time values: $\sigma = ¼, \gamma = 1/8.5,\ \rho = 1/10.$

The remaining parameters are derived for each American state by optimizing the fit of the model to historical case and death data, where the objective is to minimize a weighted sum of daily squared error over the analysis period. We utilized a search algorithm that required initialization and a constrained search space, as explained below.

We define the model function M(t; $[\beta_{start}, \beta_{end}, m, a, \alpha_{start}, \alpha_{end}, n, b]$): $t \rightarrow R^2$, where M(t; $[\beta_{start}, \beta_{end}, m, a, \alpha_{start}, \alpha_{end}, n, b]$) = $\left[\hat{I}(t) + \hat{R}(t) + \hat{D}(t), \hat{D}(t)\right]$ and the reported case number and death number at time t is [Cases(t), Deaths(t)].

Because it is unlikely for transmission and death rates to change drastically in a single day, we set upper bounds for $m\ and\ n$ at 0.33 (meaning that rates do not suddenly change in less than three days) and initialize the search at 0.25. We permit the turning point of the sigmoid function to occur on any day in the timeline; we set $a, b \in [0,125]$, where 125 is the length of the period from March 1st to July 28th, in days (as of March 1 few states had reported cases). Prior research suggests that the initial effective reproduction number is around 3 [124], equivalent to a transmission rate of 0.75, which we use for initialization. Because transmission rates vary significantly among locations due to local conditions (such as crowding), we bound $\beta_{start} \in [0.5,7.5]$ and $\beta_{end} \in [0,2.5]$, thus permitting a wide range of results.

To summarize, the parameters set $P = [\beta_{start}, \beta_{end}, m, a, \alpha_{start}, \alpha_{end}, n, b]$ is initialized as $[0.75,0.5,0.25,10,0.4,0.1,0.25,10]$. Then the parameter optimization problem is formulated in Equation (10).

$$\begin{aligned} & min_P \|M(\mathrm{t}; P) - [Cases(t), Deaths(t)]\|_2^2 \\ s.t. \quad & 0.5 \le \beta_{start} \le 7.5 \\ & 0.1 \le \beta_{end} \le 2.5 \\ & 0 \le \alpha_{\mathrm{start}} \le 1 \\ & 0 \le \alpha_{end} \le 1 \\ & 0.01 \le m \le 0.33 \\ & 0.01 \le n \le 0.33 \\ & 0 \le a \le 125 \\ & 0 \le b \le 125 \end{aligned} \tag{10}$$

The number of reported deaths is smaller than the number of reported cases in all locations. Thus, treating errors in death estimation and case estimation the same will lead to underfitting of the death data, in preference to minimizing the errors in case data. Therefore, considering the accuracy of the reported death data and the fitting accuracy, we optimized a weighted sum of squared death and case data, multiplying w by deaths during the fitting process. The adjusted objective function is formulated in Equation (11):

$$\min_p \left\| \left(\hat{I}(t) + \hat{R}(t) + \hat{D}(t) - Cases(t) \right)^2 + w * \left(\hat{D}(t) - Deaths(t) \right)^2 \right\|_2 \tag{11}$$

The parameters are estimated by solving the nonlinear constrained least-squares problem in Equation (10), utilizing the Levenberg–Marquardt algorithm (LMA). The Levenberg-Marquardt Algorithm (LMA) is a numerical optimization algorithm used to solve non-linear least squares problems by combining the steepest descent method and the Gauss-Newton method. It starts by using the steepest descent method to make large corrections in the model parameters, then switches to the Gauss-Newton method to make more accurate adjustments as the parameters get closer to the optimal solution. The LMA also uses a damping factor to balance the step size between the steepest descent and Gauss-Newton methods, ensuring convergence to the minimum of the objective function. Its hybrid approach and damping factor make it robust and efficient, and it has many practical applications in science and engineering[125]. The LMA will be implemented to our model fitting by the *lmfit* package in Python.

4.3 Model Accuracy

The first case of COVID-19 in the United States was reported on January 20, 2020[126]. As of July 31, 2020, a total of 4,665,469 cases and 155,863 deaths had been reported across the states and territories of America[102]. We fit the model with the dataset of 7-day moving average cases and deaths for the 50 states, provided by the COVID-19 tracking project lead by *The Atlantic* (derived from the Center for Disease Control), for the period from the date of the first reported cases to July 31st. The fitting accuracy across all states is presented in Figure 2, measured by the relative root mean square error (RRMSE) in Equation (12).

$$RRMSE = \frac{\left[\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 / N\right]^{1/2}}{y_N - y_1} \tag{12}$$

where y_N is the case/death on the N_{th} day.

The fitting accuracy of the reported cases ranges from 0.54% to 7.34% and of the reported deaths ranges from 0.29% to 7.28%. The average and median RRMSEs for deaths are 1.61% and 1.33%;

for cases, the average and median values are 2.30% and 1.88%. RRMSE fell below 5% by both measures for all states except Hawaii, Idaho, Louisiana, Montana and Wyoming.

Figure 2 shows that the proposed SEIRD model with time-dependent transmission rate and death rate captured the pattern of the transmission dynamics well across most states with only 8 fitted parameters.

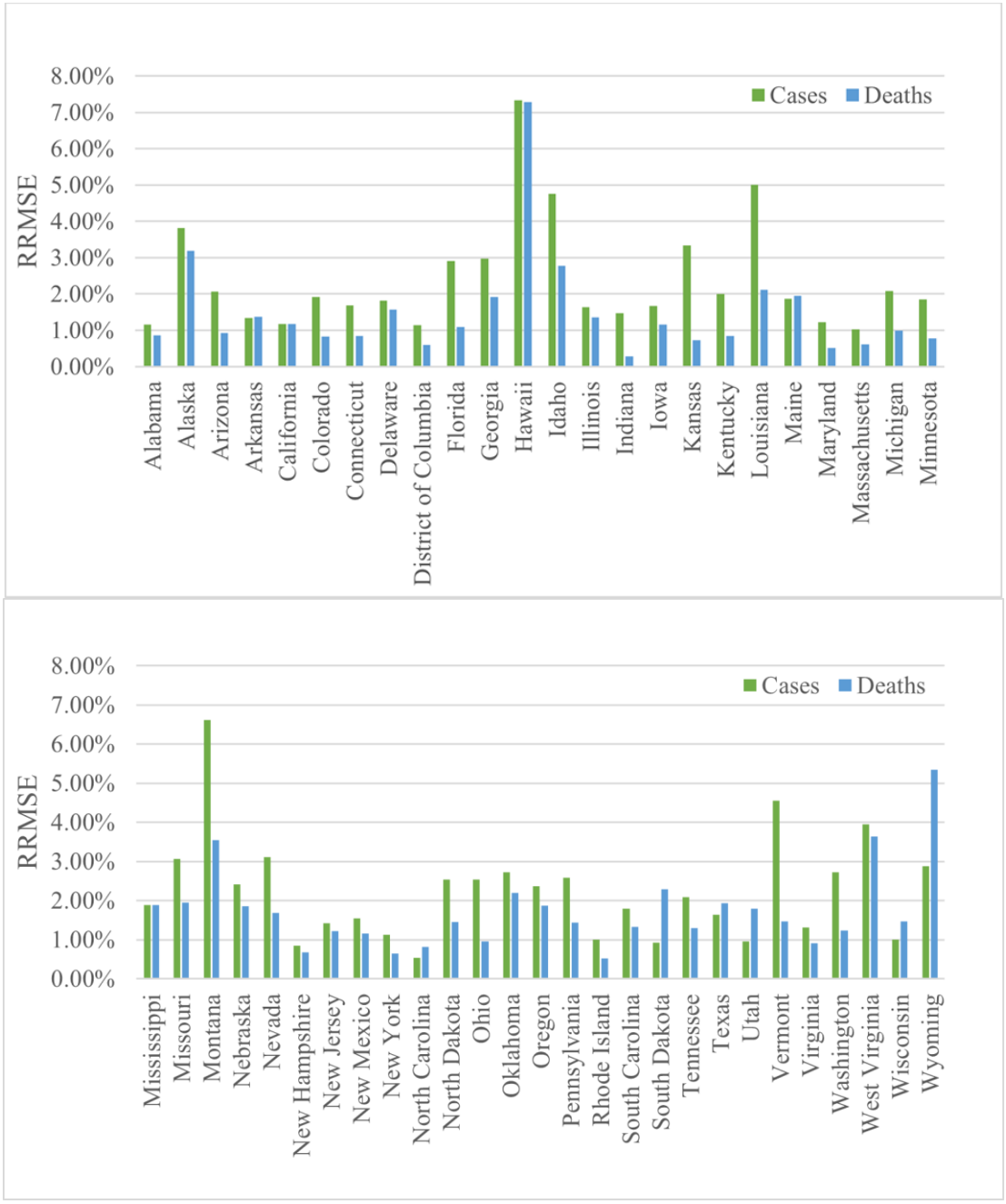

Figure 2: Fitting accuracy of the cases and fatality across all states

Figure 3 and Figure 4 show the specific fitting results for cases and deaths by day for the two states with the largest number of cases (New York and California) as well as two other states for which the fit is less accurate (Florida and Hawaii). For New York and California, the fitting results almost coincide with the CDC data. Examining Florida and Hawaii, the CDC data follows a pattern of two phases, which is not as well captured by our model. Especially for Hawaii, the curve flattened for a period and then rose. As discussed later, our model characterizes the transmission dynamic for a period with one phase, i.e. the curve should become flat at most once.

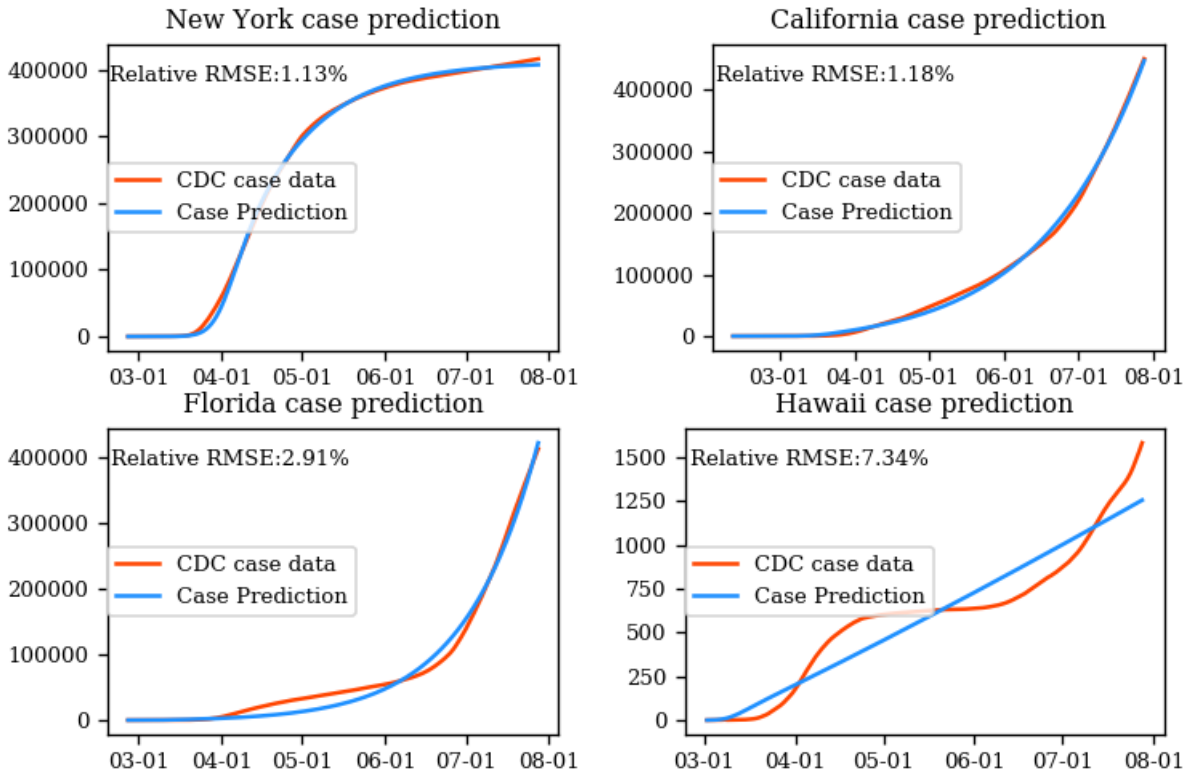

Figure 3: Fitting results of case number for New York, California, Florida, and Hawaii

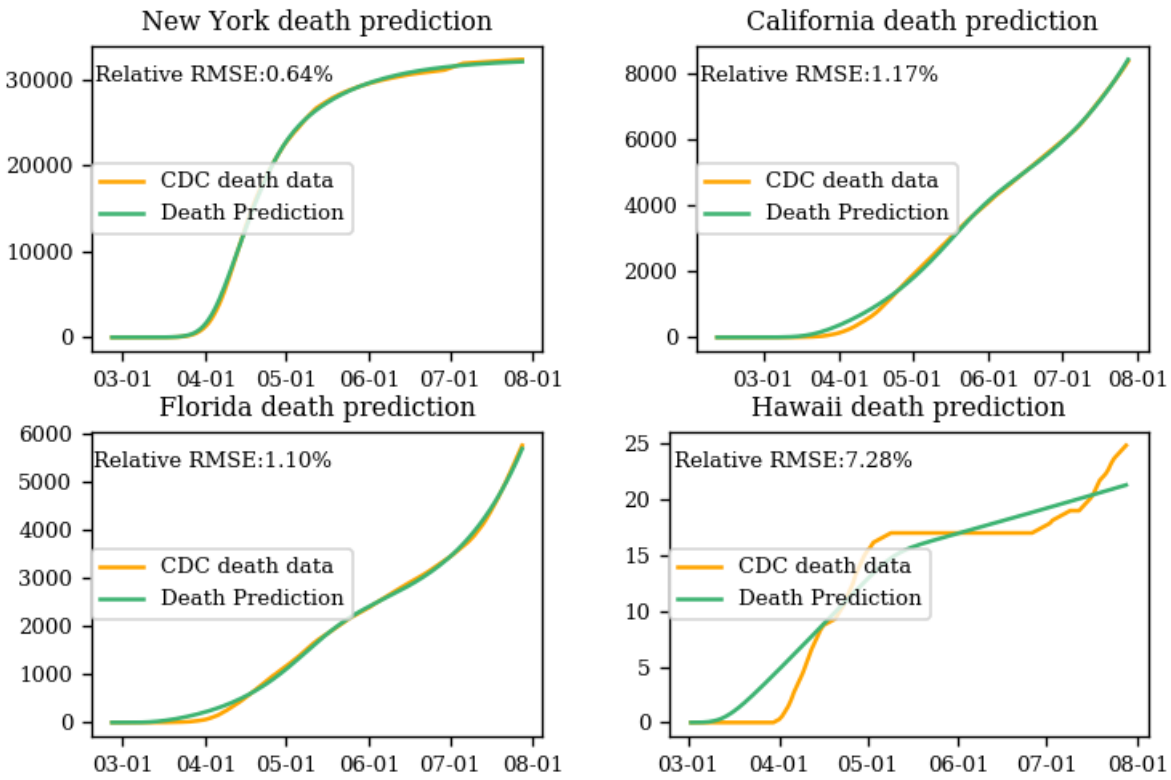

Figure 4: Fitting results of death number for New York, California, Florida, and Hawaii

4.4 Effective Reproduction Number Calculation and Trends

The effective reproduction number, which we define as $Rep(t)$, is a key metric used in epidemiology to describe the transmission potential of an infectious disease. It represents the number of secondary infections that can be caused by a single infected individual in a population that is partially susceptible to the disease. In other words, $Rep(t)$ is a measure of how many people an infected person will go on to infect, on average [127]. When $Rep(t) > 1$, the rate of new cases will increase over time, until the population loses susceptibility to the disease. When $Rep(t) < 1$, the rate of new cases will decline over time.

There are several factors that can influence $Rep(t)$, including the infectiousness of the disease, the duration of infectiousness, the population density, and the effectiveness of control measures such as social distancing, mask-wearing, and vaccination. Public health officials use $Rep(t)$ as a tool for monitoring the progress of an outbreak and for making decisions about when and how to

implement control measures. Hence, during this section we will calculate the trend of effective reproduction number and analyze the possible factors that influence the evolution of the transmission dynamics.

$Rep(t)$ can be estimated with the Next Generation Matrix(NGM) method[128]. The Next Generation Matrix method is a powerful tool for estimating $Rep(t)$ from compartmental models. The NGM method is based on the idea that the distribution of secondary infections can be described by a matrix, with each element representing the probability of an infected individual transmitting the infection to another individual in a specific population subgroup. The NGM matrix can be calculated from the parameters of the compartmental model, such as the transmission rate and the distribution of individuals in different compartments. The resulting matrix is then used to calculate the spectral radius, which is a measure of the effective reproduction number $Rep(t)$.

We define X as the vector of infected class (i.e. E, I) and Y as the vector of uninfected class (i.e. S, R, D). Let $\frac{dX}{dt} = \mathcal{F}(X,Y) - \mathcal{V}(X,Y)$, where $\mathcal{F}(X,Y)$ is the vector of new infection rates (flows from Y to X) and $\mathcal{V}(X,Y)$ is the vector of all other rates. Then for our model, the next generation matrix is expressed in Equation (13).

$$\begin{aligned} M &= \left(\frac{\partial \mathcal{F}}{\partial X}\right)_{(N,0,0,0,0)} \left(\frac{\partial V}{\partial X}\right)^{-1}_{(N,0,0,0,0)} \\ &= \begin{bmatrix} 0 & \beta(t) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma & 0 \\ \sigma & \left(1-\alpha(t)\right)\cdot\gamma + \alpha(t)\cdot\rho \end{bmatrix}^{-1} \end{aligned} \tag{13}$$

Then the effective reproduction number is the spectral radius of M, which is $\frac{\beta(t)}{(1-\alpha(t))\cdot\gamma+\alpha(t)\cdot\rho}$

At the beginning of the epidemic, $Rep(t)$ reflects the natural transmissibility of COVID-19, i.e. the basic reproduction number R_0 in the absence of intervention. With the evolution of the epidemic, $Rep(t)$ changes dynamically, as do the transmission rate $\beta(t)$ and death rate $\alpha(t)$,

which are influenced by both the intervention policy and population immunity. Figure 5 and Figure 6 show the fitted $Rep(t)$ at the start of the epidemic across all states and fitted $Rep(t)$ on July 31st. We see that $Rep(t)$ ranges from 1.27 to 16.49, with a median value of 2.87. It should be kept in mind that this optimal fit is a reflection of the reported data on cases. Increasingly aggressive testing may make it appear that $Rep(t)$ grows faster than the actual (unknown) number of cases.

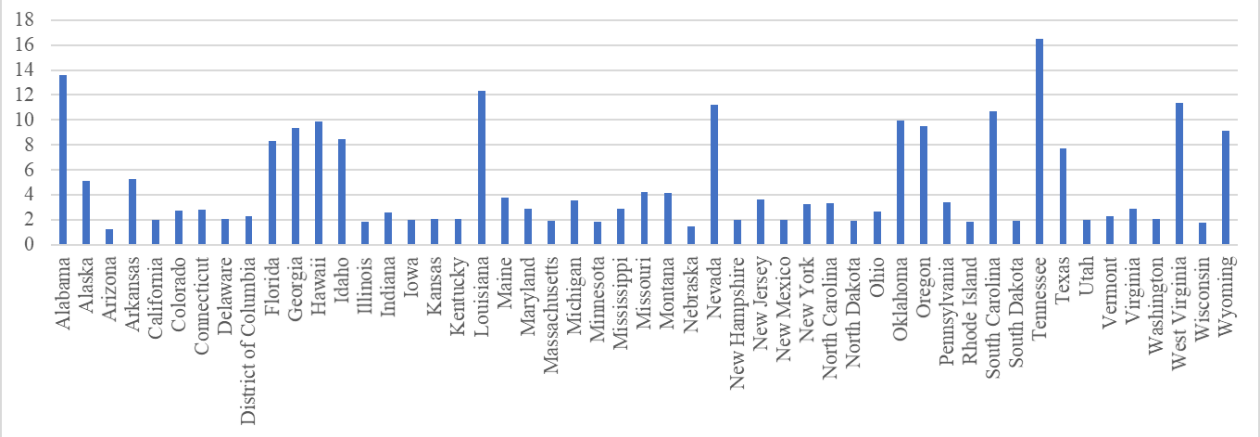

Figure 5: Fitted Rep(t) at the start of the epidemic across all states

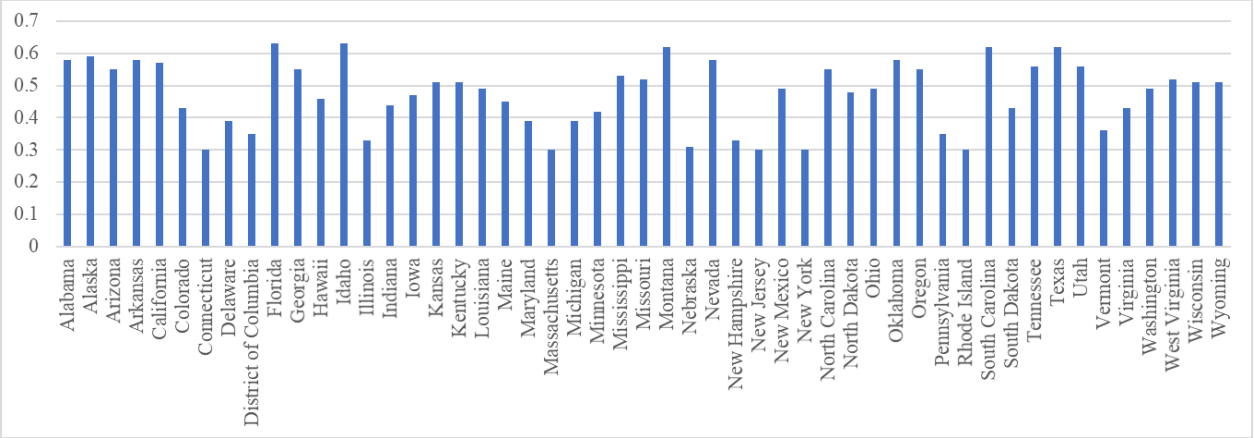

Figure 6: Fitted *Rep(t)* on July 28th

For illustration, Figure 7 shows our estimated history of for New York, California, Florida and Hawaii. Time 0 in these graphs is the day of the first reported case, which varies from state to state. In these cases, the effective reproduction number both stabilized and became smaller than 1 with time, with the change occurring over a period of 10 to 30 days.

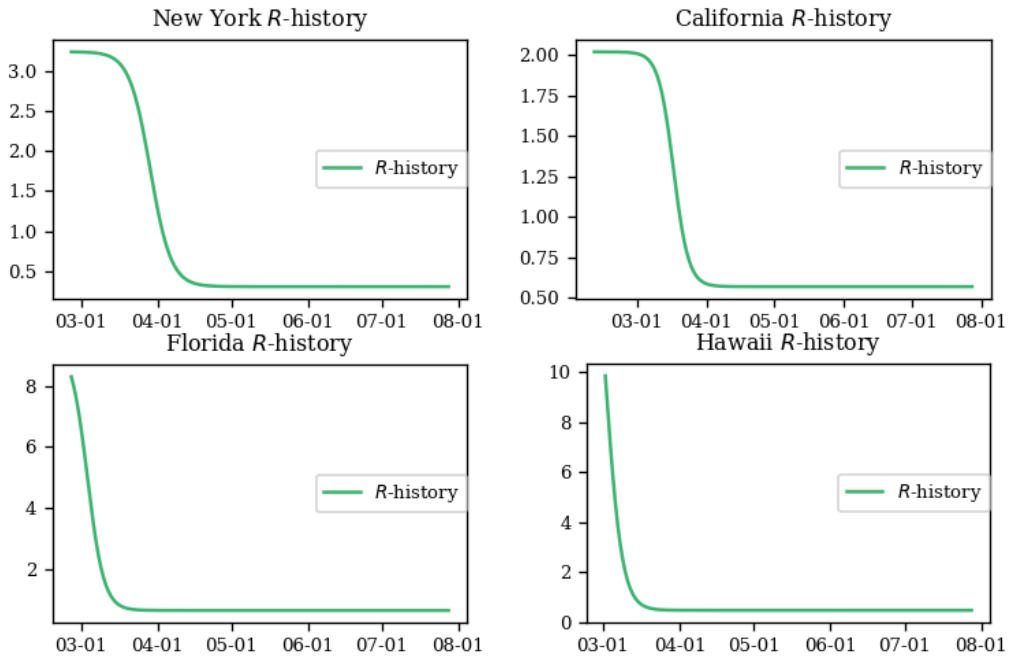

Figure 7: History of effective reproduction number for New York, California, Florida and Hawaii

As noted, in the early stages of an epidemic, the reproduction number may seem particularly large not only because the disease spreads rapidly but also because the rate of testing is increasing. In this sense, the estimated reproduction number is a reflection of both changes in the data collection process and the actual spread of disease.

Death rate is another measure that shows the change in virus outcomes over time, reflecting the health system's ability to deal with the flood of infected people. Figure 8 provides examples. From the historical plot, we see the hardest-hit states, like New York and Florida, experienced a much higher death rate in the early stage than the average 3% death rate in the United States. The relatively high death rate could be caused by the lack of effective medical treatment and hospital overload. It could also reflect limited testing of patients, whereby only the sickest patients were recorded as cases. With improvement of medical treatment, and increased testing, the death rate per confirmed case for most states decreased to a much smaller value.

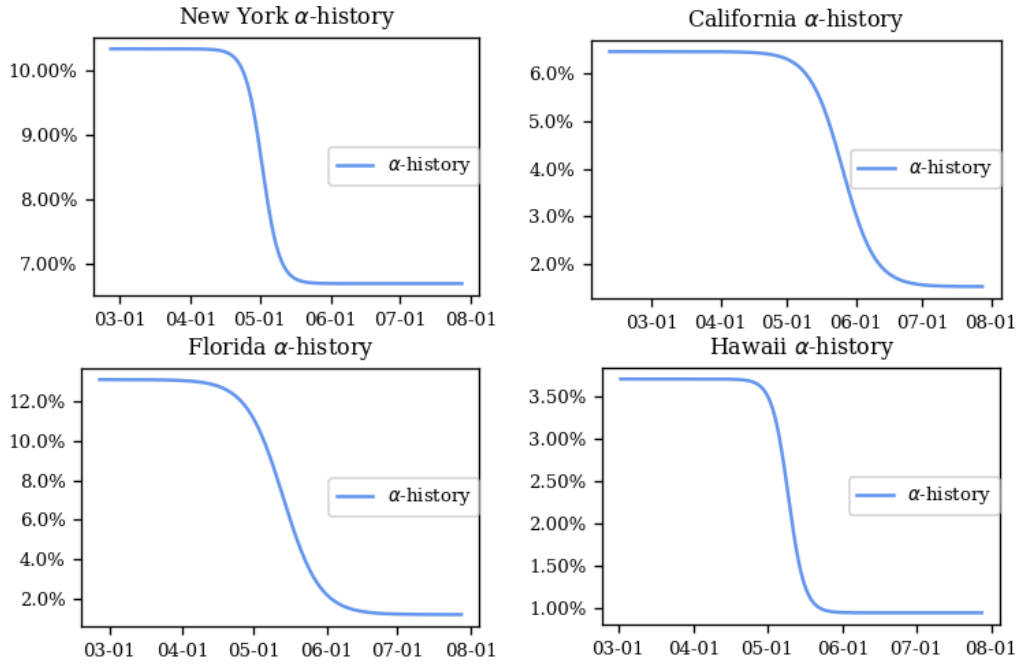

Figure 8: History of death rate for New York, California, Florida and Hawaii

4.5 Multi-Phase Model

Our model, as initially presented, demonstrates a strong fit for reported data on cases and deaths, with an error margin of less than 2% in the majority of states. Nevertheless, the model's foundational assumption—that transmission rates do not initially decrease before eventually increasing—necessitates modification when applied to states that have experienced multiple waves of the disease. Data from Hawaii, where our model exhibits the least accurate fit, exemplifies this pattern.

In response to these limitations, we introduce a multi-phase model specifically designed for locations exhibiting multiple waves of the disease. The first positive COVID-19 case in Hawaii was announced on Oahu on March 6th, prompting the Hawaii Department of Health to implement a stay-at-home order on March 25th. As a result, the case curve flattened between April 19th and May 7th. Upon announcing the commencement of the first phase of reopening on May 7th, data began to reflect a second wave of the virus.

To account for this pattern, we divided the Hawaii timeline into two distinct periods: the first from March 6th to May 7th, and the second from May 7th to July 28th. For the first phase, we fit the model with the assumption of only one exposed individual at the beginning. To initialize the second phase, we used the predicted numbers of exposed, infectious, and recovered individuals from the first phase, incorporating the reported deaths as of May 7th. This modification resulted in a decrease in the RRMSE for cases to below 2.5% and the RRMSE for deaths to below 2.7%. As illustrated in Figure 9, our two-phase model more accurately captures the transmission pattern in Hawaii than the single-phase model. This approach allows us to better account for fluctuations in transmission rates, reflecting the complexities of disease spread in locations with multiple waves of infection. Further research and refinements to our multi-phase model may enable even more precise predictions and facilitate better-informed policy decisions for managing the ongoing pandemic.

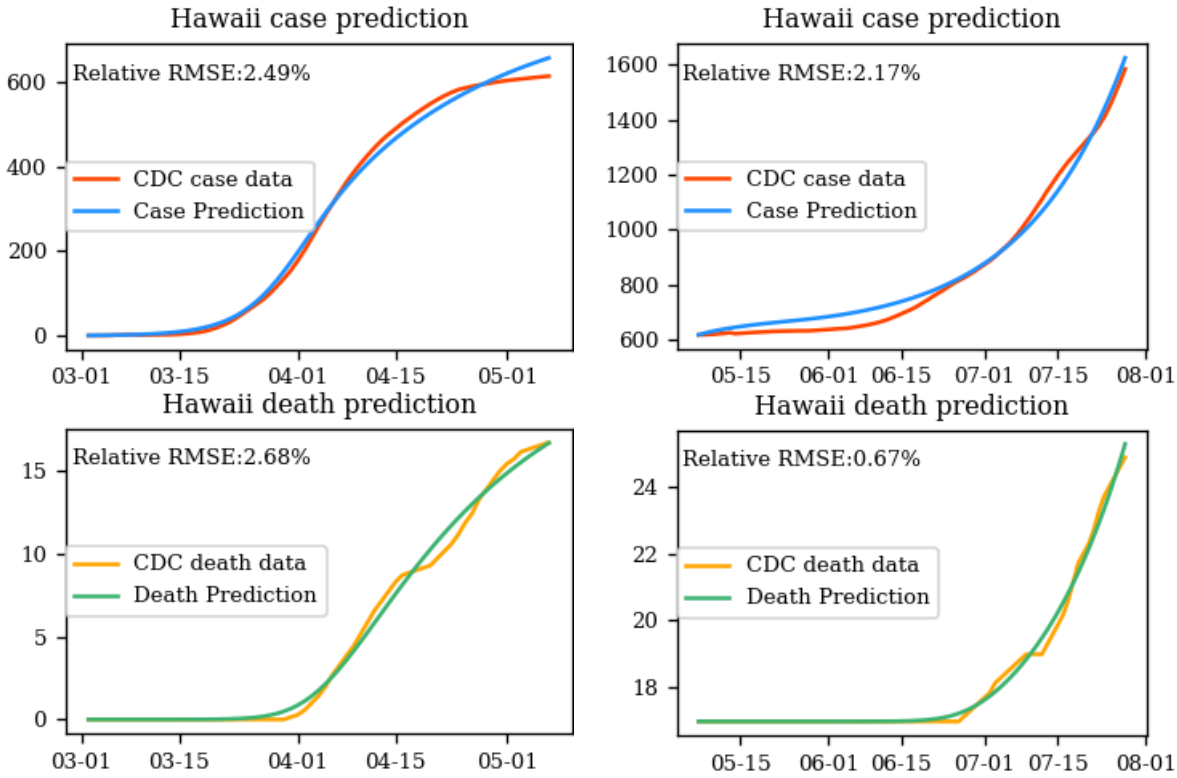

Figure 9: Fitting results for the two phases Hawaii

The history of effective reproduction number and death rate are shown in Figure 10. The first phase showed a decline in the reproduction number after the initial announcement of the stay-at-home order. However, with the reopening, the reproduction number increased, explaining increases in case rates. Death rates, by contrast, exhibit a peculiar behavior, increasing over time in each phase, with a discontinuity when transitioning from the first phase to the second. Beyond exhibiting two phases, Hawaii has a small number of deaths, with no deaths occurring in the transition period between phases. We surmise that the function, while representing the data well, is peculiar because of the unusual pattern in deaths within Hawaii.

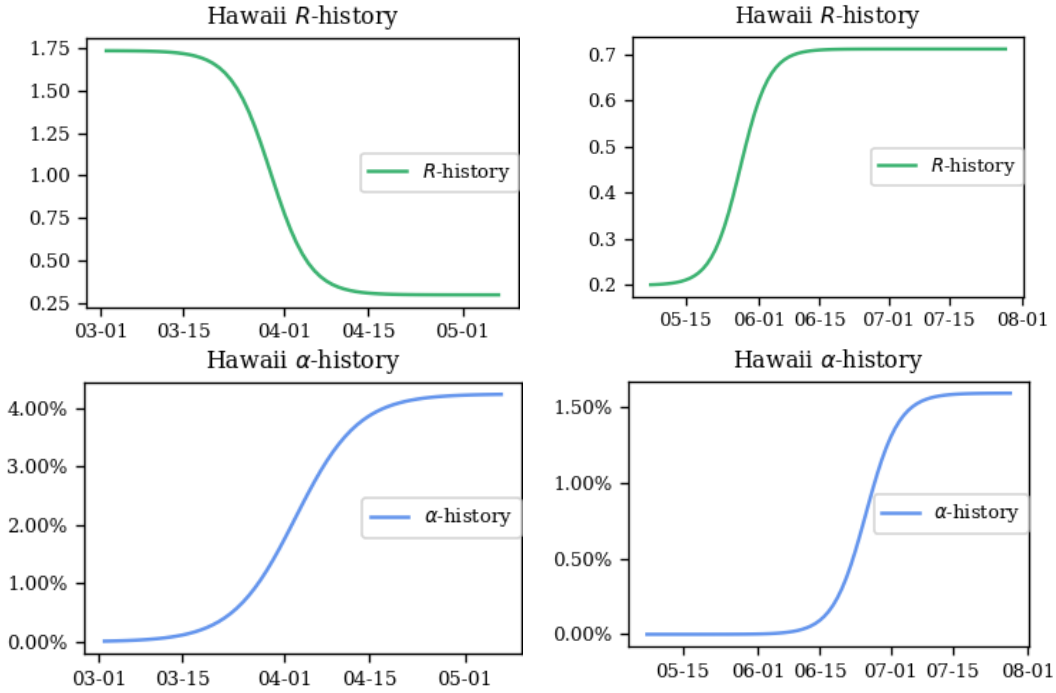

Figure 10: Historical results of the effective reproduction number and death rate

4.6 Sensitivity Analysis for Basic Time-varying Model

Sensitivity analysis is a critical step in transmission modeling of pandemics. Through sensitivity analysis, researchers and policymakers can determine the impact of changes in input parameters or assumptions on the model output. By varying input parameters and assessing the resulting output, sensitivity analysis can identify which parameters or assumptions have the most significant

impact on the model's results, such as the number of cases or deaths predicted. This process helps researchers and policymakers understand how changes in key parameters can affect the spread of the virus and the effectiveness of various interventions.

One common method for conducting sensitivity analysis is one-way sensitivity analysis. This method involves varying one input parameter while keeping all other parameters constant to evaluate its effect on the model output. For example, one-way sensitivity analysis can be used to assess how changes in the rate of transmission affect the number of cases or deaths predicted by the model. In our model, we have eight parameters with four parameters (i.e. $\beta_{start}, \beta_{end}, m, a$) related to case and four parameters (i.e. $\alpha_{start}, \alpha_{end}, n, b$) related to death. In order to check the significance of each parameter to the case and death, we vary each parameter by $\pm 5\%$ while keep other parameters constant as the best-fitted values. We sample each parameter 500 times within this ±5% range, and compute the maximum discrepancy in case/death numbers on the 30th day since the onset. This discrepancy is represented as $\Delta_{para} = O_{max} - O_{min}$, where O_{max} and O_{min} denote the highest and lowest case/death numbers among the 500 samples on the 30th day, respectively. For a comparative analysis of significance across different parameters, we normalize each parameter's maximum discrepancy (Δ_{para}) among the eight parameters. This normalized significance is denoted as $\alpha_{para} = \frac{\Delta_{para}}{\Sigma \Delta_{para}}$. Figure 11 and Figure 12 illustrate the significance of the transmission parameter on case and death numbers. Here, the height of each color bar signifies the level of significance associated with the corresponding parameter.

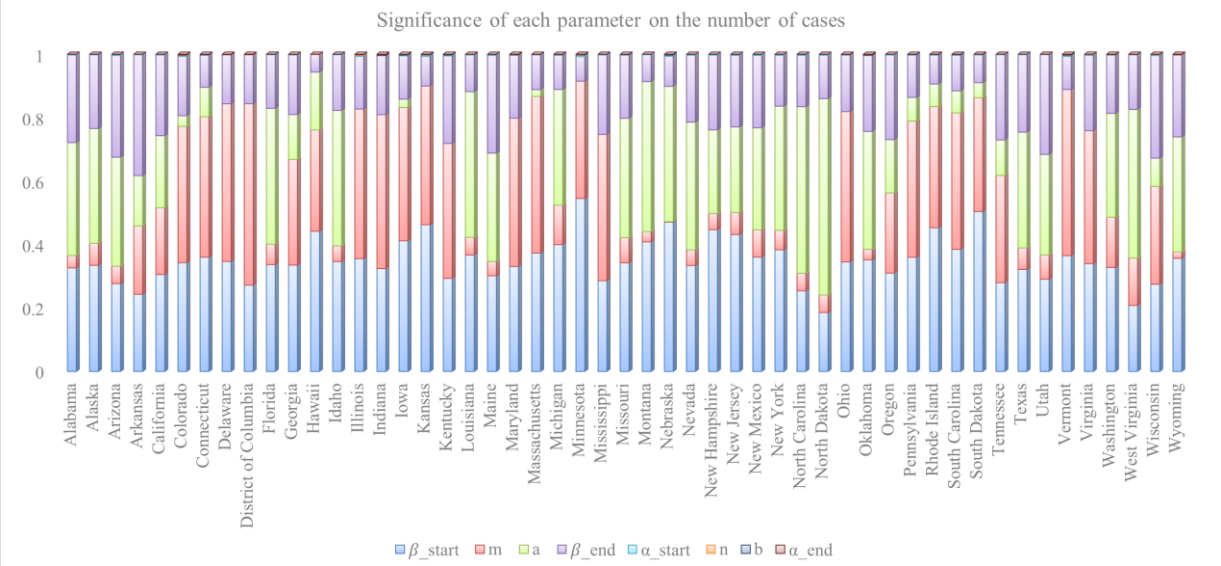

Figure 11: Significance of transmission parameter on the cases

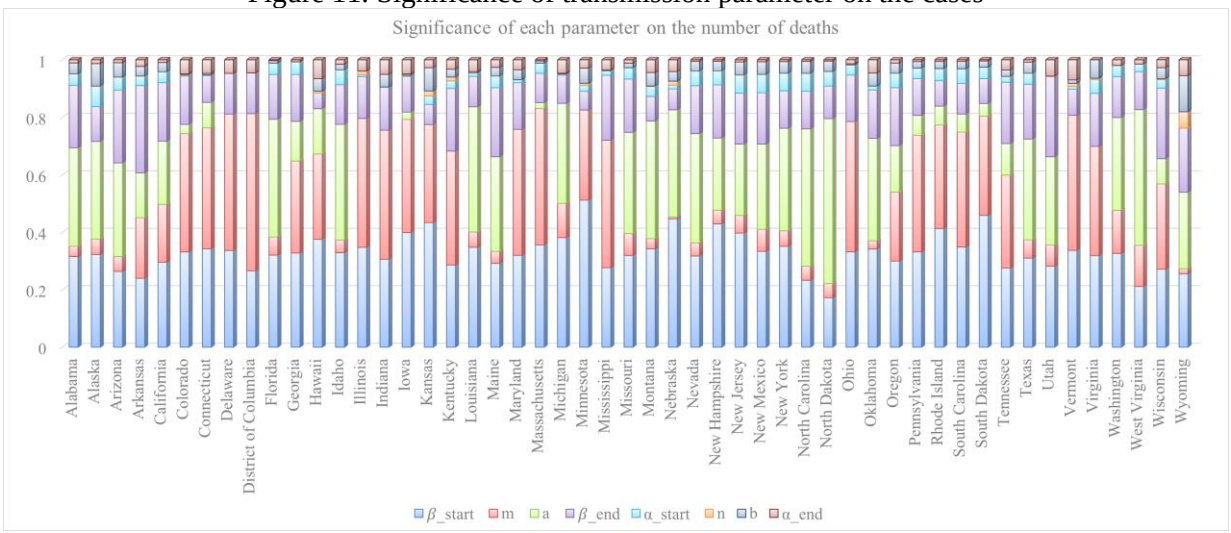

Figure 12: Significance of transmission parameter on the deaths

In Table 5, we observe that the four most significant parameters influencing the number of cases are $\beta_{start}, \beta_{end}, m, a$, which are all related to case transmission. Conversely, the parameters associated with death ($\alpha_{start}, \alpha_{end}, n, b$) exhibit minimal impact on the number of cases. Although the ranking of the first two parameters ($\beta_{start}, \beta_{end}$) may vary across states, their significance levels remain comparable.

Table 5: Rank of importance of transmission parameter on the cases

State	Rank1	Rank2	Rank3	Rank4	Rank5	Rank6	Rank7	Rank8
Alabama	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Alaska	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α

Arizona	a	β_{end}	β_{start}	m	α_{start}	x_α	α_{end}	k_α
Arkansas	β_{end}	β_{start}	m	a	α_{start}	x_α	α_{end}	k_α
California	β_{start}	β_{end}	a	m	α_{start}	x_α	α_{end}	k_α
Colorado	m	β_{start}	β_{end}	a	α_{start}	α_{end}	x_α	k_α
Connecticut	m	β_{start}	β_{end}	a	α_{start}	α_{start}	x_α	k_α
Delaware	m	β_{start}	β_{end}	α_{end}	a	x_α	α_{start}	k_α
District of Columbia	m	β_{start}	β_{end}	α_{end}	a	x_α	α_{start}	k_α
Florida	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Georgia	β_{start}	m	β_{end}	a	α_{start}	α_{end}	x_α	k_α
Hawaii	β_{start}	m	a	β_{end}	x_α	α_{end}	k_α	α_{start}
Idaho	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Illinois	m	β_{start}	β_{end}	α_{start}	k_α	α_{end}	x_α	a
Indiana	m	β_{start}	β_{end}	α_{end}	x_α	a	α_{start}	k_α
Iowa	m	β_{start}	β_{end}	a	α_{start}	x_α	α_{end}	k_α
Kansas	β_{start}	m	β_{end}	α_{start}	x_α	α_{end}	k_α	a
Kentucky	m	β_{start}	β_{end}	α_{start}	x_α	α_{end}	k_α	a
Louisiana	a	β_{start}	β_{end}	m	α_{start}	α_{end}	x_α	k_α
Maine	a	β_{end}	β_{start}	m	α_{start}	x_α	α_{end}	k_α
Maryland	m	β_{start}	β_{end}	α_{end}	x_α	α_{start}	a	k_α
Massachusetts	m	β_{start}	β_{end}	a	α_{start}	α_{end}	x_α	k_α
Michigan	β_{start}	a	m	β_{end}	α_{end}	α_{start}	x_α	k_α
Minnesota	β_{start}	m	β_{end}	α_{start}	x_α	α_{end}	k_α	a
Mississippi	m	β_{start}	β_{end}	α_{start}	α_{end}	x_α	k_α	a
Missouri	a	β_{start}	β_{end}	m	α_{end}	α_{start}	x_α	k_α
Montana	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Nebraska	β_{start}	a	β_{end}	α_{start}	x_α	m	α_{end}	k_α
Nevada	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
New Hampshire	β_{start}	a	β_{end}	m	α_{start}	x_α	α_{end}	k_α
New Jersey	β_{start}	a	β_{end}	m	α_{start}	x_α	α_{end}	k_α
New Mexico	β_{start}	a	β_{end}	m	α_{start}	x_α	α_{end}	k_α
New York	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
North Carolina	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
North Dakota	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Ohio	m	β_{start}	β_{end}	α_{end}	α_{start}	x_α	a	k_α
Oklahoma	a	β_{start}	β_{end}	m	α_{end}	x_α	α_{start}	k_α
Oregon	β_{start}	β_{end}	m	a	α_{start}	x_α	α_{end}	k_α
Pennsylvania	m	β_{start}	β_{end}	a	x_α	α_{start}	α_{end}	k_α
Rhode Island	β_{start}	m	β_{end}	a	x_α	α_{start}	α_{end}	k_α
South Carolina	m	β_{start}	β_{end}	a	α_{start}	x_α	α_{end}	k_α
South Dakota	β_{start}	m	β_{end}	a	α_{start}	x_α	α_{end}	k_α
Tennessee	m	β_{start}	β_{end}	a	α_{start}	x_α	α_{end}	k_α

Texas	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Utah	a	β_{end}	β_{start}	m	α_{end}	x_α	k_α	α_{start}
Vermont	m	β_{start}	β_{end}	α_{start}	x_α	α_{end}	k_α	a
Virginia	m	β_{start}	β_{end}	α_{start}	x_α	k_α	a	α_{end}
Washington	β_{start}	a	β_{end}	m	α_{end}	α_{start}	x_α	k_α
West Virginia	a	β_{start}	β_{end}	m	α_{end}	α_{start}	x_α	k_α
Wisconsin	β_{end}	m	β_{start}	a	α_{start}	x_α	α_{end}	k_α
Wyoming	a	β_{start}	β_{end}	m	x_α	k_α	α_{end}	α_{start}

Regarding the influence of the eight parameters on the number of deaths in Table 6, it is evident that the four case-related parameters ($\beta_{start}, \beta_{end}, m, a$) continue to play a critical role. This is primarily because the number of cases determines the baseline number of deaths. However, when comparing the local sensitivity analysis for cases and deaths, it becomes apparent that the death-related parameters ($\alpha_{start}, \alpha_{end}, n, b$) have a more pronounced effect on the number of deaths.

Table 6: Rank of importance of transmission parameter on the deaths

State	Rank1	Rank2	Rank3	Rank4	Rank5	Rank6	Rank7	Rank8
Alabama	a	β_{start}	β_{end}	α_{start}	x_α	m	α_{end}	k_α
Alaska	a	β_{start}	β_{end}	x_α	α_{start}	m	α_{end}	k_α
Arizona	a	β_{start}	β_{end}	m	x_α	α_{start}	α_{end}	k_α
Arkansas	β_{end}	β_{start}	m	a	x_α	α_{start}	α_{end}	k_α
California	β_{start}	a	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Colorado	m	β_{start}	β_{end}	α_{end}	a	k_α	α_{start}	x_α
Connecticut	m	β_{start}	β_{end}	a	α_{end}	x_α	α_{start}	k_α
Delaware	m	β_{start}	β_{end}	α_{end}	a	x_α	α_{start}	k_α
District of Columbia	m	β_{start}	β_{end}	α_{end}	a	x_α	α_{start}	k_α
Florida	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Georgia	β_{start}	m	β_{end}	a	α_{start}	α_{end}	x_α	k_α
Hawaii	β_{start}	m	a	α_{end}	β_{end}	x_α	k_α	α_{start}
Idaho	a	β_{start}	β_{end}	α_{start}	m	x_α	α_{end}	k_α
Illinois	m	β_{start}	β_{end}	α_{end}	k_α	α_{start}	x_α	a
Indiana	m	β_{start}	β_{end}	α_{end}	x_α	a	α_{start}	k_α
Iowa	β_{start}	m	β_{end}	α_{end}	a	x_α	α_{start}	k_α
Kansas	β_{start}	m	x_α	β_{end}	α_{start}	α_{end}	k_α	a
Kentucky	m	β_{start}	β_{end}	α_{end}	x_α	α_{start}	k_α	a
Louisiana	a	β_{start}	β_{end}	m	α_{end}	α_{start}	x_α	k_α
Maine	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Maryland	m	β_{start}	β_{end}	α_{end}	x_α	α_{start}	a	k_α
Massachusetts	m	β_{start}	β_{end}	α_{start}	a	x_α	α_{end}	k_α
Michigan	β_{start}	a	m	β_{end}	α_{end}	x_α	α_{start}	k_α

Minnesota	β_{start}	m	β_{end}	x_α	α_{end}	α_{start}	k_α	a
Mississippi	m	β_{start}	β_{end}	α_{end}	α_{start}	x_α	k_α	a
Missouri	a	β_{start}	β_{end}	m	α_{start}	x_α	α_{end}	k_α
Montana	a	β_{start}	β_{end}	x_α	α_{end}	m	α_{start}	k_α
Nebraska	β_{start}	a	β_{end}	α_{end}	x_α	k_α	α_{start}	m
Nevada	a	β_{start}	β_{end}	α_{start}	m	x_α	α_{end}	k_α
New Hampshire	β_{start}	a	β_{end}	α_{start}	m	x_α	α_{end}	k_α
New Jersey	β_{start}	a	β_{end}	α_{start}	m	x_α	α_{end}	k_α
New Mexico	β_{start}	a	β_{end}	m	α_{start}	x_α	α_{end}	k_α
New York	a	β_{start}	β_{end}	α_{start}	m	x_α	α_{end}	k_α
North Carolina	a	β_{start}	β_{end}	α_{start}	m	x_α	α_{end}	k_α
North Dakota	a	β_{start}	β_{end}	α_{start}	m	x_α	α_{end}	k_α
Ohio	m	β_{start}	β_{end}	α_{start}	α_{end}	x_α	a	k_α
Oklahoma	a	β_{start}	β_{end}	α_{end}	x_α	m	α_{start}	k_α
Oregon	β_{start}	m	β_{end}	a	α_{start}	x_α	α_{end}	k_α
Pennsylvania	m	β_{start}	β_{end}	a	α_{start}	x_α	α_{end}	k_α
Rhode Island	β_{start}	m	β_{end}	a	α_{start}	x_α	α_{end}	k_α
South Carolina	m	β_{start}	β_{end}	a	α_{start}	x_α	α_{end}	k_α
South Dakota	β_{start}	m	β_{end}	a	α_{start}	x_α	α_{end}	k_α
Tennessee	m	β_{start}	β_{end}	a	α_{end}	x_α	α_{start}	k_α
Texas	a	β_{start}	β_{end}	m	x_α	α_{start}	α_{end}	k_α
Utah	a	β_{start}	β_{end}	m	α_{end}	x_α	k_α	α_{start}
Vermont	m	β_{start}	β_{end}	α_{end}	x_α	k_α	α_{start}	a
Virginia	m	β_{start}	β_{end}	x_α	α_{start}	k_α	a	α_{end}
Washington	β_{start}	a	m	β_{end}	α_{start}	α_{end}	x_α	k_α
West Virginia	a	β_{start}	m	β_{end}	α_{start}	α_{end}	x_α	k_α
Wisconsin	m	β_{start}	β_{end}	a	x_α	α_{start}	α_{end}	k_α
Wyoming	a	β_{start}	β_{end}	x_α	k_α	α_{end}	m	α_{start}

Another method is multi-way sensitivity analysis, which examines the interactions between multiple input parameters and their effect on the model output. Multi-way sensitivity analysis can help identify the combined effects of multiple input parameters, which can provide more comprehensive insights into the factors driving the spread of the virus. Monte Carlo simulation methods are increasingly popular in the field of epidemiology and data science for conducting sensitivity analysis. Monte Carlo methods involve generating multiple sets of input parameters by randomly sampling from probability distributions. This approach allows for the assessment of the uncertainty associated with input parameters and the propagation of that uncertainty through the

model to obtain a distribution of model outputs. Monte Carlo sensitivity analysis can provide a more comprehensive understanding of the uncertainty associated with model predictions, which is critical for informing decision-making in pandemic response planning.

One of the advantages of Monte Carlo methods over other sensitivity analysis methods is that they can handle complex models with many input parameters and non-linear relationships between those parameters. Monte Carlo methods can also account for correlation between input parameters and identify the most influential parameters on the model output. This approach enables the identification of the most critical factors that drive the spread of the virus, such as the reproduction number (R0), the rate of transmission, the incubation period, and the severity of the disease.

In pandemic modeling, Monte Carlo sensitivity analysis can help policymakers develop more effective strategies for controlling the spread of the virus. By understanding the impact of changes in key parameters on the model output, policymakers can develop targeted interventions to mitigate the spread of the virus. For example, policymakers can use Monte Carlo sensitivity analysis to assess the impact of different social distancing measures, vaccination campaigns, or targeted testing on the spread of the virus. Overall, sensitivity analysis is an essential tool for modeling the spread of pandemics, and Monte Carlo methods can provide a more comprehensive understanding of the uncertainty associated with model predictions, enabling informed decision-making in pandemic response planning.

In our analysis, we utilize the Monte Carlo simulation method to assess the combinational effect of parameter uncertainties on the two primary outputs of each state: Cases and Deaths. Each parameter is assumed to follow a uniform distribution within a range of ±5% of their optimal fitted values. For every simulation iteration, we randomly draw samples of these parameters from their respective distributions and incorporate them into our time-varying transmission model. By

conducting 3,000 Monte Carlo simulations, we are able to effectively evaluate the collective impact of parameter uncertainties and provide a robust analysis of the model's outcomes. We highlight the results from four representative states (California, Florida, New York, Michigan), showcasing the 5_{th} to 95_{th} percentile range of both cases and deaths in Figure 13 and Figure 14.

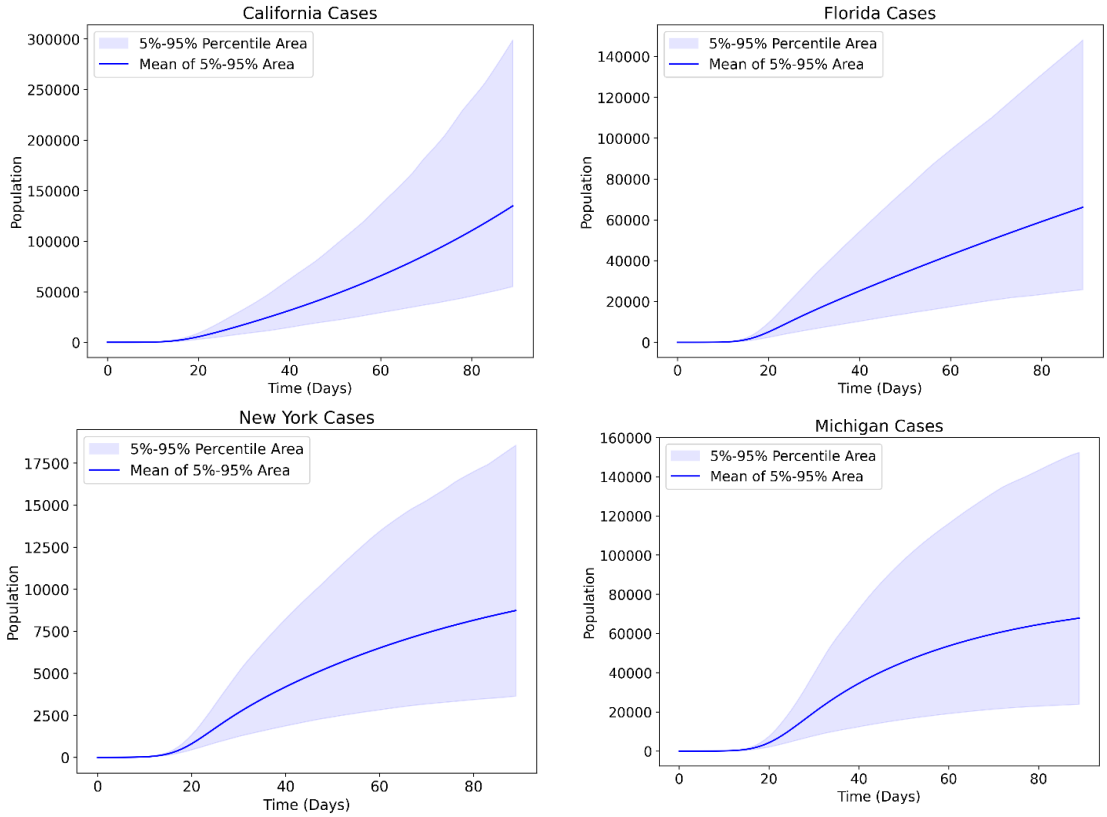

Figure 13: Simulation results of the 5th to 95th percentile range of cases

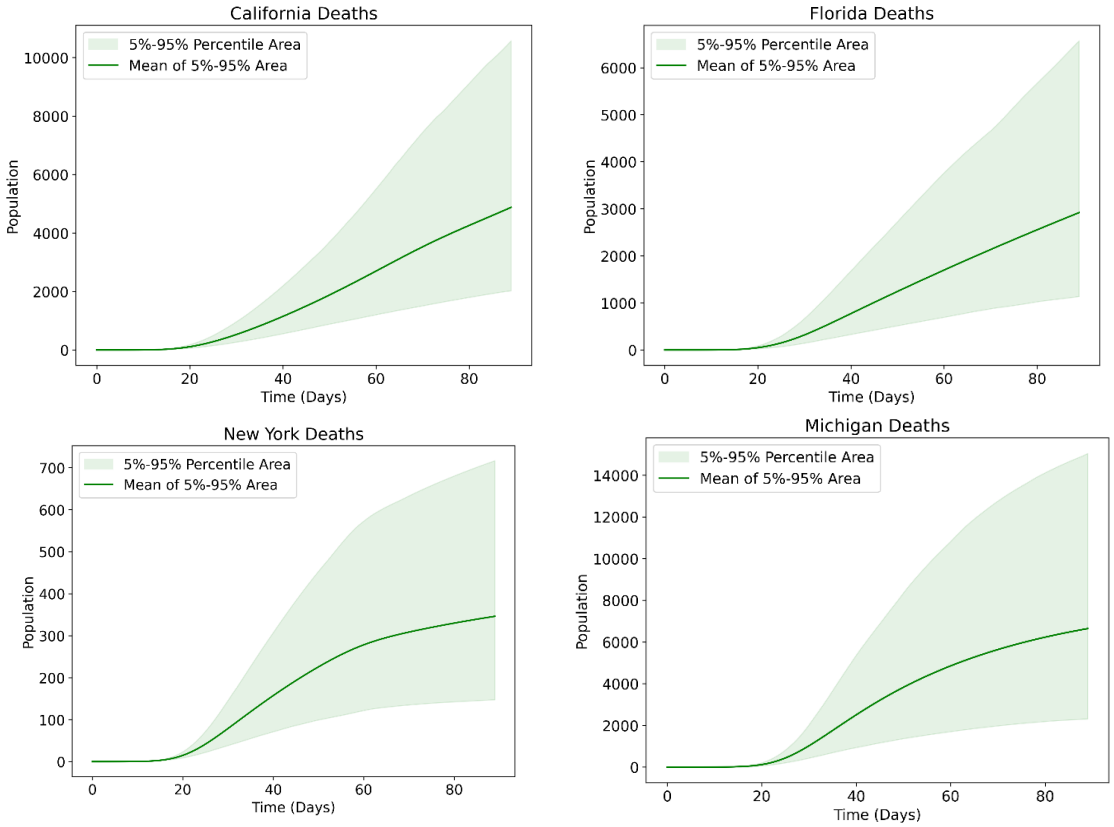

Figure 14: Simulation results of the 5th to 95th percentile range of deaths

In our Monte Carlo Analysis, we discovered that even minor fluctuations ($\pm 5\%$) in the parameters can result in significant changes in healthcare outcomes. This finding highlights the inherent challenges associated with forecasting future transmission dynamics of infectious diseases, as transmission parameters can be easily altered by a multitude of factors. For example, the emergence of new virus strains, local events, government policies, and environmental conditions can all lead to sudden shifts in transmission rates.

Likewise, death rates can be influenced by various factors, such as the availability of healthcare resources, advancements in treatment protocols, or the emergence of more virulent strains. The dynamic nature of these factors further complicates the prediction of disease transmission, as each element can have a considerable impact on the overall trajectory of the outbreak. Additionally, changes in individual and collective behaviors, such as adherence to social distancing guidelines,

mask-wearing, and hygiene practices, can have substantial effects on transmission rates, further contributing to the complexity of predicting future disease dynamics.

Considering the multitude of factors that can influence transmission parameters, it is not implausible that these parameters may change by more than 5% within a day or two. Such sudden shifts can significantly increase the uncertainty associated with forecasting future disease trends.

4.7 Conclusion and Discussion

In this chapter, we presented an extension of the SEIRD model, which incorporates dynamic changes in death and transmission rates over time using a continuous Sigmoid function. We hypothesized that these rates change continuously, rather than abruptly, in response to public health policies or treatment implementation. Our model demonstrated a strong fit to historical data for the early months of the pandemic in the United States, with a median RRMSE of 1.33% for deaths and 1.88% for cases across the 50 states. The median effective reproduction rate at the onset of the pandemic was 2.87, which we estimated had dropped below 1 for all states by July 28, 2020. We observed that states with poorer model fits typically experienced multiple waves of the disease. To account for these discrepancies, we proposed a multi-phase extension of the model, in which transitions between phases are marked by changes in public health policy. Applying this two-phase model to Hawaii as a case study, we observed a significant improvement in the model's accuracy, with RRMSE values decreasing to 2.5% for cases and 2.7% for deaths.

A key advantage of our model is the minimal number of parameters required to represent dynamic changes in transmission and death rates, enabling efficient quantification of regional and temporal differences in disease spread and outcomes. By examining historical trends, our model offers valuable insights into how variations in simple parameters can influence the number of cases and deaths, informing future policy and intervention strategies.

Additionally, our sensitivity analysis underscores the importance of certain parameters in shaping the predicted number of cases and deaths. This understanding can aid researchers and policymakers in anticipating potential shifts in disease transmission and devising effective interventions. Moreover, our Monte Carlo simulations reveal the substantial impact of uncertainties in parameter values on model outcomes, highlighting the difficulty in forecasting of the future trend.

In conclusion, our extended SEIRD model, which accounts for continuous changes in transmission and death rates, provides a powerful tool for understanding the complex dynamics of infectious diseases and guiding evidence-based public health decision-making.

Chapter 5 Integration of Dynamic Modeling with Spatial Interaction and Effect Analysis

5.1 Introduction

Transportation plays a crucial role in the transmission of pandemics such as COVID-19. The virus can be spread through respiratory droplets that are released when an infected person talks, coughs, or sneezes. These droplets can then be inhaled by other individuals in close proximity to the infected person. Transportation modes such as buses, trains, and airplanes are high-risk areas for the transmission of the virus, as they often involve large numbers of people in enclosed spaces for extended periods of time. In addition, long-distance travel can significantly impact the transmission of diseases to new areas. When people travel long distances, they can bring infectious agents with them, such as bacteria, viruses, and parasites. Modes of transportation, such as planes, trains, and ships, can increase the risk of disease transmission due to the close proximity of travelers and the confined spaces they share. Additionally, travelers may come into contact with infectious agents through contaminated surfaces or by interacting with infected individuals. These agents can then spread to new populations and new area, potentially causing outbreaks of disease. Therefore, studying the impact of long-distance travel on the transmission of diseases is crucial for understanding how infectious diseases can spread across different regions and populations. This knowledge can inform public health policies and strategies aimed at preventing and controlling the spread of infectious diseases. Studying the impact of long-distance travel on disease transmission can also inform broader discussions about global health and the interconnectedness of populations around the world. As travel becomes more common and widespread, it is

increasingly important to understand how diseases can be transmitted across borders and how to prevent the spread of infectious agents.

This chapter aims to investigate the transmission of diseases through spatial interaction, with a specific focus on state-level travel. We propose transmission export index related to transportation to assess the impact of long-distance travel on disease transmission. By gaining a better understanding of the effects of travel on disease transmission, we can analyze historical disease outbreaks and develop effective strategies for preventing and controlling future outbreaks.

5.2 Multi-Regional Dynamic Modeling with Spatial Interaction

5.2.1 Transportation Data Collection and Its Challenge

In addition to fundamental data pertaining to disease transmission, including case and mortality statistics and basic information about the causative pathogen, transportation data is an essential variable for examining the effect of spatial interactions on the spread of infectious diseases. The COVID-19 pandemic has presented significant challenges for collecting transportation data and analyzing the impact of travel on disease transmission. The pandemic has caused a sharp decrease in travel activity, with many countries implementing travel restrictions and lockdown measures to slow the spread of the virus. This has made it difficult to collect reliable and comprehensive transportation data to analyze the effects of travel on disease transmission.

First, many traditional data sources, such as surveys and manual counts, may not be feasible due to social distancing measures and restrictions on non-essential activities. This makes it difficult to obtain accurate information on travel patterns and transportation demand.

Moreover, the pandemic has also changed the way people travel. With more people working from home, there has been a shift from public transit to private vehicles, which may not be captured in traditional transportation data. Additionally, the pandemic has also led to changes in the timing

and frequency of travel, which may further complicate the collection and analysis of transportation data.

Another challenge is the issue of privacy and data protection. Collecting transportation data, particularly data related to individuals' movements, raises concerns about privacy and data protection. This can make it difficult to obtain the necessary data to conduct meaningful analyses of the impact of travel on disease transmission, particularly in countries with strict data protection laws.

To address these challenges, transportation researchers and practitioners have turned to alternative data sources, such as mobile phone location data, to track changes in transportation patterns during the pandemic. With the help of innovative data sources and methods, it is possible to gain insights into the impact of COVID-19 on transportation systems. This knowledge is crucial for developing effective policies and strategies to address the ongoing pandemic and future public health crises. After a thorough search, we found two main data sources for the spatial interactions: i) Trips by Distance Data, ii) The COVID-19 Impact Analysis Platform.

i) Trips by Distance Data

Bureau of Transportation Statistics collected and curated data on the number of trips taken in the United States by distance, mode of transportation, and purpose of trip, provided on the website of https://data.bts.gov/Research-and-Statistics/Trips-by-Distance/w96p-f2qv. The data is available for the years 2019 to 2022, and the daily travel estimates are based on a merged mobile device data panel that addresses issues with geographic and temporal variation.

Trips are defined as movements that include a stay of longer than 10 minutes at an anonymized location away from home, and the data captures travel by all modes of transportation. It is considered multiple trips when a movement involves multiple stops of more than 10 minutes

before returning home. The data is analyzed at the national, state, and county levels, and a weighting procedure is used to ensure the sample of mobile devices is representative of the entire population in a given area. To protect confidentiality and support data quality, data for a county is not reported if there are fewer than 50 devices in the sample on any given day.

It is important to note that the data is experimental and may not meet all quality standards. However, these experimental data products are created to provide valuable insights to data users in the absence of other relevant products. We will combine the dataset with the following dataset provided by the COVID-19 Impact Analysis Platform to generate the spatial interaction flow.

ii) The COVID-19 Impact Analysis Platform

The COVID-19 Impact Analysis Platform is developed by Maryland Transportation Institute (MTI) and Center for Advanced Transportation Technology Laboratory (CATT Lab) cooperatively, which is a comprehensive data analysis tool designed to provide insights into the impact of the COVID-19 pandemic on communities across the United States. The platform provides a range of data and analytical tools, including interactive maps, visualizations, and dashboards, to help users better understand the spread of the virus and its impact on various social and economic indicators. One of the key features of the platform is its ability to integrate multiple data sources, including public health data, mobility data, and socioeconomic data, to provide a more comprehensive picture of the pandemic's impact. Specifically, the mobility data tracks daily visits to different types of locations, such as retail and recreation areas, transit stations, workplaces, and grocery stores, and compares them to pre-pandemic levels. The mobility data is derived from anonymized and aggregated data from mobile devices, such as smartphones and tablets, that have opted into location tracking services. The data is aggregated at the county level in the United States, and at the national level for other countries. For the analysis of spatial interactions, the platform

specifically provides the state/county level percentage of out-of-state/out-of-county trips per day from Jan 1st, 2020 to April 30th, 2021.

We can combine the mobility data provided by the COVID-19 Impact Analysis Platform and the daily trips from BTS to calculate the daily out-of-state trips for each state, which can be useful for understanding how people's movements have changed during the pandemic, and for identifying areas where social distancing measures may be effective or where there may be increased risk of COVID-19 transmission. However, it's important to note that both the mobility datasets are based on a sample of mobile devices and may not be representative of the entire population, and that the datasets are anonymized and aggregated to protect user privacy.

5.2.2 Gravity Modeling of the State-Level Transportation

Based on the best datasets we mention above, we can only get the daily out-of-state trips for each state. However, for the analysis of spatial interaction, we should also calculate the daily trips that go into each state. One common model that is widely applied in various transportation contexts to estimate the flow of passenger or goods between cities, which is based on Newton's law of universal gravitation. The model is relatively straightforward and involves estimating the transportation flow between two locations based on their mass and distance.

The basic principle behind the gravity model is that the flow of transportation between two locations is proportional to the product of their masses and inversely proportional to the distance between them.

The gravity model is widely used in transportation modeling due to its simplicity and ability to provide accurate predictions of transportation flows. One of the key advantages of the gravity model is that it can be applied to a wide range of transportation contexts, including freight transportation, passenger transportation, and tourism. The model can also be applied to various

transportation modes, including air, sea, and land transportation. Additionally, the gravity model can be adapted to include other variables that may influence transportation flows, such as population density, income, or trade barriers. The model's simplicity and ability to provide accurate predictions of transportation flows make it a valuable tool for transportation planners and policymakers seeking to understand and optimize transportation networks.

Despite its usefulness, the gravity model has some limitations that should be considered. One of the main limitations of the model is that it assumes that transportation flows are solely dependent on the mass and distance between two locations, and it does not consider other discrete factors that may affect transportation demand, such as differences in regional economies, cultural preferences, or political factors. In addition, another limitation of the gravity model is that it may be challenging to estimate accurate values for the model parameters. For instance, estimating the exponents of the model can be challenging since they are typically determined through statistical regression analysis, and the estimates may vary depending on the data used for calibration. Furthermore, the parameters' values may also vary depending on the transportation mode or the nature of the goods being transported.

Despite these limitations, the gravity model remains a valuable tool for transportation modeling and has been extensively used in research and policy analysis. Various extensions of the model have been proposed to address some of the limitations mentioned earlier. For example, some researchers have proposed incorporating the network topology of transportation systems or accounting for heterogeneity in transportation preferences across different population groups.

In conclusion, the gravity model is a powerful and versatile tool for transportation modeling that has been widely used in various transportation contexts. The model's simplicity and ability to provide accurate predictions of transportation flows make it a valuable tool for transportation

planners and policymakers seeking to optimize transportation networks. However, it is essential to consider the model's limitations and potential extensions when applying it to real-world transportation planning and policy analysis.

The gravity model has several common features, including the Gross Domestic Product (GDP) and distance. The GDP is a measure of the economic activity within a particular country. The GDP is used in the gravity model to represent the mass of a particular location. A higher GDP is often associated with a higher demand for transportation goods and services. This is because a higher GDP typically implies higher economic activity, which may increase the demand for transportation of goods and services. In the gravity model, the GDP of each location is used to estimate the mass of the location, which is then used to estimate the flow of transportation between the two locations. Distance is another essential feature in the gravity model. The model posits that transportation flows between two locations are inversely proportional to the distance between them. This means that as the distance between two locations increases, the flow of transportation between them decreases. Distance is often considered a critical factor in transportation demand and can significantly impact transportation infrastructure planning and investment decisions. In the gravity model, distance is incorporated into the model through the denominator of the formula, where transportation flows decrease as the distance between two locations increases.

The formula for the gravity model of trip distribution is expressed in Equation (14):

$$M_{ij} = M_i * \frac{\frac{G_j^{\alpha}}{D_{ij}^{\gamma}}}{\sum_k \frac{G_k^{\alpha}}{D_{ik}^{\gamma}}} \tag{14}$$

Where M_{ij} represents the flow of trips from region i to region j, G_k is the GDP of the location k, D_{ik} represents the distance between the region i and region k, and α and γ are exponents that determine the relative influence of the variables. $\frac{G_j^{\alpha}}{D_{ij}^{\gamma}}$ indicates the attraction index of region j

calculated by gravity model. The ratio $\frac{\frac{G_j^{\alpha}}{D_{ij}^{\gamma}}}{\sum_k \frac{G_k^{\alpha}}{D_{ik}^{\gamma}}}$ shows the proportion of the total trips goes from region i to region j. Due to the limitation of real trip flow data from region i to region j, we cannot fit the best-fitted values of α and γ for the real case. However, a reasonable assumption that follows the correlation is both the α and γ equal to 2, which is same as the Newton's gravity model.

5.2.3 Dynamic Modeling with Multi-Regional Spatial Interaction

Based on the dynamic modeling with time-varying transmission and fatality rates, we will further increase the applicability of model (1) to incorporate the effect of multi-regional spatial interaction. In general, state-level transportation can influence the movement of people between regions, increasing the likelihood of contact between individuals from different regions and leading to the spread of the virus.

For the susceptible population, who have not been infected with the virus and can become infected if exposed, state-level transportation can increase the size of the susceptible population by bringing individuals from different regions into contact with each other, increasing the likelihood of exposure to the virus.

For the exposed population, who have been infected with the virus but have not yet developed symptoms, state-level transportation can increase the size of the exposed population by facilitating the movement of infected individuals across regions, increasing the likelihood of exposure to susceptible individuals in other regions. For example, individuals who are infected with the virus and travel through airports or highways may spread the virus to other regions, leading to an increase in the number of individuals in the exposed population.

Meanwhile, some states that are geographically close to each other tend to have more spatial interactions compared to far-connected states. For example, New York states have higher volume

of transportation with New Jersey due to geographic proximity, strong transportation infrastructure, and cultural and social connections. Hence, for the simplicity of the modeling, it is reasonable to assume the effect of state-level transportation on the transmission is consistent within the same region.

We use the Standard Federal Regions to aggregate the 50 united states into 10 parts. The regions were defined based on geographic, economic, and cultural factors. They were designed to promote efficient and effective delivery of federal programs and services by bringing together federal agencies, state and local governments, and private organizations to work collaboratively and address regional issues and concerns. The states within the same region have proven to be able to share resources, expertise, and best practices across state lines and jurisdictions. The 10 Standard Federal Regions (shown in Figure 15) are as follows:

Region 1: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont.

Region 2: New Jersey, New York, Puerto Rico, and the Virgin Islands.

Region 3: Delaware, Maryland, Pennsylvania, Virginia, West Virginia, and the District of Columbia.

Region 4: Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee.

Region 5: Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin.

Region 6: Arkansas, Louisiana, New Mexico, Oklahoma, Texas.

Region 7: Iowa, Kansas, Missouri, and Nebraska.

Region 8: Colorado, Montana, North Dakota, South Dakota, Utah, and Wyoming.

Region 9: Arizona, California, Hawaii, Nevada, the Pacific Trust Territories.

Region 10: Alaska, Idaho, Oregon, and Washington.

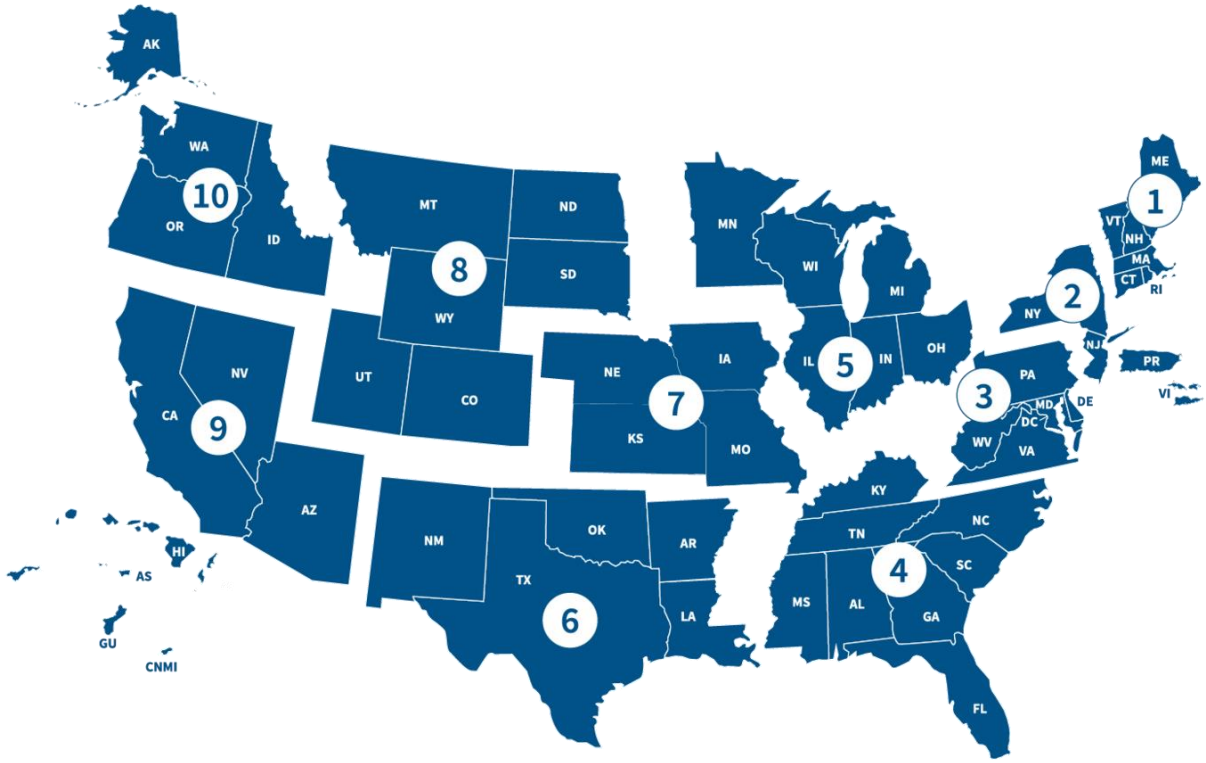

Figure 15: Ten regions defined by the Federal Emergency Management Agency in the continental United States and territories

According to the above assumptions, the modified model is shown in Equation (15).

$$
\begin{aligned}
\frac{\partial \mathrm{S_i(t)}}{\partial \mathrm{t}} &= -\beta(t)\cdot I_i(t)\cdot\frac{S_i(t)}{N_i}+\sum_{j\neq i}\mu_{kj}\sum_{j\neq i}\frac{M_{ij}S_j}{N_j-I_j}-\mu_{ki}\sum_{j\neq i}\frac{M_{ji}S_i}{N_i-I_i}\\
\frac{\partial E_i(t)}{\partial t} &= \beta(t)\cdot I_i(t)\cdot\frac{S_i(t)}{N_i}-\sigma\cdot E_i(t)+\sum_{j\neq i}\mu_{\mathrm{kj}}\sum_{j\neq i}\frac{M_{ij}E_j}{N_j-I_j}-\mu_{\mathrm{ki}}\sum_{j\neq i}\frac{M_{ji}E_i}{N_i-I_i}\\
\frac{\partial I_i(t)}{\partial t} &= \sigma\cdot E_i(t)-\big(1-\tau\cdot\alpha(t)\big)\cdot\gamma I_i(t)-\tau\cdot\alpha(t)\cdot\rho\cdot I_i(t)\\
\frac{\partial R_i(t)}{\partial t} &= \tau\cdot\big(1-\alpha(t)\big)\cdot\gamma\cdot I_i(t)\\
\frac{\partial D_i(t)}{\partial t} &= \tau\cdot\alpha(t)\cdot\rho\cdot I_i(t)
\end{aligned}
\tag{15}
$$

where $S_i(t), E_i(t), I_i(t), R_i(t), D_i(t)$ *and* N_i are the susceptible, exposed, infected, recovered, dead and total population in region i at time t. θ is defined as the protection rate of the vaccine, indicating the percentage of vaccinated people who are truly immune. τ is a scalar factor that shows the reduction effect of the vaccination on the death rate. Spatial interaction between cities is represented by the daily number of people traveling from region j to region i and an adjustable

factor μ_{kj}. It is still possible that infectious people, who have symptoms and is conscious about their status infection, still travel locally and sometimes between different regions. However, for simplicity, we assume the infectious people will not commit a state-level travel in general. The rare case of the traveling infected people will be represented by the adjustable factor μ_{kj}. Meanwhile, given that the total number of individuals moving in and out of region i is substantially smaller than the region's overall population, we've simplified our model by disregarding changes in total population due to multi-regional transportation.

Comparing to the parameter estimation for a single state, finding the best-fitted values for the transmission parameters simultaneously for all 50 states is more complicated and time-consuming. For a single state, we only have 8 parameters to be fitted. However, due to the travel of exposed and susceptible population, multi-regional spatial interactions make the compartments (i.e. susceptible, exposed, infectious, recovered) for one state depend on the compartments of other states. In order to get the best-fitted values for the transmission parameters considering the spatial interaction, we need to simulate the 50 states together and find out all 410 parameters at the same time. Parameter estimation procedure from Chapter 4 basically consists of two parts: i) calculate the value for the cases and death by solving the transmission differential equations with the scipy odeint solver, ii) iteratively optimize the transmission parameters with Levenberg–Marquardt algorithm. However, solving the 200 differential equations using the fourth-order Runge-Kutta method with the odeint solver is time-consuming. Additionally, the fourth-order Runge-Kutta method necessitates interpolation of data points between two days, which significantly prolongs the simulation time and introduces the issue of gradient vanishing due to the implicit function involved in the interpolation process. Therefore, we opted to replace the third-party ode solver with our own custom function that employs the first order Runge-Kutta method to solve the

ordinary differential equations. This modification reduces the simulation time from 10 hours to 2 hours while yielding improved fitting results.

5.2.4 Model Accuracy

We fit the model with the dataset of 7-day moving average cases and deaths for the 50 states, provided by the COVID-19 tracking project lead by *The Atlantic* (derived from the Center for Disease Control), for each 30 days from 03/15/2020 to 10/15/2020. The fitting accuracy across all states is presented in Figure 16, measured by the relative root mean square error (RRMSE) defined in Equation (12).

The average fitting accuracy of the reported cases over 7 months ranges from 0.54% to 3.78% and of the reported deaths ranges over 7 months from 0.24% to 2.49%. The average and median RRMSEs for cases are 1.54% and 1.48%; for deaths, the average and median values are 1.20% and 1.14%. Figure 16 shows the average RRMSE for cases and deaths over 7 months across 50 states of the dynamic model with multi-regional spatial interaction.

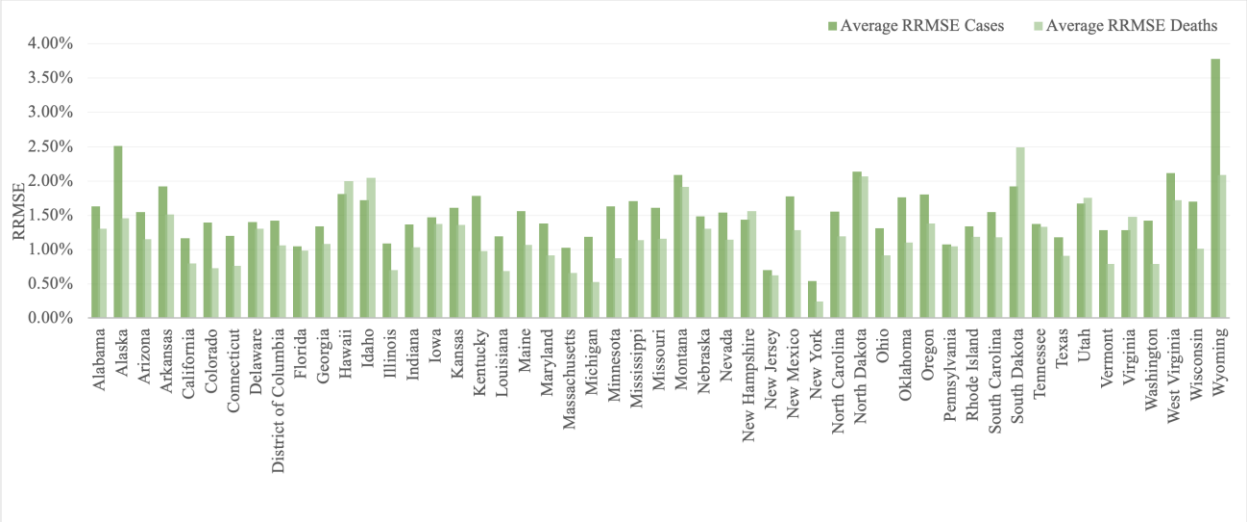

Figure 16: Average RRMSE for cases and deaths over 7 months across 50 states

Figure 17 and Figure 18 display the fitted results for COVID-19 cases and deaths in example states (Georgia, New Jersey, Florida, and Maryland), during the period from October 15, 2020, to November 15, 2020.

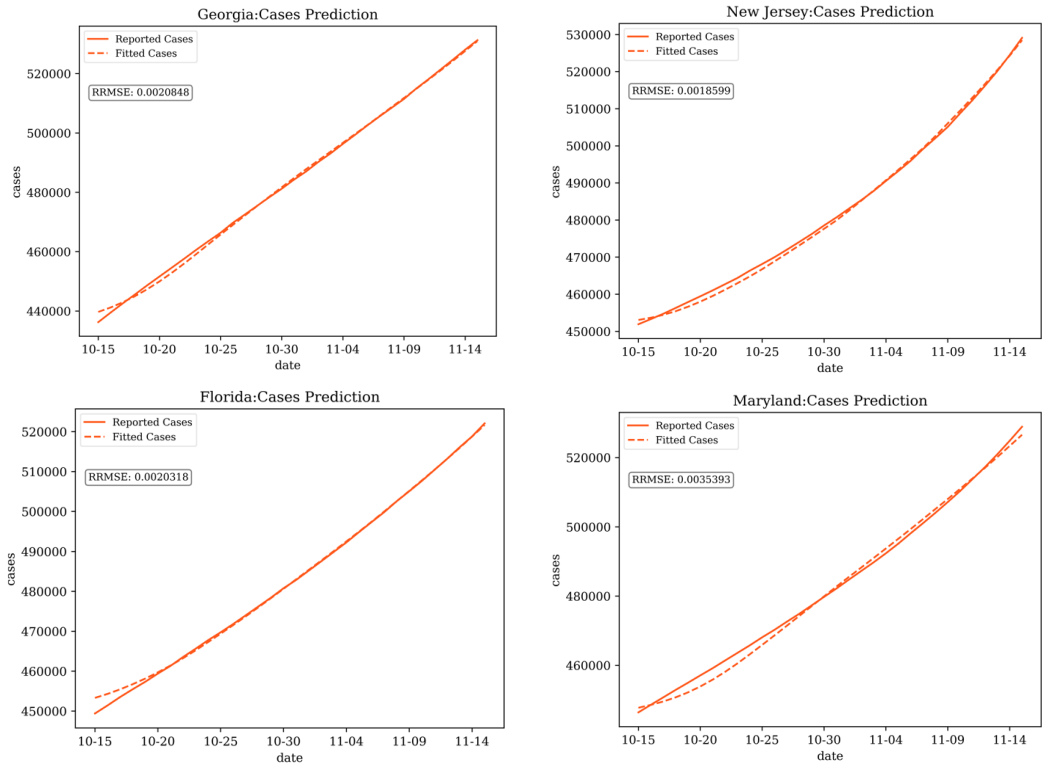

Figure 17:Fitting results of case number for Georgia, New Jersey, Florida, and Maryland

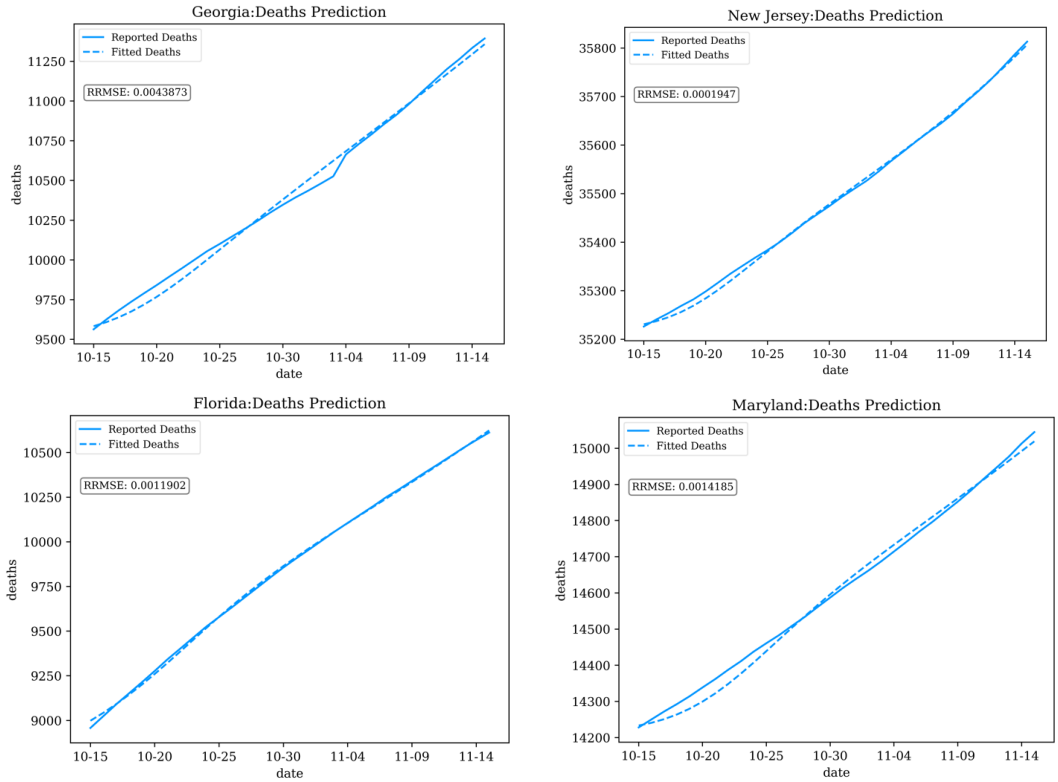

Figure 18:Fitting results of death number for Georgia, New Jersey, Florida, and Maryland

In summary, the dynamic modeling with multiregional spatial interaction, demonstrates a high degree of accuracy in capturing the historical transmission dynamics of infectious diseases. This method effectively accounts for the complexities and interactions between various regions, leading to a more comprehensive understanding of the factors influencing disease spread and the effectiveness of control measures.

5.3 Effect Evaluation of Transportation on the Multi-Regional Transmission

5.3.1 Transmission Export Index

According to the model (15), the parameter μ_i represents the average number of individuals from other regions who could potentially contract the infection upon contact with a single infectious traveler originating from region i. The parameter β_i, on the other hand, denotes the average number of people who could become infected through contact with a single infectious individual within region i, thus reflecting the local transmissibility of the disease.

To evaluate the risk of disease transmission from one region to another due to multi-regional spatial interaction, we introduce a transmission export index for region i, defined as $\beta_i * \mu_i$. This index incorporates both μ_i and β_i, with μ_i accounting for the potential spread of the infection to new regions by an infectious traveler and β_i indicating the local increase in the number of infected individuals.

The transmission export index provides valuable insights into the extent to which a region poses a transmission risk to other regions through travel. If a region is experiencing a surge in transmission (i.e., high β_i) and travelers from that region exhibit a higher propensity to spread the disease to other areas (i.e. high μ_i), it is considered to pose a greater risk to other regions via travel.

Policymakers in low-risk regions must be proactive in implementing strict travel regulations or quarantine measures for travelers originating from such high-risk areas.

In summary, the transmission export index is a vital tool in assessing the potential risk of disease transmission from one region to another through multi-regional spatial interaction. By considering both the local transmissibility (β_i) and the ability of infectious travelers to spread the disease to new regions (μ_i), this index serves as a valuable guide for healthcare experts and policymakers alike. By identifying high-risk regions, appropriate interventions such as stringent travel restrictions or quarantine measures can be put in place, ultimately mitigating the spread of infectious diseases and safeguarding public health. Figure 19 shows the heatmap of infectious export index for all 50 states in the US from 03/15/2020 to 04/15/2020.

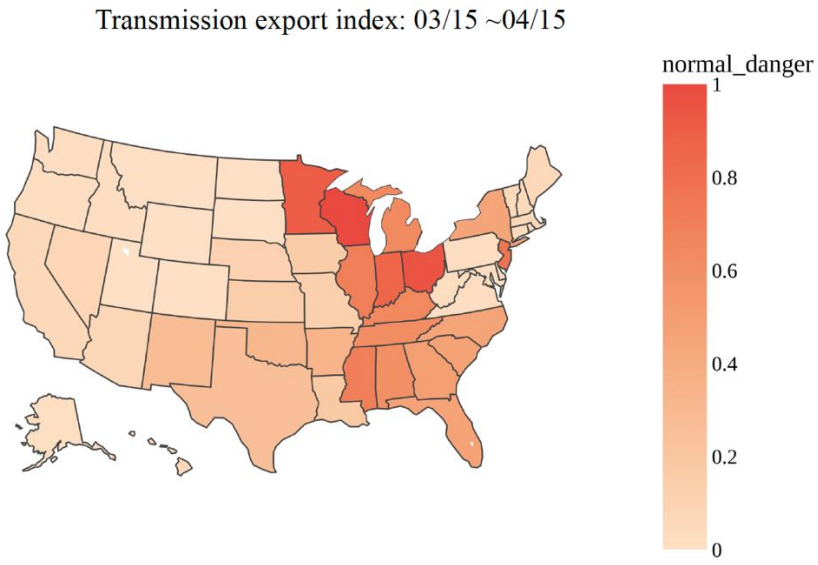

Figure 19: Heatmap of infectious export index for all 50 states in the US from 03/15/2020 to 04/15/2020

5.3.2 Causal Analysis for Increment of Transmission Export Index

In the following discussion, we will delve into the possible reasons for the increase in the transmission export index, focusing on three primary aspects: Coronavirus State Actions, Political Events, and Festivals/Entertainment events. Our objective is to shed light on the complex interplay

of these factors and their potential impact on the transmission of infectious diseases, with the goal of informing future decision-making processes and guiding public health policy.

To begin, we will examine the various state-level actions and policies enacted in response to the coronavirus pandemic. These actions encompass a wide range of measures, including the imposition and relaxation of social distancing protocols, the closure and reopening of public spaces and businesses, and the adjustment of transportation capacities. By assessing the effectiveness and potential consequences of these measures, we aim to identify the ways in which they may have contributed to the observed fluctuations in the transmission export index.

Next, we will explore the role of political events in influencing the transmission of infectious diseases. During the pandemic, numerous political gatherings, rallies, and protests have taken place, often attracting large crowds and creating environments conducive to disease transmission. By examining the specific contexts and circumstances surrounding these events, we hope to better understand the extent to which they may have affected the transmission export index and the broader public health landscape.

Finally, we will investigate the impact of festivals and entertainment events on disease transmission dynamics. These gatherings, which frequently involve large numbers of attendees in close proximity, have the potential to serve as major drivers of infectious disease spread. We will consider various factors, such as event size, location, and duration, as well as the implementation of preventive measures, in order to assess the potential influence of festivals and entertainment events on the transmission export index.

Through a comprehensive analysis of these three aspects, we aim to provide a nuanced understanding of the factors contributing to the increase in the transmission export index during the coronavirus pandemic. This in-depth exploration will not only shed light on the complex

dynamics at play but also serve as a valuable resource for policymakers and public health officials as they navigate the ongoing challenges posed by infectious diseases.

5.3.2.1 State Actions in Response to Transmission Export Index Increases

Table 7 summaries the top 2 states with the most-increased transmission export index for each month and their related state-level actions and policies that could potentially lead to an increase in the transmission export index of infectious diseases. These actions can be grouped into several categories: relaxation of social distancing measures, reopening of public spaces and businesses, expansion of transportation capacity, and easing of regulations for certain industries.

Firstly, some actions, such as the closure of state parks and forests (Executive Order 118) and New York City playgrounds, initially aimed to strengthen social distancing measures. However, the subsequent reopening of county beach parks in the County of Hawaii and the approval of businesses and operations on O'ahu represent a relaxation of these measures, which could potentially contribute to an increased transmission export index.

Similarly, the expansion of allowable outdoor recreational activities (Executive Order 20-38) and the resumption of contact practices and competitions in outdoor settings for organized sports (Executive Order No. 168) may also lead to increased transmission rates. These activities may involve close contact between individuals and a higher likelihood of disease transmission.

Another factor that could contribute to the increase of the transmission export index is the reopening of public spaces, such as allowing food trucks to operate at highway rest stops in Minnesota (Executive Order 20-49). This may encourage people to gather in these locations, increasing the possibility of disease transmission.

In the realm of transportation, the lifting of 50% capacity limits on NJ TRANSIT and private-carrier buses, trains, and light rail vehicles (Executive Order No. 165) could lead to increased

transmission risks. As more individuals use public transportation, the likelihood of close contact between passengers and subsequent disease transmission may rise.

Moreover, easing regulations for certain industries, such as providing emergency relief from regulations for motor carriers and drivers operating in Minnesota (Executive Order 20-80), could potentially contribute to the increase of the transmission export index. This might result from increased movement and interactions between individuals in these industries.

On the other hand, some state-level actions may help mitigate transmission risks. For example, the distribution of more than 4 million masks to businesses, their customers, and those who are unable to afford or easily obtain one (July 29, 2020) promotes the use of face coverings, which can reduce disease transmission.

In conclusion, while some state-level actions and policies have been implemented to control the spread of infectious diseases, others may inadvertently contribute to an increase in the transmission export index. These actions include the relaxation of social distancing measures, reopening of public spaces and businesses, expansion of transportation capacity, and easing of regulations for certain industries. To effectively manage and mitigate disease transmission risks, it is crucial for state authorities to balance the need for economic recovery with public health considerations.

Table 7: Major state actions in response to transmission export index increases

Date	State	Change of Transmission Export index (beta*mu)	Coronavirus State Actions
0320-0415	New Jersey	1.931885	April 7, 2020 - State & County Parks, Statewide - Executive Order 118 announced, closing all parks and forests to enforce social distancing measures.
0320-0415	New York	1.561676	
0415-0515	Minnesota	1.610151	April 1, 2020 - New York City, New York - Governor announced the closure of all playgrounds due to insufficient social distancing compliance. April 3, 2020 - New York State - Governor introduced a

0415-0515	Wisconsin	1.239867	website with daily updates on the state's comprehensive coronavirus testing data. April 20, 2020 - Wisconsin - Health Services Secretary-designee issued an Emergency Order for the Badger Bounce Back reopening plan. April 27, 2020 - Wisconsin - Governor signed Emergency Order #34, expanding operations for essential businesses and permitting curbside drop-off for nonessential businesses.
0515-0615	Hawaii	0.2480408	May 19, 2020 - County of Hawaii - Governor approved reopening of county beach parks island-wide with social distancing restrictions. May 27, 2020 - Oahu, Hawaii - Governor approved Mayor's proposal to reopen more businesses and operations under safety guidelines.
0515-0615	Nevada	0.1858456	May 26, 2020 – Nevada -The Governor announced that Nevada is ready to move into Phase 2 of the state's Nevada United: Roadmap to Recovery reopening plan on Friday, May 29.
0615-0715	New Jersey	2.86914	June 19, 2020 - Statewide - Administration announced outdoor visits for long-term care facility residents starting June 21. July 13, 2020 - New Jersey - Executive Order No. 165 signed, lifting 50% capacity limits on public transit and requiring carriers to limit vehicles to maximum seated capacity.
0615-0715	Minnesota	1.196351	July 13, 2020 - Statewide - Executive Order 20-78 signed, extending the COVID-19 peacetime emergency. July 14, 2020 - Statewide - Governor announced $100 million housing assistance program, funded by CARES Act, to prevent evictions, homelessness, and maintain housing stability.
0715-0815	New Jersey	0.48593	July 20, 2020 - New Jersey - Executive Order No. 168 signed, permitting resumption of outdoor contact practices and competitions for high-risk sports.
0715-0815	Minnesota	0.4159602	July 17, 2020 - Minnesota - Executive Order 20-80 signed, extending provisions in Executive Order 20-76 for emergency relief for motor carriers and drivers transporting livestock. July 29, 2020 - Statewide - Governor highlighted distribution of over 4 million masks to businesses, customers, and those unable to afford or obtain a mask.

5.3.2.2 Major Political Events in Response to Transmission Export Index Increases

During the study period from March 15, 2020 to Oct, several major events occurred in the top 2 states with the most-increased transmission export index, summarized in Table 8, which may have influenced the transmission of the virus.

1. May 26 George Floyd Protests:

Major protests began in the Minneapolis–Saint Paul area following the murder of George Floyd. These protests quickly spread to other cities across the United States and around the world. Large gatherings of protesters, often in close proximity and sometimes without masks or face coverings, created an environment conducive to the transmission of COVID-19. The risk of transmission increased due to the difficulty of maintaining physical distancing and proper hygiene practices during the protests. Additionally, law enforcement's use of tear gas and other crowd control measures could have exacerbated respiratory issues, further contributing to the spread of the virus.

2. July 7 Primary Elections in New Jersey:

The primary elections in New Jersey, rescheduled from June 2, also posed a potential risk for COVID-19 transmission. The act of voting typically involves people gathering in polling stations, standing in lines, and touching shared surfaces such as voting machines and pens. Although election officials implemented various safety measures, such as social distancing, providing hand sanitizer, and encouraging the use of face coverings, the risk of transmission could not be entirely eliminated. Moreover, some voters might have been discouraged from participating due to fear of infection, impacting overall voter turnout.

3. June 12 Minneapolis City Council Vote:

The Minneapolis City Council's vote to disband the police department and replace it with a "community" safety department was a significant political event. Although the vote itself likely had minimal direct impact on the transmission of COVID-19, the associated public meetings, discussions, and debates could have contributed to the spread. Such gatherings often involve people in close proximity, speaking passionately and potentially projecting respiratory droplets.

The risk of transmission would be higher if these gatherings took place indoors or if proper safety measures, such as wearing masks and maintaining physical distance, were not followed.

In summary, all three of these events had the potential to influence the transmission of COVID-19. The George Floyd protests, in particular, posed a significant risk due to the large gatherings and close contact between participants. The primary elections and the Minneapolis City Council vote also presented potential risks, although safety measures were likely implemented to mitigate the spread. These events highlight the challenges of balancing essential social and political activities with public health concerns during a pandemic.

Table 8: Major political events in response to transmission export index increases

Date	State	Change of Transmission attack index (beta*mu)	Major Political Events
0320-0415	New Jersey	1.931885	
0320-0415	New York	0.561676	
0415-0515	Minnesota	1.610151	May 26 - Minneapolis-Saint Paul, Minnesota - Major protests begin in response to George Floyd's murder.
0415-0515	Wisconsin	1.239867	
0515-0615	Hawaii	0.2480408	
0515-0615	Nevada	0.1858456	
0615-0715	New Jersey	2.86914	July 7 - New Jersey - Primary elections held, rescheduled from June 2.
0615-0715	Minnesota	1.196351	June 12 - Minneapolis, Minnesota - City Council votes to disband Police Department and replace with a community safety department, but is prevented by city charter.
0715-0815	New Jersey	0.48593	
0715-0815	Minnesota	0.4159602	

5.3.2.3 Major Festivals/Entertainment Events in Response to Transmission Export Index Increases

The potential for infectious disease transmission at large-scale events and festivals is a significant concern, given the unique challenges these gatherings present. Several factors contribute to the risk of disease spread, including the close contact between attendees, indoor venues, shared surfaces and equipment, food and beverage consumption, travel and accommodation arrangements, and insufficient hygiene practices. In the subsequent analysis, we will delve into these factors comprehensively and offer suggestions to reduce the risks linked to festivals and entertainment events in the top two states with the highest monthly increase in transmission export indices, as presented in

Table 9.

Large-scale events and festivals, such as CineKink NYC, Queens World Film Festival, Tribeca Film Festival, Roots and Bluegrass Music Festival, Lakes Jam, Stone Arch Bridge Festival, and Uptown Art Fair, are characterized by the congregation of vast crowds, leading to increased close contact between attendees. This proximity can facilitate the transmission of infectious diseases, particularly those that are airborne or spread through respiratory droplets. Notably, events held in indoor venues, like the Roots and Bluegrass Music Festival and CineKink NYC, pose an elevated risk of disease transmission due to limited air circulation and confined spaces.

Another factor that contributes to disease transmission at these events is the presence of shared surfaces and equipment. For example, events like Ecofest, where visitors interact with exhibits, or the Jersey Surf Film Festival, which offers surf lessons using shared equipment, may expose attendees to contaminated surfaces. Food and beverage consumption at events featuring tastings, such as Restaurant Week La Crosse or the Roots and Bluegrass Music Festival, can also increase

the risk of transmission due to shared utensils, plates, or cups, as well as close contact during food preparation and serving.

Furthermore, the travel and accommodation arrangements associated with these events can contribute to the spread of infectious diseases. Attendees often travel from different regions or countries, potentially introducing new disease threats to the event location. Shared accommodations, such as hotels or hostels, can further facilitate disease transmission. Compounding these factors, insufficient hygiene practices at large events can exacerbate the risk of disease spread. Maintaining proper hygiene can be challenging, especially in areas like restrooms, food service stations, or communal spaces where inadequate handwashing or sanitizing facilities may be present.

To mitigate the risk of disease transmission during such events, both organizers and attendees should consider implementing preventive measures. These measures may include promoting proper hand hygiene, providing sanitizing stations, enforcing physical distancing, requiring masks or face coverings, implementing health screenings, and ensuring adequate ventilation in indoor venues. Additionally, event organizers should stay informed about emerging disease threats and collaborate with local health authorities to make informed decisions regarding event planning and execution. By adopting these strategies, the risk of disease transmission at large-scale events and festivals can be significantly reduced.

Table 9: Major festivals/entertainment events in response to transmission export index increases

Date	State	Change of Transmission Export index (beta*mu)	Festivals /Entertainment events
0320-0415	New Jersey	1.931885	

0320-0415	New York	0.561676	March 18-22, 2020 - New York City, NY - CineKink NYC, a four-day film event celebrating diverse sexuality with movies, panel discussions, and parties. March 19-29, 2020 - Queens, NY - Queens World Film Festival, showcasing innovative films by maverick filmmakers from around the world. April 4-5, 2020 - New York City, NY - Ecofest, a free event featuring alternative energy exhibits, green vehicles, food, and entertainment in Times Square. April 15-26, 2020 - Lower Manhattan, NY - Tribeca Film Festival, offering movie screenings, celebrity talks, and exclusive content.
0415-0515	Minnesota	1.610151	
0415-0515	Wisconsin	1.239867	April 17-19, 2020 - Wisconsin - Roots and Bluegrass Music Festival, a free three-day indoor event with regional and local bands, workshops, and tastings. April 20-26, 2020 - La Crosse, WI - Restaurant Week La Crosse, a week-long food festival celebrating local restaurants and eateries.
0515-0615	Hawaii	0.2480408	
0515-0615	Nevada	0.1858456	
0615-0715	New Jersey	2.86914	June 19-20 - Mont Grantez - Sunset Concerts, summer music event with pop, blues, and jazz in a picturesque setting. July - New Jersey - Jersey Surf Film Festival, celebrating surfing with outdoor film screenings, workshops, surf lessons, and talks (2020 festival unconfirmed).
0615-0715	Minnesota	1.196351	June 25-27 - Minnesota - Lakes Jam, featuring two days of country music and a day of rock. June 19-21 - Minneapolis, MN - Stone Arch Bridge Festival, a three-day event celebrating art, food, and live music over Father's Day weekend.
0815-0915	New Jersey	0.48593	July 18 - New Jersey - Wonky Town, a post-apocalyptic themed one-day music festival with immersive experiences.
0915-1015	Minnesota	0.4159602	August 7-9 - Uptown, Minnesota - Uptown Art Fair, a juried arts festival celebrating the Uptown community with over 380,000 visitors.

5.3.3 Causal Analysis for Decrement of Transmission Export Index

According to Table 10 about the top 2 states with most decrement in transmission export index every month, a variety of state-level actions and policies have been implemented to curb the transmission of infectious diseases and reduce the transmission export index. Some common

strategies include promoting public health guidance, enacting travel restrictions, and implementing social distancing measures. The following paragraphs provide a summary and discussion of these policies.

One common policy implemented by state governments is urging residents to follow guidance from health authorities such as the Centers for Disease Control and Prevention (CDC) and state health departments. By promoting adherence to these guidelines, states aim to minimize the spread of infectious diseases within their jurisdictions. Examples of such policies include encouraging residents to stay at home as much as possible and extending stay-at-home orders for specified durations.

Another approach taken by states is the implementation of social distancing measures to limit the spread of infections. Some of these measures include suspending in-person voting for elections, issuing shelter-in-place orders for specific counties with increased cases, closing schools for the remainder of the academic year, and limiting social, community, recreational, leisure, and sporting gatherings. In some cases, states have permitted the reopening of certain businesses and establishments, such as salons, barbershops, massage and tattoo parlors, restaurants, and fitness centers, but only with strict public health measures in place.

States have also introduced phased plans for reopening their economies in a gradual and safe manner. These plans typically involve assessing the reopening of businesses and activities based on the level of disease transmission and essential classification. For example, some states have adopted multi-stage approaches to reopening, with each stage permitting a specific set of businesses or activities to resume operations under certain conditions. Additionally, some states have authorized businesses to deny entry to individuals who do not wear masks or face coverings, further emphasizing the importance of personal protective measures.

Travel restrictions have been another key policy employed by states to reduce the transmission export index. These restrictions often involve mandatory quarantines for travelers entering the state or requiring travelers to present proof of a negative COVID-19 test prior to their arrival. In some cases, states have delayed the implementation of pre-travel testing programs or extended the duration of mandatory quarantines for incoming travelers.

In conclusion, state governments have adopted a range of policies to mitigate the transmission of infectious diseases and lower the transmission export index. These strategies include promoting public health guidance, implementing social distancing measures, introducing phased plans for economic reopening, and imposing travel restrictions. By adopting these measures, states aim to protect their residents and limit the spread of infections both within their borders and across the nation.

Table 10: Major state actions in response to transmission export index decreases

Date	State	Change of Transmission attack index (beta*mu)	Coronavirus State Actions
0320-0415	Wisconsin	-4.1314	March 20, 2020 - Wisconsin: The Governor urged residents to follow CDC and state health department guidance to stay home as much as possible. April 6, 2020 - Wisconsin: The Governor signed an executive order postponing in-person voting for the April 7 election until June 9 and called for a special legislative session to address the election date.
0320-0415	Mississippi	-3.91885	March 31, 2020 - Mississippi: The Governor issued a shelter in place order for a county due to increased cases in the region. April 14, 2020 - Mississippi: The Governor announced that schools will remain closed for the rest of the school year.
0415-0515	Kansas	-0.855282	April 15, 2020 - Kansas: The Governor extended the stay-at-home order until May 1st. April 30, 2020 - Kansas: The Governor presented a detailed framework for gradually reopening the economy starting May 4, 2020, with Executive Order 20-29, lifting the statewide stay-home order in Executive Order 20-16. May 14, 2020 - Kansas: The Governor signed Executive Order 20-31, establishing a new "1.5" Phase effective May

			18, 2020, continuing reopening efforts with some restrictions to prevent community transmission of COVID-19.
0415-0515	Iowa	-0.849734	April 16, 2020 - Iowa: The Governor signed a proclamation continuing the State Public Health Emergency Declaration, requiring additional protective measures in Region 6 (Northeastern Iowa), including limiting social and recreational gatherings. April 19, 2020 - Iowa: The Governor announced that all schools will be closed for the remainder of the school year. May 13, 2020 - Iowa: The Governor signed a proclamation continuing the Public Health Disaster Emergency, allowing certain businesses to reopen with restrictions and extending the prohibition on gatherings of more than 10 people until 11:59 p.m. on May 27, 2020.
0515-0615	New Jersey	-4.353733	May 18, 2020 - New Jersey: The Governor unveiled a multi-stage approach for a responsible and strategic economic reopening and signed Executive Order No. 147, allowing certain outdoor activities at recreational businesses and community gardens with social distancing measures. June 1, 2020 - New Jersey: The Governor announced that the state will enter Stage Two on June 15, including outdoor dining for restaurants and indoor, non-essential retail.
0515-0615	New York	-2.264473	May 20, 2020 - New York: The Governor announced that religious gatherings of no more than 10 people will be allowed starting May 21. May 21, 2020 - New York The Governor announced that summer school will be conducted through distance learning and that meal programs and childcare services for essential employees will continue. May 28, 2020 - New York: The Governor issued an executive order allowing businesses to deny entry to individuals not wearing masks or face coverings.
0615-0715	Hawaii	-0.1641658	June 24, 2020 - Hawai‘i: The Governor announced a pre-travel testing program for out-of-state travelers starting Aug. 1 and approved the proposal to allow singing and playing of wind instruments at indoor and outdoor restaurants/bars with restrictions. July 13, 2020 - Hawai‘i: The Governor delayed the launch of the pre-travel testing program to Sept. 1, extending the mandatory 14-day quarantine for travelers entering the state until then.

0615-0715	Arizona	-0.1007183	June 29, 2020 - Arizona: The Governor issued an executive order prohibiting large gatherings, ceasing new special event licenses, and pausing operations of bars, gyms, movie theaters, waterparks, and tubing rentals, and delayed in-person learning until August 17, 2020. July 9, 2020 - Arizona: The Governor issued an executive order requiring restaurants with indoor seating to operate at less than 50% percent capacity.
0715-0815	Hawaii	-3.147501	July 17, 2020 - Hawai‘i: The Governor signed the 10th Emergency Proclamation, keeping the mandatory 14-day quarantine in effect for travelers entering the state, and travelers will continue to undergo mandatory screening at airports. July 20, 2020 - Hawai‘i: The Governor confirmed the state's plans to move ahead with school reopening for students on August 4. July 29, 2020 - Hawai‘i: The Governor announced plans to reinstate some of the measures relaxed in recent weeks to combat COVID-19 in Hawaii.
0715-0815	Arizona	-1.167903	July 23, 2020 - Arizona: The Governor extended an executive order pausing operations on gyms, bars, nightclubs, movie theaters, water parks, and tubing and announced a statewide campaign promoting mask use and other precautions. July 30, 2020 - Arizona: The Governor extended a statewide mask order until August 31, mandating masks in schools and colleges for employees and students in second grade and above.

5.4 Conclusion and Discussion

The investigation of disease transmission through spatial interaction, particularly state-level travel, has provided valuable insights into the complex dynamics that govern the spread of infectious diseases. By developing a multi-regional dynamic model with spatial interaction, we have been able to accurately capture historical transmission patterns and evaluate the impact of long-distance travel on disease transmission. This understanding is crucial for developing effective prevention and control strategies for future outbreaks.

The introduction of the transmission export index, which combines local transmissibility and the potential for infectious travelers to spread diseases to new regions, has proven to be an important tool for assessing the risk of disease transmission between regions. By identifying high-risk areas,

appropriate interventions, such as travel restrictions or quarantine measures, can be put in place to mitigate disease spread and protect public health.

Furthermore, our causal analysis of factors influencing the transmission export index has highlighted the importance of considering a wide range of influences, including state-level actions, political events, and festivals/entertainment events. By examining the interplay between these factors and their potential impact on disease transmission, we can better inform future decision-making processes and guide public health policy.

Additionally, our analysis of state-level policies aimed at reducing the transmission export index has demonstrated the effectiveness of various measures, such as promoting public health guidance, implementing social distancing measures, introducing phased plans for economic reopening, and imposing travel restrictions. These strategies play a vital role in protecting residents and limiting the spread of infections both within individual states and across the nation.

In conclusion, this chapter has shed light on the critical role of spatial interaction in disease transmission and the importance of understanding these dynamics for effective prevention and control efforts. The methods and findings presented here can serve as a foundation for future research, policy development, and public health interventions aimed at mitigating the impact of infectious diseases on a regional, national, and global scale.

Chapter 6 Integration of Dynamic Modeling with Vaccination Reallocation

6.1 Introduction

Vaccination is one of the most effective public health interventions for controlling the spread of infectious diseases. Vaccines work by inducing immunity to a pathogen, thereby reducing the likelihood of transmission. Vaccination can have a significant impact on disease transmission. When a large proportion of the population is vaccinated, the likelihood of transmission decreases, as the pathogen has fewer hosts in which to replicate and spread. This phenomenon is known as herd immunity. The level of herd immunity required to control the spread of a disease varies depending on the disease and the vaccine efficacy. In general, a higher proportion of the population needs to be vaccinated for diseases with higher transmissibility.

The introduction of vaccines has led to a renewed interest in transmission modeling in epidemiology. Modeling the impact of vaccination on disease transmission is important for understanding the effectiveness of vaccination programs and designing effective vaccine distribution policies.

Although the primary goal of vaccination is to provide immunity to individuals against a specific infectious agent, it is crucial to ensure that the vaccination campaign is designed to maximize the benefits of vaccination while minimizing the risks and limitations. Thus, designing effective vaccine distribution policies is essential for successful vaccination programs in controlling the transmission of an epidemic. One crucial factor that must be considered is the age of the population targeted for vaccination. Vaccines typically have different efficacy rates and side effects in different age groups, and this must be taken into account when designing a distribution policy. For

example, vaccines such as the Pfizer-BioNTech and Moderna vaccines have been found to be highly effective in preventing infection and severe disease in all age groups, including adolescents, adults, and elders. On the other hand, the Johnson & Johnson vaccine has shown slightly lower efficacy rates in preventing infection, but still high efficacy in preventing severe disease and death. When it comes to vaccine distribution, it is essential to prioritize those who are at the highest risk of developing severe disease or dying from the virus. This includes the elderly, those with underlying medical conditions, and healthcare workers. By vaccinating these high-risk groups first, the transmission of the virus can be significantly reduced. The Centers for Disease Control and Prevention (CDC) has recommended that frontline essential workers, including those in education, transportation, and food service, also be prioritized for vaccination due to their increased risk of exposure to the virus.

Another critical factor in vaccine distribution policy is the availability of vaccines. Limited vaccine supply can make it difficult to prioritize groups effectively, and it may be necessary to implement a phased approach. In such a situation, it may be appropriate to prioritize those at the highest risk of severe disease, including the elderly and those with underlying medical conditions, followed by essential workers and then the general population.

In this chapter, we will delve deeper into the development of a transmission model that takes into account the varying age groups and vaccination statuses of individuals, utilizing dynamic modeling with time-varying transmission parameters. This approach enables us to capture the complex interactions between different age groups and vaccination coverage, providing a more accurate representation of disease transmission dynamics.

To validate the accuracy and reliability of our model, we will compare its predictions against historical case and death data from all 50 states. This comparison serves as an essential benchmark,

ensuring that our model is capable of capturing the true dynamics of disease transmission and providing reliable insights for public health decision-making.

Building upon the age-structured dynamic model that incorporates vaccination, we will then propose a novel method for optimizing vaccine allocation across different regions. This method takes into consideration the varying transmission severity and population structures among different areas, as well as the limited resources available for vaccine distribution. By dynamically optimizing the allocation of vaccines, we aim to minimize the overall impact of the disease while maximizing the efficient use of available resources.

Finally, we will analyze the implications of different vaccine allocation policies under various scenarios of vaccine availability. This analysis will provide valuable insights into how different strategies perform under a range of circumstances, informing decision-makers on the most effective approaches for managing the disease and mitigating its impact on public health. Through the development and application of our age-structured dynamic model with vaccination, we hope to contribute to a better understanding of disease transmission dynamics and inform evidence-based decision-making for vaccine allocation and distribution strategies. This, in turn, will ultimately help minimize the adverse effects of infectious diseases on populations and ensure a more efficient and equitable use of limited resources.

6.2 Age-Structured Dynamic Modeling with Vaccination

6.2.1 Age-Structured Transmission Data and Vaccine Data

Age-structured case and death data provide crucial insights into the COVID-19 pandemic's impact on different age groups. The age-specific case data reveals how the infection rate varies among different age brackets. Early in the pandemic, it became clear that older populations were more susceptible to severe illness and death from COVID-19. According to the CDC, individuals aged

65 years and older accounted for the majority of COVID-19-related deaths, while younger age groups experienced significantly fewer fatalities. This age distribution also influenced the policy recommendations for vaccination priority, with older individuals and those with underlying health conditions receiving vaccinations first.

Vaccine data is another critical component in understanding the progression of the pandemic and the effectiveness of public health interventions. The CDC tracks vaccine distribution, administration, and coverage across different age groups, geographic regions, and demographic categories. This information allows for the identification of disparities in vaccine access and uptake, as well as areas that may require targeted outreach and education efforts. Age-stratified vaccine data also allows for the assessment of vaccine effectiveness in preventing severe illness and death, especially among high-risk age groups.

In our research about vaccine distribution, we will use CDC as the primary source of COVID-19 data in the United States. They collect and publish data on cases, deaths, and vaccinations through various channels, including state and local health departments, laboratories, and healthcare providers. One key resource provided by the CDC is the COVID Data Tracker (https://covid.cdc.gov/covid-data-tracker), an interactive web-based platform that presents up-to-date information on cases, deaths, and vaccinations, as well as other relevant metrics such as testing, hospitalizations, and variant tracking. The CDC continually updates these data sources to provide the most accurate and comprehensive information possible, helping inform policy decisions and public health interventions.

Although the CDC has been diligent in tracking and reporting COVID-19 vaccination data, there was a period between December 16, 2020, and March 4, 2021, when age-structured vaccination data was not readily available. In order to estimate the daily administered vaccinations in each age

group for every state during this period, we utilized a data-driven approach. First, the ratio of eligible persons in each age group to the total number of eligible persons was calculated based on the vaccination policies in place at the time, which primarily prioritized older adults and individuals with underlying health conditions. Next, the daily administered vaccinations for each age group were estimated by multiplying the total amount of daily administered vaccines by the calculated ratio for each age group.

This approach provides a valuable approximation of age-specific vaccination trends during this critical period when vaccine distribution was in its early stages. It is essential to acknowledge that these estimations come with a degree of uncertainty, as they are based on the assumption that the proportion of eligible individuals among different age groups directly corresponds to the actual vaccine uptake in each age group. Nevertheless, this method offers a useful framework for analyzing vaccination patterns in the absence of complete age-structured data from the CDC during this specific time frame.

6.2.2 Model Structure and Parameter Estimation

In order to better understand the impact of vaccination on COVID-19 transmission and mortality, we have expanded upon dynamic modeling with time-varying transmission and fatality rates in Chapter 4 while taking into account different age groups and the effect of vaccination. We divided the whole population into three age groups: 0-17, 18-64, and 65 plus. Each age group has different transmission and death rates, reflecting the observed disparities in the susceptibility to the virus and the severity of the disease.

To capture these differences, we introduced two scalars, ω_i and τ_i, for each age group i. The scalar ω_i represents the difference of transmission rate between age group i and the reference age group 2 (18-64 years), while τ_i denotes the difference in fatality rate between age group i and the

reference age group 3 (65+ years). These scalars allow us to quantify the relative transmission and mortality risk for each age group in comparison to the reference groups.

Next, we incorporated the effect of vaccination by introducing a vaccine compartment (V_i) for each age group. The vaccination process reduces the number of susceptible individuals in each age group, with the reduction being proportional to the number of vaccinated individuals and the effectiveness of the vaccine (θ). By considering the vaccine's effectiveness, we can account for the fact that vaccinated individuals may still be at risk of infection, albeit at a lower level than those who are unvaccinated. In our modeling approach that incorporates the effect of vaccines, we have disregarded the vaccine's impact on reducing the death rate among vaccinated individuals. This is primarily due to the absence of specific mortality data tracking the number of vaccinated individuals who succumbed to the disease across different age groups. Additionally, we have not factored in inter-state transportation within our model due to the unavailability of age-specific mobility data. Taking these factors into account, the system of equations of the proposed SEIRD_V model is given by Equation (16):

$$\begin{aligned}
\frac{\partial S_i(t)}{\partial t} &= -\omega_i \cdot \beta(t) \cdot I_i(t) \cdot \frac{S_i(t) - \theta \cdot V_i(t)}{N} \\
\frac{\partial E_i(t)}{\partial t} &= \omega_i \cdot \beta(t) \cdot I_i(t) \cdot \frac{S_i(t) - \theta \cdot V_i(t)}{N} - \sigma \cdot E_i(t) \\
\frac{\partial I_i(t)}{\partial t} &= \sigma \cdot E_i(t) - \big(1 - \tau_i \cdot \alpha(t)\big) \cdot \gamma I_i(t) - \tau_i \cdot \alpha(t) \cdot \rho \cdot I_i(t) \\
\frac{\partial R_i(t)}{\partial t} &= \tau_i \cdot \big(1 - \alpha(t)\big) \cdot \gamma \cdot I_i(t) \\
\frac{\partial D_i(t)}{\partial t} &= \tau_i \cdot \alpha(t) \cdot \rho \cdot I_i(t)
\end{aligned} \tag{16}$$

where:

$S_i(t)$: the number of susceptible individuals in age group i over time.

$E_i(t)$: the number of exposed individuals in age group i over time.

$I_i(t)$: the number of infectious individuals in age group i over time.

$R_i(t)$: the number of recovered individuals in age group i over time.

$D_i(t)$: the number of dead individuals in age group i over time.

$V_i(t)$: the number of individuals vaccinated in age group i over time.

N: the total population size.

β(t): the effective contact rate, a measure of how many people to whom an infected person can transmit the disease at time t.

α(t): the fraction of infectious individuals detected and isolated at time t.

γ: the recovery rate of infected individuals.

δ: the rate at which exposed individuals become infectious.

ρ: the fatality rate among infected individuals.

ω_i: scalars representing the difference in transmission rate between age groups i with respect to age group 2.

τ_i: scalars representing the difference in fatality rate between age groups i with respect to age group 3.

The model we developed aims to provide a more comprehensive understanding of the interplay between age-specific transmission dynamics, vaccine rollout, and disease outcomes. By considering the heterogeneous nature of COVID-19 transmission and death rates across different age groups and accounting for the impact of vaccination, we can generate more accurate and nuanced predictions about the pandemic's progression. This, in turn, can help inform public health policies and interventions tailored to the specific needs of each age group, ultimately contributing to more effective control of the pandemic.

6.2.3 Fitting Results of Age-Structured Dynamic Modeling with Vaccination

Figure 20 and Figure 21 summarize the model fitting accuracy of the transmission model with vaccination for three age groups (0-17, 18-64, and 65+) across all 50 states in the United States

from December 16, 2020, to June 30, 2021. The model's accuracy is evaluated by the relative root mean square error (RRMSE) defined in Equation (12) for both the number of COVID-19 cases and deaths within each age group for each state. The results provide valuable insights into the model's performance and the effectiveness of incorporating vaccination data into the transmission model.

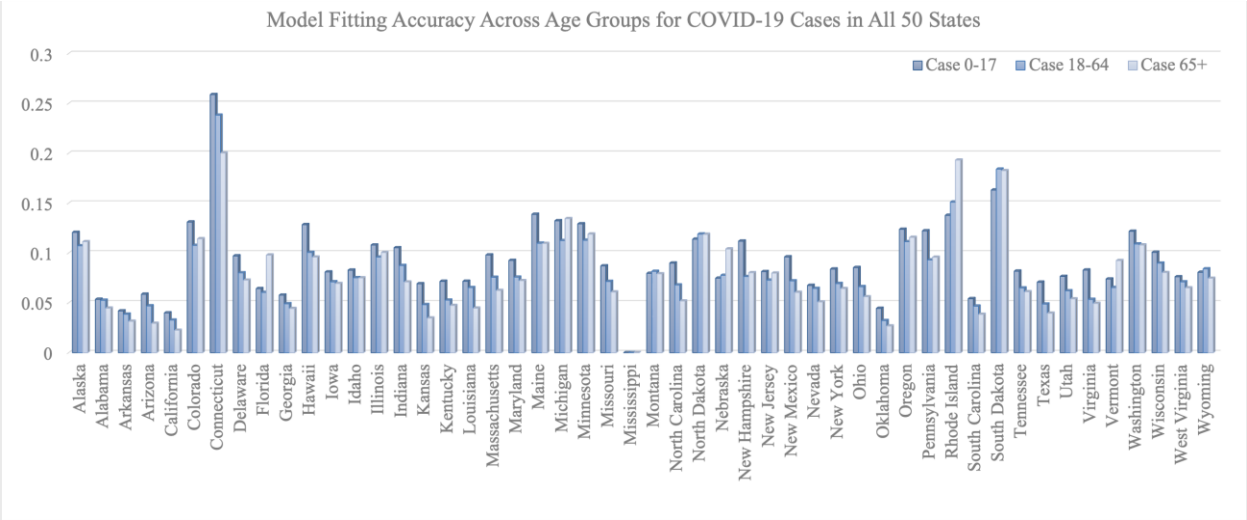

Figure 20: Model fitting accuracy across age groups for covid-19 cases in all 50 states

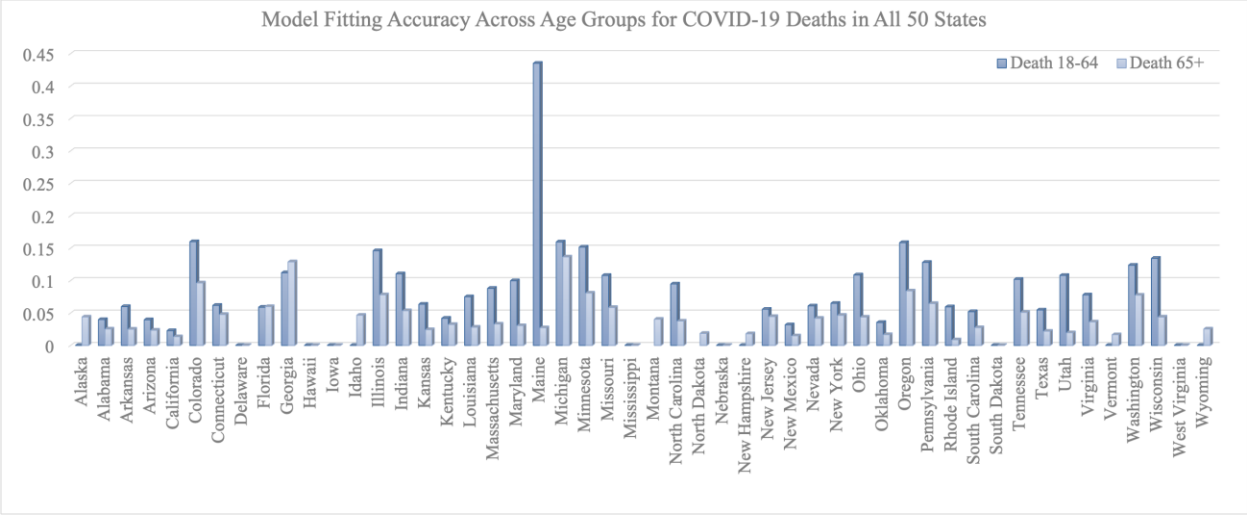

Figure 21: Model fitting accuracy across age groups for covid-19 deaths in all 50 states

The model fitting accuracy varies across states and age groups, which may be attributed to factors such as differences in state-level vaccination rates, adherence to public health guidelines, population density, and other regional factors influencing transmission and death rates. On average,

the fitting accuracy for the number of cases is 0.092 for the 0-17 age group, 0.080 for the 18-64 age group, and 0.078 for the 65+ age group. The average fitting accuracy for the number of deaths is 0.009 for the 0-17 age group, 0.073 for the 18-64 age group, and 0.038 for the 65+ age group. The model generally exhibits higher fitting accuracy for the number of cases than for the number of deaths. This may be because the number of cases is typically higher and more consistently reported than the number of deaths, making it easier for the model to fit case data. Furthermore, the relatively low fitting accuracy for the number of deaths in the 0-17 age group could be due to the rarity of COVID-19-related deaths in this demographic, resulting in fewer data points and greater uncertainty in the model.

It is notable that some states, such as Mississippi, have zero fitting accuracy for all age groups in both cases and deaths. This may be caused by a lack of available data, inconsistencies in reporting, or potential issues with the model's assumptions for that specific state. Further investigation would be required to determine the cause of these discrepancies and improve the model's performance.

Moreover, the lower fitting accuracy for certain states may be attributed to the discrete daily variation of administered vaccines. Daily fluctuations in vaccination numbers can add complexity to the modeling process, making it more challenging for the model to generate smooth transmission rate and death rate functions that accurately capture historical trends. This issue can be particularly pronounced in states with inconsistent vaccination rollouts or disruptions due to supply chain issues, logistical challenges, or changes in vaccine eligibility criteria. In such cases, the model may struggle to accurately account for the impacts of these fluctuations on overall transmission and death rates. The daily variation in administered vaccines can lead to inconsistencies in the model's predictions, which may contribute to lower fitting accuracy observed in some states.

In summary, the model fitting results demonstrate that incorporating vaccination data into a transmission model can provide reasonably accurate estimates of COVID-19 cases and deaths across different age groups and states. The model's varying accuracy across states highlights the importance of considering regional factors when evaluating its performance and potential improvements. This analysis also emphasizes the need for continued data collection and reporting to refine the model and better understand the impact of vaccination on the pandemic's trajectory. Despite its limitations, the transmission model with vaccination data offers valuable insights into the progression of the COVID-19 pandemic in the United States, particularly in terms of age-specific trends. By accounting for the effects of vaccination and different age group transmission dynamics, this model can help inform public health policies and interventions that are tailored to the specific needs of each age group. This, in turn, can contribute to more effective control of the pandemic and better health outcomes for all.

6.3 Vaccine Allocation Optimization with Dynamic Transmission Pattern

The COVID-19 pandemic has demonstrated the importance of efficient and equitable vaccine allocation strategies to control the spread of the virus and reduce morbidity and mortality rates. In this section, we propose a dynamic optimization framework to allocate vaccines among 50 states in the United States, taking into account the transmission patterns and the impact of vaccination on disease transmission and death rates.

Previously, we developed a transmission model incorporating vaccination data to provide reasonably accurate estimates of COVID-19 cases and deaths across different age groups and states. The model considers age-structured case and death data, vaccine data, and time-varying transmission and death rates, accounting for the effects of vaccination on susceptible populations

in each age group. The fitting results demonstrated that the model effectively captured the historical trends of COVID-19 cases and deaths and the impact of vaccination on these trends.

Building upon this transmission model, we now aim to optimize vaccine allocation among the 50 states by solving an optimization problem. The objective function for the optimization problem is defined as the sum of the weighted case numbers and death numbers. The constraints are the bi-weekly available amount of vaccination for each state. For the vaccine allocation optimization problem, we consider the objective function and constraints in Equation (17):

$$
\begin{aligned}
&\min_{V_{i,t}} \sum_{i,t} w_1 \cdot Cases_{i,t} + w_2 \cdot Deaths_{i,t} \\
&\text{s.t.} \sum_{i} V_{i,t} \le Q_t \ \ \forall t \\
&\quad 0 \le V_{i,t} \le N_i \\
&\quad \frac{\partial S_i(t)}{\partial t} = -\omega_i \cdot \beta(t) \cdot I_i(t) \cdot \frac{S_i(t) - \theta \cdot V_i(t)}{N} \\
&\quad \frac{\partial E_i(t)}{\partial t} = \omega_i \cdot \beta(t) \cdot I_i(t) \cdot \frac{S_i(t) - \theta \cdot V_i(t)}{N} - \sigma \cdot E_i(t) \\
&\quad \frac{\partial I_i(t)}{\partial t} = \sigma \cdot E_i(t) - \big(1 - \tau_i \cdot \alpha(t)\big) \cdot \gamma I_i(t) - \tau_i \cdot \alpha(t) \cdot \rho \cdot I_i(t) \\
&\quad \frac{\partial R_i(t)}{\partial t} = \tau_i \cdot \big(1 - \alpha(t)\big) \cdot \gamma \cdot I_i(t) \\
&\quad \frac{\partial D_i(t)}{\partial t} = \tau_i \cdot \alpha(t) \cdot \rho \cdot I_i(t) \\
&\quad Cases_{i,t} = I_i(t) + R_i(t) + D_i(t) \\
&\quad Deaths_{i,t} = D_i(t)
\end{aligned}
\tag{17}
$$

Where $V_{i,t}$ refers to the vaccine number in region i on day t, Q_t refers to the total amount of available vaccine on day t, w_1 and w_2 are the weight the policymaker put on the case number and death number. The dynamic nature of this optimization problem lies in the fact that it considers the evolving transmission patterns and vaccination rates over time. By accounting for these dynamics, the optimization process can be adjusted as new data on the pandemic's progression and vaccination efforts becomes available, allowing for a more adaptive and responsive allocation strategy.

The optimization problem is highly nonlinear since the case number and death number are solved using implicit ordinary differential equations (ODEs) with time-varying transmission parameters and discrete vaccine numbers. The decision variables in the functions are the bi-weekly vaccine numbers allocated to each state. To solve the optimization problem, we employ the Sequential Least Squares Quadratic Programming (SLSQP) method [129], a gradient-based optimization algorithm. The SLSQP algorithm is well-suited for this problem because it can handle both equality and inequality constraints and is capable of solving nonlinear optimization problems with a large number of variables.

To illustrate the SLSQP solving process for vaccine allocation optimization, let's first consider the optimization problem, which aims to minimize the weighted sum of cases and deaths over a specific time horizon. The objective function consists of two components: the number of cases and the number of deaths. The decision variables in the optimization problem are the bi-weekly vaccine allocations for each state, subject to constraints on the total available vaccines and the maximum vaccination capacity of each state.

The SLSQP algorithm starts with an initial guess for the decision variables (i.e., the bi-weekly vaccine allocations) and iteratively updates these values to minimize the objective function. At each iteration, the algorithm computes the gradient of the objective function with respect to the decision variables, which is essential for updating the decision variables in the right direction.

In this context, the gradient computation is challenging due to the implicit nature of the objective function, which depends on the solution of ordinary differential equations (ODEs) describing the transmission dynamics. To calculate the case/death number in the objective functions and the gradient with respect to the decision variables, we must first solve the ODEs. Traditional third-party ODE solvers, such as the odeint function provided by the SciPy library, utilize the fourth-

order Runge-Kutta method to achieve a higher level of accuracy. This method approximates the daily increments with a sufficiently small step size. However, the transmission rate and death rate of the disease will remain constant within the same day, and the implicit formulation of the fourth-order Runge-Kutta method makes it challenging for the algorithm to find the derivatives with respect to the decision variables, potentially leading to gradient vanishment.

To address this issue, we utilize Euler's method, a first-order numerical method for solving ODEs, instead of the fourth-order Runge-Kutta method. By employing Euler's method, we can calculate the weekly case and death numbers with fixed decision variables (i.e. bi-weekly allocated vaccination for each state). This approach allows for a more straightforward computation of the gradient, avoiding the complexities associated with higher-order ODE solvers like the Runge-Kutta method.

Once the gradient is computed, the SLSQP algorithm updates the decision variables by moving in the direction of the negative gradient, which corresponds to the steepest descent in the objective function. The algorithm also takes into account the constraints on vaccine availability and state capacities, ensuring that the updated decision variables are feasible. This iterative process continues until the algorithm converges to a solution that minimizes the objective function, subject to the constraints.

The SLSQP-based optimization framework offers a systematic approach to determining the optimal distribution of vaccines among the 50 states, taking into account the dynamic nature of the transmission patterns, vaccine availability, and state capacities. By continuously updating the decision variables, the algorithm ensures that the vaccine allocation strategy remains aligned with the evolving pandemic landscape, ultimately leading to a more efficient and effective allocation of resources and better public health outcomes. The dynamic vaccine allocation optimization

framework proposed in this study allows for a more effective and equitable distribution of vaccines among the 50 states, considering the regional transmission patterns and the impact of vaccination on disease transmission and death rates. By dynamically adjusting the vaccine allocation based on the latest available data and evolving transmission patterns, this framework can help inform public health decision-making and guide effective pandemic response efforts at the national and state levels.

In conclusion, the dynamic optimization of vaccine allocation with transmission patterns provides a valuable tool for public health authorities and policymakers to make data-driven decisions on vaccine distribution. The use of Euler's method to solve the highly nonlinear optimization problem with implicit ODEs and the application of the SLSQP method for solving the optimization problem ensures that the algorithm can find the derivatives with respect to the decision variables, enabling an effective solution to the problem. By incorporating regional transmission patterns and the impact of vaccination on disease transmission and death rates, the proposed framework enables a more targeted and efficient allocation of vaccines among the 50 states.

6.4 Vaccine Allocation Policy under Different Scenarios

In this section, we will discuss the best vaccine distribution policy under different scenarios of vaccine availability. A thorough analysis of various vaccine distribution strategies is essential to identify the most effective approaches for allocating limited resources to minimize the impact of COVID-19 on public health. By studying the best vaccine distribution policy under different scenarios of vaccine availability, we aim to understand the implications of alternative vaccine allocation strategies, as well as to inform and improve future vaccine distribution efforts.

Studying the best vaccine distribution policy under different scenarios is essential to understand the implications of various allocation strategies and to identify the most effective approach to

controlling the spread of COVID-19. Considering different vaccine availability scenarios allows us to account for uncertainties in vaccine production and distribution, as well as potential changes in demand due to factors such as vaccine hesitancy or new variants. The objective of this analysis is to ensure that vaccines are allocated in a manner that minimizes the number of cases and deaths, while also maximizing the overall public health benefits.

By examining different vaccine distribution scenarios, we can explore the impact of prioritizing certain age groups or geographical regions, and assess the potential trade-offs between focusing on high-risk populations versus wider coverage. This information is invaluable for policymakers, public health officials, and other stakeholders involved in the decision-making process, as it enables them to make informed choices about vaccine distribution strategies that optimize resource allocation and ultimately save lives.

Furthermore, understanding the best vaccine distribution policy under various scenarios can inform future vaccination campaigns, not only for COVID-19 but also for other infectious diseases. Lessons learned from this analysis can contribute to the development of more robust vaccination strategies that can be adapted to different contexts and changing circumstances, thereby improving the overall effectiveness of public health interventions.

In the following part, we will consider several scenarios of vaccine availability, starting with the hypothetical situation of zero vaccine availability. This baseline scenario will enable us to evaluate the effectiveness of historical vaccine distribution efforts and inform our understanding of the potential benefits of optimized vaccine distribution policies. We will then explore other scenarios with varying levels of vaccine availability to identify the optimal vaccine distribution policy for each situation, ultimately leading to more efficient and effective distribution strategies that minimize the number of cases and deaths.

6.4.1 Healthcare Outcomes without Vaccination

First, let's consider the scenario with zero vaccine availability. This hypothetical situation allows us to understand the effectiveness of the historical vaccine distribution by comparing the case and death numbers under this scenario to the real historical data. By simulating the age-structured dynamic model with no vaccination, we can estimate the number of cases and deaths that would have occurred if no vaccines were distributed.

The results from this analysis show that the historical vaccine distribution has had a significant impact on reducing the spread of the virus and saving lives. In the absence of any vaccine distribution, the model estimates that there would have been an additional 1,827,631 cases and 9,180 deaths. These findings highlight the crucial role that vaccines have played in mitigating the severity of the pandemic and demonstrate the importance of an effective vaccine distribution strategy.

In order to further explore the best vaccine distribution policy, we can consider different scenarios of vaccine availability. For each scenario, the age-structured dynamic model with vaccination can be used to simulate the impact of various distribution strategies on the number of cases and deaths. By comparing the outcomes under different strategies, we can identify the optimal vaccine distribution policy for each level of vaccine availability.

The zero vaccine availability scenario serves as a baseline for understanding the effectiveness of historical vaccine distribution efforts. As we analyze other scenarios with varying vaccine availability, we can gain insights into how to optimize the allocation of vaccines across states and age groups, ultimately leading to more efficient and effective distribution strategies that minimize the number of cases and deaths. This valuable information can inform future vaccine distribution

policies, ensuring that limited resources are allocated in the most impactful way possible, contributing to better public health outcomes and the eventual control of the pandemic.

6.4.2 Vaccine Allocation Policy with Original Vaccine Availability

In the second scenario, we explore the optimal vaccine allocation policy among the 50 states to achieve better healthcare outcomes while considering the dynamic transmission patterns in each state. The current government strategy of distributing vaccines proportional to the eligible population size in each state can lead to an imbalance in vaccine allocation, with some states receiving a surplus while others face a shortage of this vital resource. To address this issue, we analyzed the best vaccine allocation policy under two different prioritizations: reducing the number of cases as much as possible and reducing the number of deaths as much as possible. These two priorities encompass the primary concerns of policymakers when managing the pandemic.

When focusing on case reduction, the optimal vaccine allocation policy suggests that a larger share of vaccines should be distributed to the younger age group (0-17 years). This is because this population is generally more active and has more frequent social interactions, leading to a higher potential for spreading the virus. Additionally, younger individuals may exhibit milder symptoms or be asymptomatic, making them more likely to unknowingly transmit the virus to others. By vaccinating this age group, the overall transmission rate within the population can be significantly reduced, ultimately lowering the total number of cases. In this prioritization, the focus is on reducing the spread of the virus, leading to an overall decrease in cases and, consequently, a lower number of associated deaths. According to the model results, this optimal allocation strategy could potentially reduce 2,042,312 cases and 1,796 deaths.

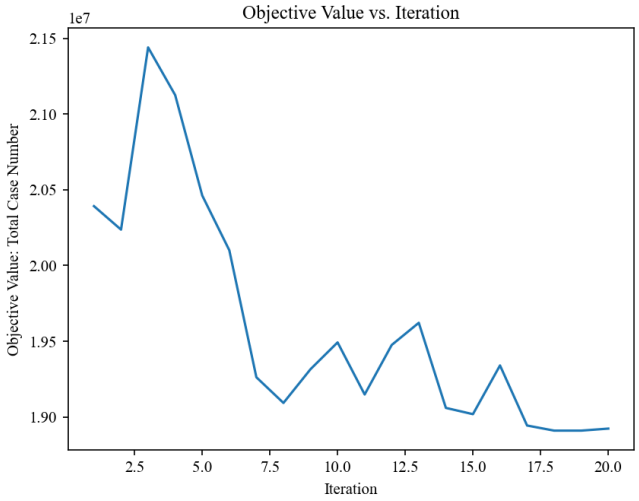

Figure 22: Training process of case-prioritized vaccine optimization with original vaccine availability

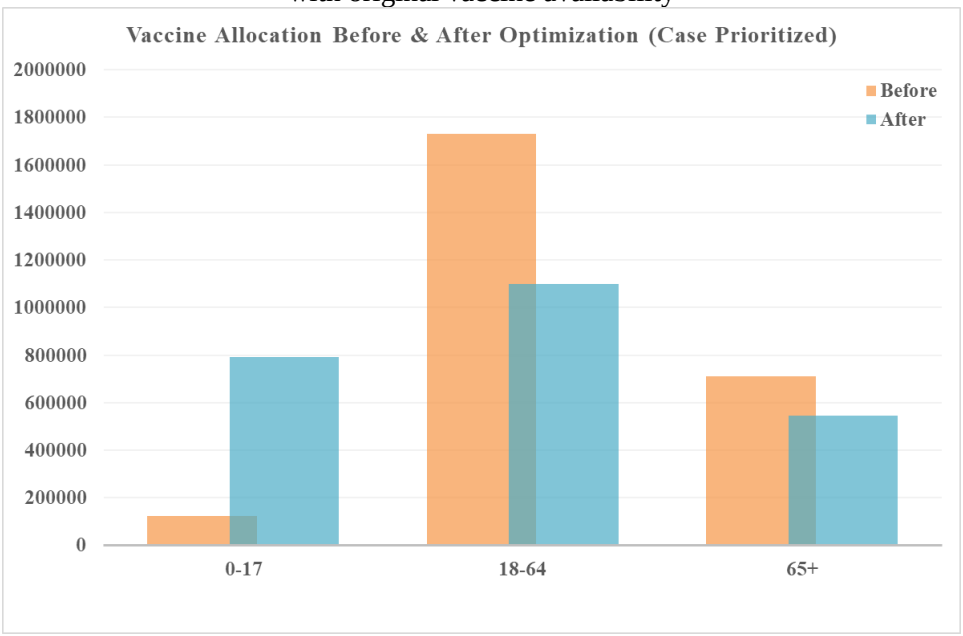

Figure 23: Vaccine allocation comparison for case-prioritized vaccine optimization with original vaccine availability

On the other hand, when prioritizing the reduction of deaths, the optimal vaccine allocation policy retains an emphasis on vaccinating the younger age group (0-17 years) due to their role in driving overall case numbers. However, this strategy also allocates more vaccines to the older population (65+ years), who face a higher risk of severe illness and death from COVID-19. Older individuals typically have weaker immune systems and may suffer from comorbidities, making them more vulnerable to the virus's severe effects. By prioritizing the vaccination of the older population, the

policy aims to protect those at the highest risk of death, leading to a substantial reduction in fatalities. This allocation strategy strikes a balance between limiting virus transmission by vaccinating the younger age group and protecting the most vulnerable members of society, resulting in a decrease in both case numbers and deaths. According to the model results, this optimal allocation strategy could reduce 220,010 cases and 6,319 deaths.

The results of the second scenario demonstrate that a more targeted vaccine allocation strategy, considering the dynamic transmission patterns and the specific priorities of policymakers, can lead to significantly better healthcare outcomes. By shifting the vaccine distribution towards age groups that have the most significant impact on transmission and death rates, it is possible to achieve substantial reductions in both cases and fatalities. This analysis highlights the importance of data-driven and adaptable vaccine allocation policies to effectively manage the ongoing pandemic and safeguard public health.

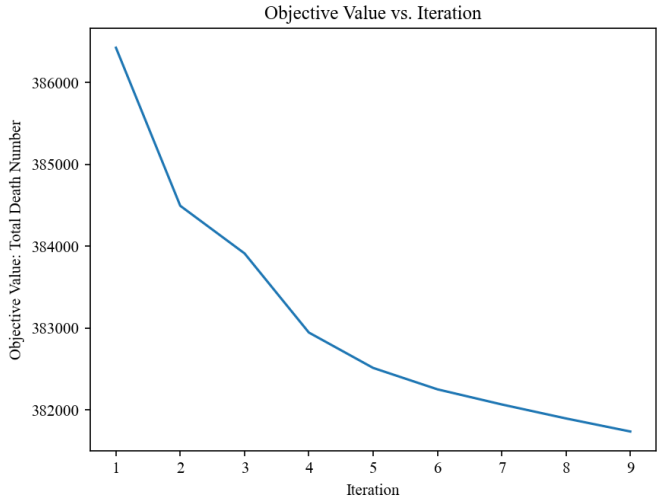

Figure 24: Training process of death-prioritized vaccine optimization with original vaccine availability

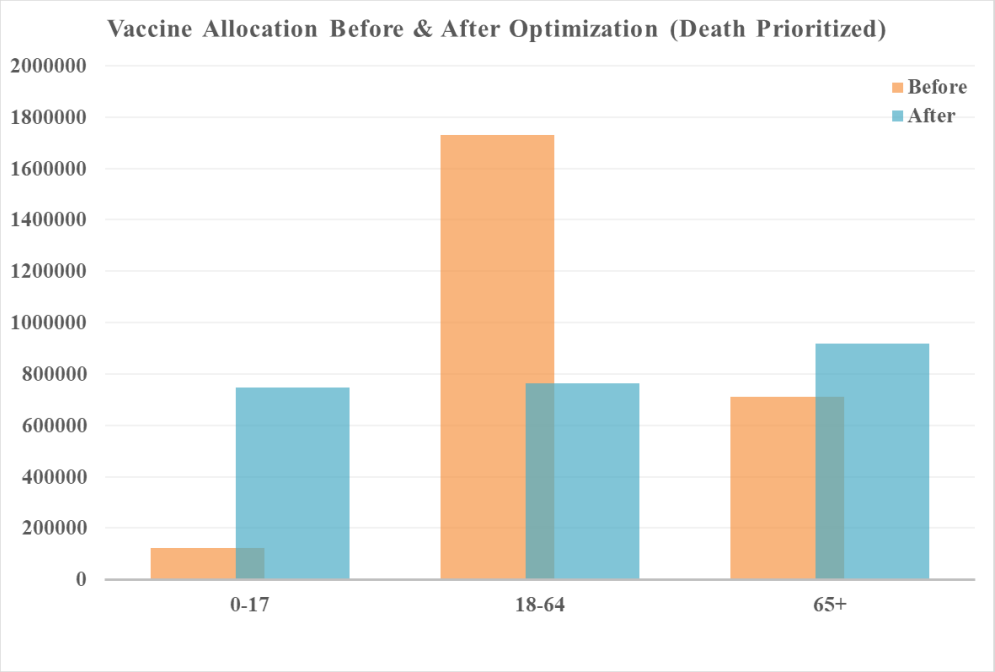

Figure 25: Vaccine allocation comparison for death-prioritized vaccine optimization with original vaccine availability

Figure 26 and Figure 27 show the redistribution of original amount of vaccine among 50 states for case/death-prioritized scenario. According to the comparison, some states, like Maine, Vermont, Montana, South Dakota, and New Hampshire, need more vaccines than distributed. There are several factors contribute to their increased need for vaccines to effectively reduce the number of cases and deaths. These states generally have smaller populations, fewer resources, and limited healthcare infrastructure, particularly in rural areas, which can affect their ability to quickly identify, treat, and manage cases. An increased allocation of vaccines could help to compensate for these limitations by reducing the number of severe cases that require hospitalization and specialized care. Moreover, the low population density and sparse distribution of the population in these states create challenges in vaccine distribution and administration, leading to slower immunization rates. Increasing the allocation of vaccines to these states can help overcome logistical challenges and ensure that more people receive the vaccine. Additionally, the age distribution of the population in these states may play a role in the increased need for vaccines, as

some have a higher proportion of older adults who are at greater risk of severe illness and death due to COVID-19. Prioritizing vaccine allocation to these states can help protect their most vulnerable citizens and reduce fatalities. Lastly, the effectiveness of public health policies and their implementation varies between states, and those with a higher need for vaccines may have less stringent public health measures or lower compliance. By increasing the vaccine allocation, these states can mitigate the impact of less effective public health policies on case numbers and deaths. In summary, a targeted approach to vaccine allocation that considers the unique needs and challenges faced by states like Maine, Vermont, Montana, South Dakota, and New Hampshire can lead to a more effective reduction in both cases and deaths through a data-driven and adaptable vaccine allocation strategy that addresses disparities in vaccine distribution and helps to better manage the ongoing pandemic.

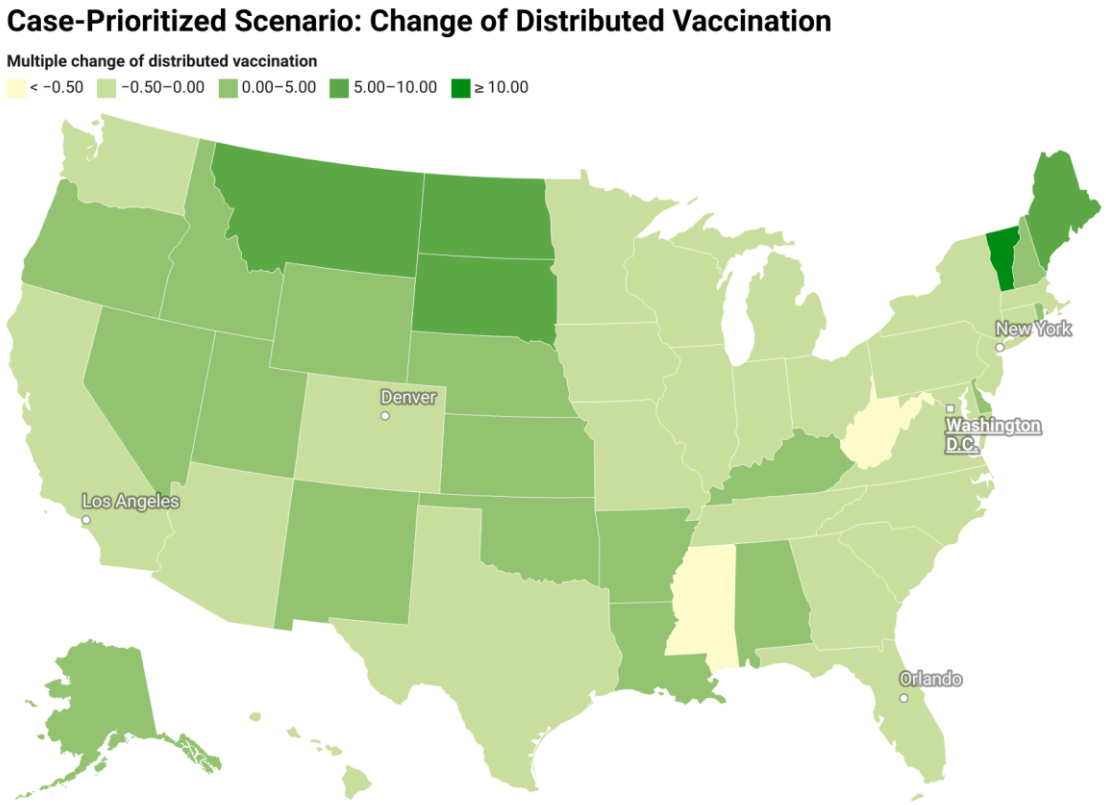

Figure 26: Redistribution of original amount of vaccine among 50 states for case-prioritized scenario (Change of the vaccine distribution divided by original vaccine number)

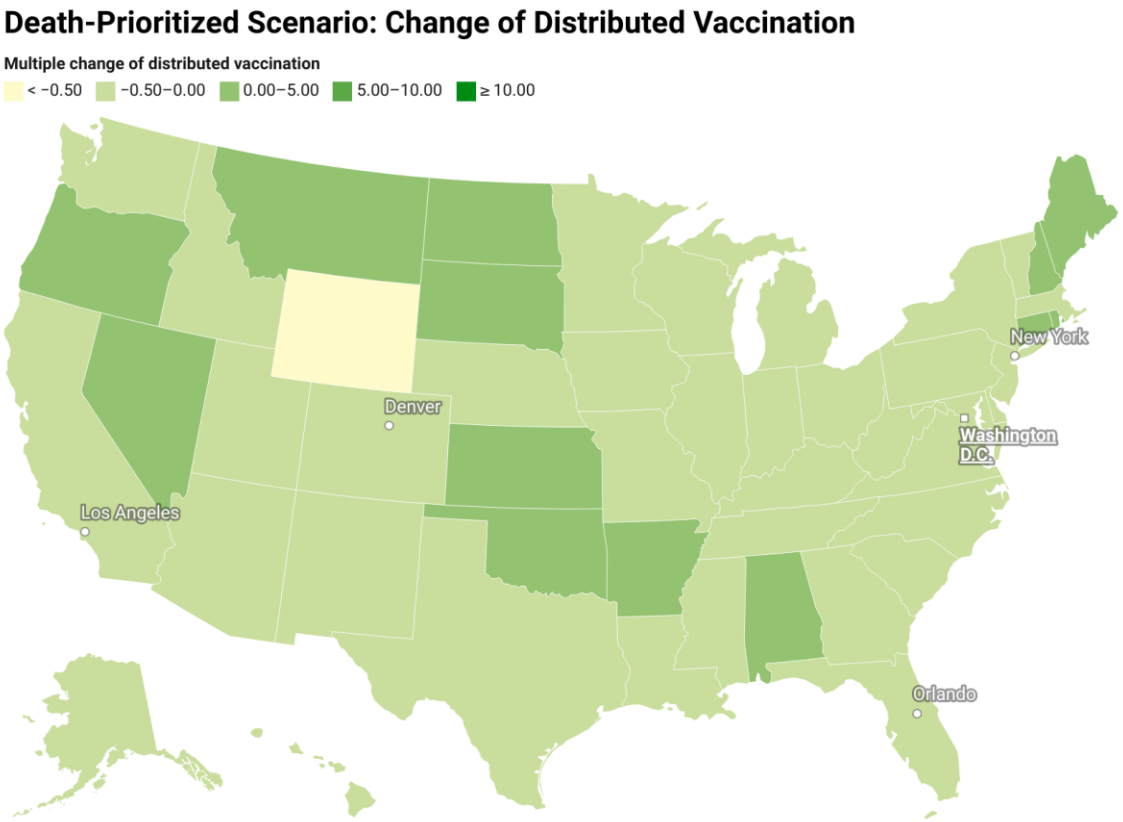

Figure 27: Redistribution of original amount of vaccine among 50 states for death-prioritized scenario (Change of the vaccine distribution divided by original vaccine number)

6.4.3 Vaccine Allocation Policy with 10 Times of Original Vaccine Availability

In the third scenario, we explore the impact of a substantial increase in vaccine availability, specifically 10 times the original weekly availability. This scenario aims to understand the optimal vaccine allocation policy and the corresponding healthcare outcomes, given this significant increase in resources. Similar to the second scenario, we analyze the best vaccine allocation policy under two different prioritizations: reducing the number of cases as much as possible and reducing the number of deaths as much as possible.

The vaccine allocation results for both prioritizations remain consistent with the second scenario. They maintain an emphasis on vaccinating the younger age group (0-17 years) due to their role in driving overall case numbers. Additionally, prioritizing the vaccination of the older population is

recommended if policymakers emphasize the reduction in fatalities. This demonstrates the robustness of the allocation strategies across different levels of vaccine availability.

According to the model results, the optimal allocation strategy could potentially reduce 2,561,885 cases and 6,735 deaths when prioritizing the reduction of cases. Although this represents a significant reduction in case numbers, the increment in vaccination resources only leads to an additional reduction of 519,573 cases. This is primarily because, even with 10 times the vaccine availability for each week, the available amount of vaccine is still too small to control the epidemic during the first several months. The limited vaccination resources at the beginning of a new wave of transmission make it difficult to substantially reduce the number of cases once the disease has spread widely throughout the population.

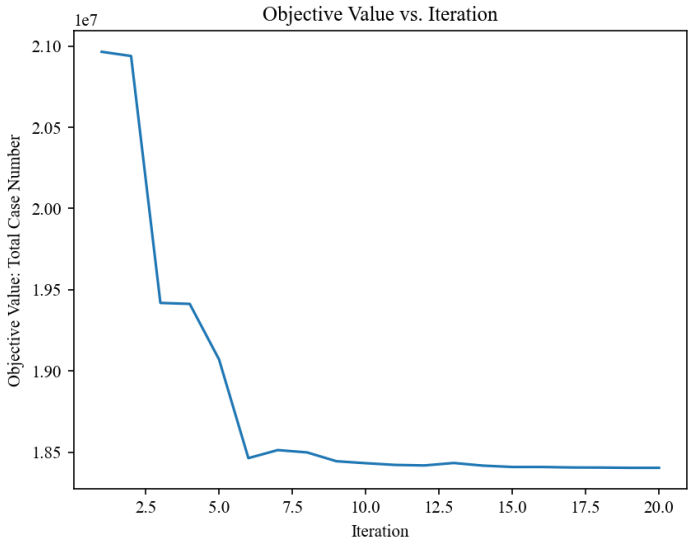

Figure 28: Training process of case-prioritized vaccine optimization with 10 times vaccine availability

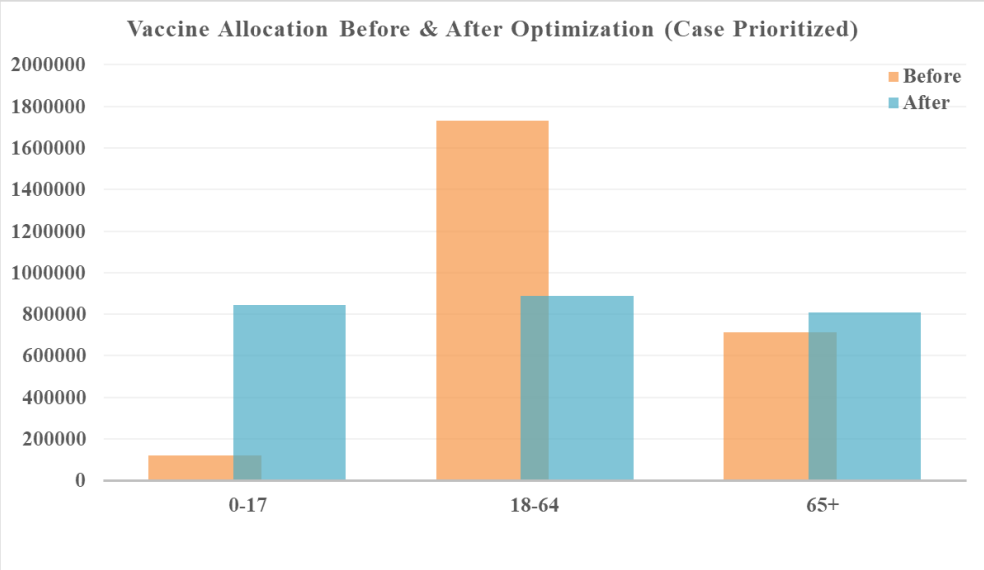

Figure 29: Vaccine allocation comparison for case-prioritized vaccine optimization with 10 times vaccine availability

In contrast, the optimal allocation strategy could potentially reduce 1,537,008 cases and 16,014 deaths when prioritizing the reduction of deaths, leading to an additional 9,695 lives saved. These finding highlights that, even though a 10-fold increase in vaccination resources may not result in a dramatic decrease in case numbers, it can have a substantial impact on saving lives. The allocation of sufficient vaccination resources can protect vulnerable populations, particularly the older population, who are at the highest risk of severe illness and death from COVID-19.

When analyzing the third scenario, it is crucial to acknowledge the inherent limitations of increasing vaccine availability. A 10-fold increase in weekly availability may not be feasible due to production constraints, logistical challenges, and the need for a rapid and efficient rollout. Nonetheless, this scenario provides valuable insights into the potential impact of increased vaccination resources on the overall healthcare outcomes during the pandemic.

In conclusion, the third scenario demonstrates the importance of strategic vaccine allocation, particularly when resources are limited. Although a substantial increase in vaccine availability can lead to significant reductions in both case numbers and deaths, it is essential to prioritize the most vulnerable populations and target age groups that play a significant role in virus transmission. By

understanding the potential outcomes under different prioritizations and levels of vaccine availability, policymakers can make informed decisions to optimize healthcare outcomes and mitigate the impact of the COVID-19 pandemic.

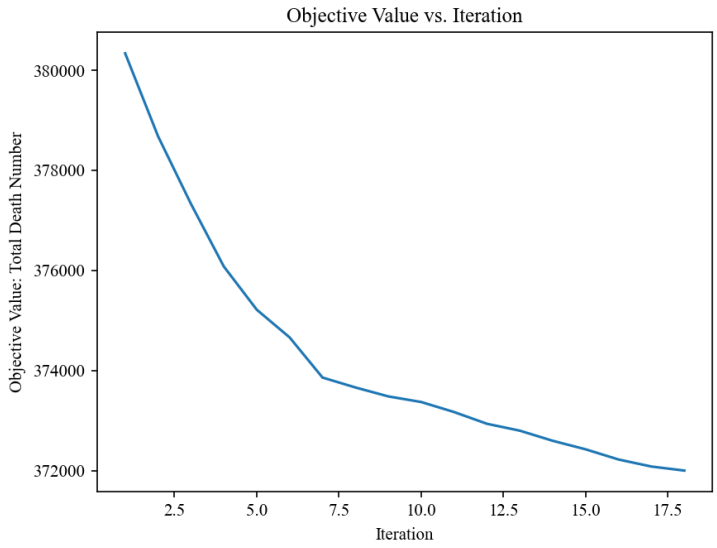

Figure 30: Training process of death-prioritized vaccine optimization with 10 times vaccine availability

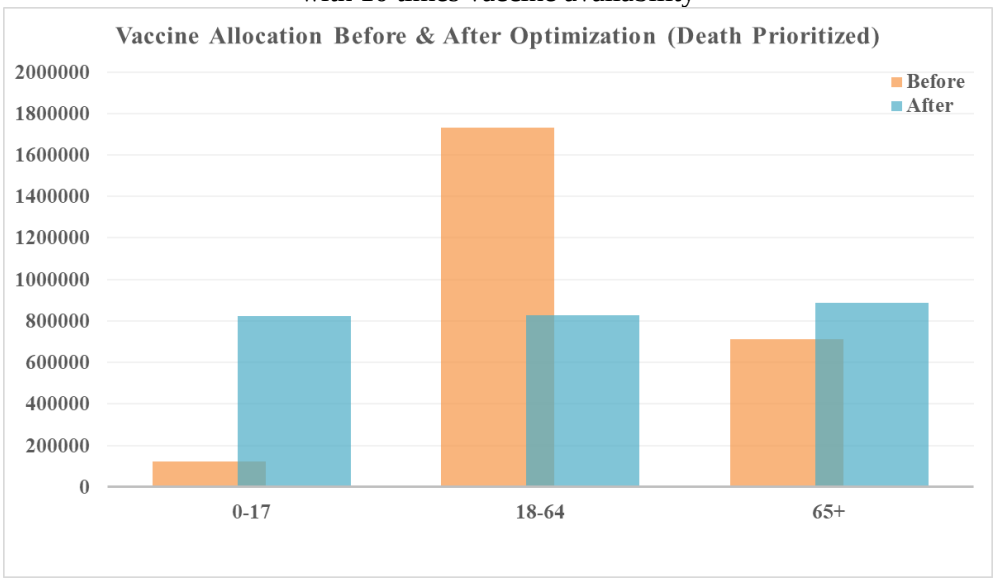

Figure 31: Vaccine allocation comparison for death-prioritized vaccine optimization with 10 times vaccine availability

6.5 Conclusion

In this chapter, we have presented an age-structured dynamic model that incorporates vaccination data, providing a more accurate and nuanced understanding of disease transmission dynamics

among different age groups. We have also proposed a novel method for optimizing vaccine allocation across regions, taking into consideration the varying transmission severity, population structures, and limited resources for vaccine distribution. By analyzing the implications of different vaccine allocation policies under various scenarios of vaccine availability, we have provided valuable insights into the most effective strategies for managing the disease and mitigating its impact on public health.

Our work has demonstrated the importance of data-driven and adaptable vaccine allocation policies in managing the ongoing pandemic and safeguarding public health. By incorporating regional transmission patterns, population structures, and the impact of vaccination on disease transmission and death rates, our age-structured dynamic model with vaccination and optimization framework offers a valuable tool for public health authorities and policymakers to make informed decisions on vaccine distribution.

The analysis of different vaccine allocation policies under various scenarios of vaccine availability highlights the need for targeted and strategic vaccine distribution, particularly when resources are limited. By prioritizing the most vulnerable populations and targeting age groups that play a significant role in virus transmission, it is possible to achieve substantial reductions in both cases and fatalities. Our findings emphasize the crucial role that vaccines play in mitigating the severity of the pandemic, and the importance of an effective vaccine distribution strategy in controlling the spread of the virus and reducing morbidity and mortality rates.

As the COVID-19 pandemic continues to evolve, it is essential for public health authorities and policymakers to remain vigilant and adapt to changing circumstances. Our age-structured dynamic model with vaccination and optimization framework can serve as a valuable foundation for future research and decision-making, as new data becomes available and new challenges arise. The

ongoing refinement and application of our model will help inform evidence-based decision-making for vaccine allocation and distribution strategies, contributing to a better understanding of disease transmission dynamics and ensuring a more efficient and equitable use of limited resources. In summary, our work contributes to a deeper understanding of the complex interplay between age-specific transmission dynamics, vaccine rollout, and disease outcomes. By developing and applying our age-structured dynamic model with vaccination, we hope to inform evidence-based decision-making for vaccine allocation and distribution strategies, ultimately helping to minimize the adverse effects of infectious diseases on populations and ensure a more efficient and equitable use of limited resources.

Chapter 7 Conclusion & Discussion

7.1 Conclusion and Discussion

The present thesis has focused on advancing the understanding of infectious disease dynamics and providing tools for effective intervention and optimization strategies. By integrating dynamic modeling with time-varying transmission and fatality rates, spatial interaction analysis, and vaccination reallocation, we have presented a comprehensive and adaptable framework for studying the complex interplay of factors that influence the spread and control of infectious diseases.

7.1.1 Dynamic Modeling with Time-Varying Transmission and Fatality Rates

Our work on dynamic modeling with time-varying transmission and fatality rates highlights the importance of employing flexible and data-driven approaches to understanding and managing infectious disease outbreaks. The SEIRD model with time-dependent transmission and death rates offers a robust framework for analyzing the historical disease transmission and informing evidence-based public health decision-making. The insights gained from this research can help researchers and policymakers develop targeted and effective interventions to mitigate the devastating impacts of infectious diseases on global health and wellbeing.

In particular, our research demonstrates the effectiveness of the SEIRD model with time-dependent transmission and death rates in capturing the transmission dynamics of COVID-19 across the United States. The model's fitting accuracy is relatively high for both the number of cases and deaths, with a majority of states exhibiting an RRMSE below 2%. This demonstrates the value of incorporating time-varying transmission and fatality rates into epidemic modeling efforts. Moreover, the multi-phase model introduced to account for locations exhibiting multiple waves of

infection further improves the model's accuracy, allowing for better prediction and management of the ongoing pandemic.

Moreover, our sensitivity analysis highlights the significance of certain parameters in influencing the predicted number of cases and deaths. Understanding the impact of changes in these key parameters can help researchers and policymakers anticipate potential shifts in disease transmission and develop appropriate interventions. Furthermore, our Monte Carlo simulations demonstrate the extent to which uncertainties in parameter values can impact model outcomes, emphasizing the need for robust sensitivity analyses in infectious disease modeling.

7.1.2 Integration of Dynamic Modeling with Spatial Interaction and Effect Analysis

By incorporating spatial interaction analysis, we have demonstrated the critical role of understanding complex disease transmission dynamics across regions for effective prevention and control efforts. Our multi-regional dynamic model with spatial interaction provides a valuable tool for identifying high-risk areas and assessing the impact of various interventions, such as travel restrictions or quarantine measures. Furthermore, our research on the transmission export index and the interplay between various factors influencing disease transmission can help guide future decision-making processes and inform public health policy.

The investigation of disease transmission through spatial interaction, particularly state-level travel, has provided valuable insights into the complex dynamics that govern the spread of infectious diseases. By developing a multi-regional dynamic model with spatial interaction, we have been able to accurately capture historical transmission patterns and evaluate the impact of long-distance travel on disease transmission. This understanding is crucial for developing effective prevention and control strategies for future outbreaks.

The introduction of the transmission export index, which combines local transmissibility and the potential for infectious travelers to spread diseases to new regions, has proven to be an important tool for assessing the risk of disease transmission between regions. By identifying high-risk areas, appropriate interventions, such as travel restrictions or quarantine measures, can be put in place to mitigate disease spread and protect public health.

Furthermore, our causal analysis of factors influencing the transmission export index has highlighted the importance of considering a wide range of influences, including state-level actions, political events, and festivals/entertainment events. By examining the interplay between these factors and their potential impact on disease transmission, we can better inform future decision-making processes and guide public health policy.

Additionally, our analysis of state-level policies aimed at reducing the transmission export index has demonstrated the effectiveness of various measures, such as promoting public health guidance, implementing social distancing measures, introducing phased plans for economic reopening, and imposing travel restrictions. These strategies play a vital role in protecting residents and limiting the spread of infections both within individual states and across the nation.

In conclusion, this chapter has shed light on the critical role of spatial interaction in disease transmission and the importance of understanding these dynamics for effective prevention and control efforts. The methods and findings presented here can serve as a foundation for future research, policy development, and public health interventions aimed at mitigating the impact of infectious diseases on a regional, national, and global scale.

7.1.3 Integration of Dynamic Modeling with Vaccination Reallocation

The integration of an age-structured dynamic model with vaccination data and optimization techniques allows us to better understand age-specific transmission dynamics and develop efficient

strategies for vaccine allocation and distribution. Our analysis demonstrates the importance of targeted and strategic vaccine distribution, especially when resources are limited. By prioritizing vulnerable populations and targeting age groups that play a significant role in virus transmission, it is possible to achieve substantial reductions in both cases and fatalities.

In this chapter, we have presented an age-structured dynamic model that incorporates vaccination data, providing a more accurate and nuanced understanding of disease transmission dynamics among different age groups. We have also proposed a novel method for optimizing vaccine allocation across regions, taking into consideration the varying transmission severity, population structures, and limited resources for vaccine distribution. By analyzing the implications of different vaccine allocation policies under various scenarios of vaccine availability, we have provided valuable insights into the most effective strategies for managing the disease and mitigating its impact on public health.

Our work has demonstrated the importance of data-driven and adaptable vaccine allocation policies in managing the ongoing pandemic and safeguarding public health. By incorporating regional transmission patterns, population structures, and the impact of vaccination on disease transmission and death rates, our age-structured dynamic model with vaccination and optimization framework offers a valuable tool for public health authorities and policymakers to make informed decisions on vaccine distribution.

The analysis of different vaccine allocation policies under various scenarios of vaccine availability highlights the need for targeted and strategic vaccine distribution, particularly when resources are limited. By prioritizing the most vulnerable populations and targeting age groups that play a significant role in virus transmission, it is possible to achieve substantial reductions in both cases and fatalities. Our findings emphasize the crucial role that vaccines play in mitigating the severity

of the pandemic, and the importance of an effective vaccine distribution strategy in controlling the spread of the virus and reducing morbidity and mortality rates.

7.1.4 Overall Summary

Our work contributes to a deeper understanding of the complex interplay between disease transmission dynamics, intervention strategies, and optimization techniques. The frameworks and models developed in this thesis can serve as a foundation for future research, policy development, and public health interventions aimed at mitigating the impact of infectious diseases on a regional, national, and global scale.

As the COVID-19 pandemic continues to evolve, and new diseases arise, it is essential for public health authorities and policymakers to remain vigilant and adapt to changing circumstances. The ongoing refinement and application of the models and approaches presented in this thesis will help inform evidence-based decision-making for intervention strategies and resource allocation, ultimately helping to minimize the adverse effects of infectious diseases on populations and ensure a more efficient and equitable use of limited resources.

In summary, our work highlights the potential of dynamic modeling with time-varying transmission and fatality rates, spatial interaction analysis, and vaccination reallocation in understanding and managing infectious disease outbreaks. By leveraging the insights gained from our models and approaches, researchers and policymakers can work together to develop targeted and effective interventions to address infectious diseases and safeguard public health.

7.2 Future Study

The findings from the main chapters of this thesis have provided valuable insights into disease transmission dynamics and informed public health decision-making. By integrating dynamic modeling with time-varying transmission and fatality rates, spatial interaction and effect analysis, and vaccination reallocation, the research has demonstrated the potential for significant improvements in our understanding of disease transmission and the development of targeted intervention strategies. The following future study outlines several key areas for further exploration, building on the foundation established by the current research.

(1) Developing a Unified Modeling Framework

An important direction for future research is the development of a unified modeling framework that integrates the components from the three chapters. This would involve incorporating time-varying transmission and fatality rates, spatial interaction and effect analysis, and vaccination reallocation into a single, comprehensive model. Such a framework could provide a more holistic understanding of disease transmission dynamics and support the development of more effective and targeted public health policies. Simultaneously, as part of our sensitivity analysis, we intend to scrutinize the impact of fixed parameters on the model outputs. These fixed parameters include key elements such as the incubation period and recovery time. By analyzing these factors, we aim to understand their influence on the model and gain insights into how variations in these parameters might affect the model's predictions. This will also help us assess the robustness of our model under different scenarios and conditions.

(2) Incorporating Additional Factors and Granularity

Expanding the model to include additional factors influencing disease transmission, such as population density, age distribution, socioeconomic status, cultural factors, climate, and healthcare

infrastructure, could provide a more comprehensive understanding of disease transmission dynamics. Moreover, enhancing the granularity of the model to operate at finer spatial resolutions, such as county or city-level, would help identify localized transmission hotspots and inform targeted intervention strategies.

(3) Investigating the Impact of Emerging Variants, Waning Immunity, and Vaccine Hesitancy

Future research should explore the impact of emerging virus variants, waning immunity, and vaccine hesitancy on transmission and fatality rates, as well as the effectiveness of various intervention strategies. This could involve modeling the potential influence of emerging variants, assessing the impact of public health interventions aimed at addressing vaccine hesitancy, and exploring the use of booster shots in response to waning immunity.

(4) Analyzing and Comparing the Effects of Different Intervention Strategies

Future studies could adapt the dynamic model to assess the impact of various intervention strategies, such as vaccination campaigns, social distancing measures, and testing, tracing, and isolation protocols. This would involve comparing the effectiveness of these strategies under different scenarios and providing valuable insights for policymakers and public health officials.

(5) Assessing the Impact of Public Health Policies on Transmission Dynamics

Systematically evaluating the effectiveness of various public health policies in reducing transmission risk and optimizing resource allocation can help inform policy decisions. Future research could develop computational models to simulate the impact of different policy scenarios or conduct comparative analyses of regions with varying policy approaches.

By pursuing these future research directions, the scientific community can continue to refine and expand upon the insights gained from integrating dynamic modeling with time-varying

transmission and fatality rates, spatial interaction and effect analysis, and vaccination reallocation. This ongoing research will contribute to a more comprehensive understanding of disease transmission dynamics and support the development of evidence-based public health policies and interventions in the face of emerging infectious diseases.

References

[1] J. Snow, “On the Mode of Communication of Cholera, London: John Churchill, 1849,” Snow Follow. this up with a number Artic. that Refin. his theory.

[2] M. G. Roberts and J. A. P. Heesterbeek, Mathematical models in epidemiology. EOLSS, 2003.

[3] W. O. Kermack and A. G. McKendrick, “Contributions to the mathematical theory of epidemics—I,” Bull. Math. Biol., vol. 53, no. 1–2, pp. 33–55, 1991.

[4] L. Sattenspiel and A. Lloyd, The geographic spread of infectious diseases: models and applications, vol. 5. Princeton University Press, 2009.

[5] M. Y. Li, H. L. Smith, and L. Wang, “Global dynamics of an SEIR epidemic model with vertical transmission,” SIAM J. Appl. Math., vol. 62, no. 1, pp. 58–69, 2001.

[6] K. Linka, M. Peirlinck, F. Sahli Costabal, and E. Kuhl, “Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions,” Comput. Methods Biomech. Biomed. Engin., vol. 23, no. 11, pp. 710–717, 2020, doi: 10.1080/10255842.2020.1759560.

[7] Y. Fang, Y. Nie, and M. Penny, “Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis,” J. Med. Virol., vol. 92, no. 6, pp. 645–659, 2020, doi: 10.1002/jmv.25750.

[8] P. Yang et al., “Feasibility study of mitigation and suppression strategies for controlling COVID-19 outbreaks in London and Wuhan,” PLoS One, vol. 15, no. 8, p. e0236857, 2020, doi: 10.1371/journal.pone.0236857.

[9] R. Rustan and L. Handayani, “The Outbreak’S modeling of coronavirus (Covid-19) using the modified seir model in Indonesia,” Spektra J. Fis. dan Apl., vol. 5, no. 1, pp. 61–68, 2020.

[10] S. Sousa et al., “Efficacy and Safety of Vaccination in Pediatric Patients with Systemic Inflammatory Rheumatic Diseases: a systematic review of the literature.,” Acta Reumatol. Port., no. 1, 2017.

[11] A. R. Tuite, D. N. Fisman, J. C. Kwong, and A. L. Greer, “Optimal pandemic influenza vaccine allocation strategies for the Canadian population,” PLoS One, vol. 5, no. 5, p. e10520, 2010.

[12] H. W. Hethcote, “An age-structured model for pertussis transmission,” Math. Biosci., vol. 145, no. 2, pp. 89–136, 1997.

[13] S. Abrams et al., “Modelling the early phase of the Belgian COVID-19 epidemic using a stochastic compartmental model and studying its implied future trajectories,” Epidemics,

vol. 35, p. 100449, 2021.

[14] P. Minayev and N. Ferguson, “Improving the realism of deterministic multi-strain models: implications for modelling influenza A,” J. R. Soc. Interface, vol. 6, no. 35, pp. 509–518, 2009.

[15] B. W. Kooi, M. Aguiar, and N. Stollenwerk, “Analysis of an asymmetric two-strain dengue model,” Math. Biosci., vol. 248, pp. 128–139, 2014.

[16] K. Ameri and K. D. Cooper, “A network-based compartmental model for the spread of whooping cough in Nebraska,” AMIA Summits Transl. Sci. Proc., vol. 2019, p. 388, 2019.

[17] A. Wiratsudakul, P. Suparit, and C. Modchang, “Dynamics of Zika virus outbreaks: an overview of mathematical modeling approaches,” PeerJ, vol. 6, p. e4526, 2018.

[18] E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, Rinaldo, and A, “On spatially explicit models of cholera epidemics,” J. R. Soc. Interface, vol. 7, no. 43, pp. 321–333, 2010.

[19] N. R. Smith et al., “Agent-based models of malaria transmission: a systematic review,” Malar. J., vol. 17, no. 1, pp. 1–16, 2018.

[20] D. O’Sullivan, M. Gahegan, D. J. Exeter, and B. Adams, “Spatially explicit models for exploring COVID-19 lockdown strategies,” Trans. GIS, vol. 24, no. 4, pp. 967–1000, 2020.

[21] Y. M. Tobo, J. Bartacek, and I. Nopens, “Linking CFD and kinetic models in anaerobic digestion using a compartmental model approach,” Processes, vol. 8, no. 6, p. 703, 2020.

[22] K. Bhaskaran, A. Gasparrini, S. Hajat, L. Smeeth, and B. Armstrong, “Time series regression studies in environmental epidemiology,” Int. J. Epidemiol., vol. 42, no. 4, pp. 1187–1195, 2013.

[23] M. Saez, A. Tobias, D. Varga, and M. A. Barceló, “Effectiveness of the measures to flatten the epidemic curve of COVID-19. The case of Spain,” Sci. Total Environ., p. 138761, 2020.

[24] S. Greenland, “Bayesian perspectives for epidemiological research. II. Regression analysis,” Int. J. Epidemiol., vol. 36, no. 1, pp. 195–202, 2007.

[25] G. Hamra, R. MacLehose, and D. Richardson, “Markov chain Monte Carlo: an introduction for epidemiologists,” Int. J. Epidemiol., vol. 42, no. 2, pp. 627–634, 2013.

[26] P. C. L. Silva, P. V. C. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães, and R. C. P. Silva, “COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions,” Chaos, Solitons & Fractals, vol. 139, p. 110088, 2020.

[27] L. Bian and D. Liebner, “A network model for dispersion of communicable diseases,” Trans. GIS, vol. 11, no. 2, pp. 155–173, 2007.

[28] L. Bian, “Object-oriented representation for modelling mobile objects in an aquatic environment,” Int. J. Geogr. Inf. Sci., vol. 14, no. 7, pp. 603–623, 2000.

[29] J. A. Simoes, “An agent-based/network approach to spatial epidemics,” in Agent-based models of geographical systems, Springer, 2012, pp. 591–610.

[30] J. H. Bergstrand, “The gravity equation in international trade: some microeconomic foundations and empirical evidence,” Rev. Econ. Stat., pp. 474–481, 1985.

[31] S. Merler and M. Ajelli, “The role of population heterogeneity and human mobility in the spread of pandemic influenza,” Proc. R. Soc. B Biol. Sci., vol. 277, no. 1681, pp. 557–565, 2010.

[32] F. Simini, M. C. González, A. Maritan, and A.-L. Barabási, “A universal model for mobility and migration patterns,” Nature, vol. 484, no. 7392, pp. 96–100, 2012.

[33] A. P. Masucci, J. Serras, A. Johansson, and M. Batty, “Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows,” Phys. Rev. E, vol. 88, no. 2, p. 22812, 2013.

[34] X. Li, H. Tian, D. Lai, and Z. Zhang, “Validation of the gravity model in predicting the global spread of influenza,” Int. J. Environ. Res. Public Health, vol. 8, no. 8, pp. 3134–3143, 2011.

[35] M. U. G. Kraemer et al., “Utilizing general human movement models to predict the spread of emerging infectious diseases in resource poor settings,” Sci. Rep., vol. 9, no. 1, pp. 1–11, 2019.

[36] A. D. Cliff, P. Haggett, and J. K. Ord, Spatial aspects of influenza epidemics. Routledge, 1986.

[37] C. J. Rhodes and R. M. Anderson, “Epidemic thresholds and vaccination in a lattice model of disease spread,” Theor. Popul. Biol., vol. 52, no. 2, pp. 101–118, 1997.

[38] J. S. Koopman and J. W. Lynch, “Individual causal models and population system models in epidemiology.,” Am. J. Public Health, vol. 89, no. 8, pp. 1170–1174, 1999.

[39] T. A. Glass and M. J. McAtee, “Behavioral science at the crossroads in public health: extending horizons, envisioning the future,” Soc. Sci. \& Med., vol. 62, no. 7, pp. 1650–1671, 2006.

[40] D. J. Watts, R. Muhamad, D. C. Medina, and P. S. Dodds, “Multiscale, resurgent epidemics in a hierarchical metapopulation model,” Proc. Natl. Acad. Sci., vol. 102, no. 32, pp. 11157–11162, 2005.

[41] E. E. Holmes, “Basic Epidemiological Concepts in a Spatial Context. In ‘Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions’(D. Tilman, P. Kareiva, eds.) pp. 111-136.” Princeton University Press, Princeton, 1997.

[42] M. J. Keeling, “Metapopulation moments: coupling, stochasticity and persistence,” J. Anim. Ecol., vol. 69, no. 5, pp. 725–736, 2000.

[43] T. C. Reluga, “An SIS epidemiology game with two subpopulations,” J. Biol. Dyn., vol. 3, no. 5, pp. 515–531, 2009.

[44] R. J. Doran and S. W. Laffan, “Simulating the spatial dynamics of foot and mouth disease outbreaks in feral pigs and livestock in Queensland, Australia, using a susceptible-infected-recovered cellular automata model,” Prev. Vet. Med., vol. 70, no. 1–2, pp. 133–152, 2005.

[45] L. Mao and L. Bian, “A dynamic network with individual mobility for designing vaccination strategies,” Trans. GIS, vol. 14, no. 4, pp. 533–545, 2010.

[46] M. Smallman-Raynor and A. D. Cliff, “Epidemic diffusion processes in a system of US military camps: Transfer diffusion and the spread of typhoid fever in the Spanish-American War, 1898,” Ann. Assoc. Am. Geogr., vol. 91, no. 1, pp. 71–91, 2001.

[47] M. R. Smallman-Raynor and A. D. Cliff, “Impact of infectious diseases on war.,” Infect. Dis. Clin. North Am., vol. 18, no. 2, pp. 341–368, 2004.

[48] M. L. Martins, G. Ceotto, S. G. Alves, C. C. B. Bufon, J. M. Silva, and F. F. Laranjeira, “A cellular automata model for citrus variegated chlorosis,” Phys. A Stat. Mech. its Appl., vol. 295, no. 1–2, pp. 42–48, 2001.

[49] K. T. D. Eames and M. J. Keeling, “Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases,” Proc. Natl. Acad. Sci., vol. 99, no. 20, pp. 13330–13335, 2002.

[50] M. Ienca and E. Vayena, “On the responsible use of digital data to tackle the COVID-19 pandemic,” Nat. Med., vol. 26, no. 4, pp. 463–464, 2020.

[51] D. Brockmann and D. Helbing, “The hidden geometry of complex, network-driven contagion phenomena,” Science (80-.)., vol. 342, no. 6164, pp. 1337–1342, 2013.

[52] A. G. Wilson, “A family of spatial interaction models, and associated developments,” Environ. Plan. A, vol. 3, no. 1, pp. 1–32, 1971.

[53] Y. Liu, Z. Sui, C. Kang, and Y. Gao, “Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data,” PLoS One, vol. 9, no. 1, p. e86026, 2014.

[54] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, “Multiscale mobility networks and the spatial spreading of infectious diseases,” Proc. Natl. Acad. Sci., vol. 106, no. 51, pp. 21484–21489, 2009.

[55] C. Poletto et al., “Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic,” Eurosurveillance, vol. 19, no. 42, 2014.

[56] M. U. G. Kraemer et al., “The effect of human mobility and control measures on the

COVID-19 epidemic in China," Science (80-.)., vol. 368, no. 6490, pp. 493–497, 2020.

[57] F. Tanser, B. Gijsbertsen, and K. Herbst, "Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system," Soc. Sci. \& Med., vol. 63, no. 3, pp. 691–705, 2006.

[58] A. J. Tatem, C. A. Guerra, C. W. Kabaria, A. M. Noor, and S. I. Hay, "Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity," Malar. J., vol. 7, pp. 1–17, 2008.

[59] M. Gatto et al., "Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures," Proc. Natl. Acad. Sci., vol. 117, no. 19, pp. 10484–10491, 2020.

[60] A. Wesolowski, C. O. Buckee, L. Bengtsson, E. Wetter, X. Lu, and A. J. Tatem, "Commentary: containing the Ebola outbreak-the potential and challenge of mobile network data," PLoS Curr., vol. 6, 2014.

[61] S. T. Stoddard et al., "The role of human movement in the transmission of vector-borne pathogens," PLoS Negl. Trop. Dis., vol. 3, no. 7, p. e481, 2009.

[62] A. Wesolowski et al., "Quantifying the impact of human mobility on malaria," Science (80-.)., vol. 338, no. 6104, pp. 267–270, 2012.

[63] L. Hufnagel, D. Brockmann, and T. Geisel, "Forecast and control of epidemics in a globalized world," Proc. Natl. Acad. Sci., vol. 101, no. 42, pp. 15124–15129, 2004.

[64] C. Viboud et al., "The RAPIDD ebola forecasting challenge: Synthesis and lessons learnt," Epidemics, vol. 22, pp. 13–21, 2018.

[65] S. Flaxman et al., "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update," arXiv Prepr. arXiv2004.11342, 2020.

[66] B. J. Cowling et al., "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study," Lancet Public Heal., vol. 5, no. 5, pp. e279--e288, 2020.

[67] S. Woolhandler and D. U. Himmelstein, "Intersecting US epidemics: COVID-19 and lack of health insurance." American College of Physicians, 2020.

[68] M. McKee and D. Stuckler, "If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future," Nat. Med., vol. 26, no. 5, pp. 640–642, 2020.

[69] O. Y. Özalt\in, O. A. Prokopyev, and A. J. Schaefer, "Optimal design of the seasonal influenza vaccine with manufacturing autonomy," INFORMS J. Comput., vol. 30, no. 2, pp. 371–387, 2018.

[70] C. Gerdil, “The annual production cycle for influenza vaccine,” Vaccine, vol. 21, no. 16, pp. 1776–1779, 2003.

[71] I. M. Longini Jr, M. E. Halloran, A. Nizam, and Y. Yang, “Containing pandemic influenza with antiviral agents,” Am. J. Epidemiol., vol. 159, no. 7, pp. 623–633, 2004.

[72] A. S. Monto, J. S. Koopman, and I. R. A. M. LONGINI JR, “Tecumseh study of illness. XIII. Influenza infection and disease, 1976--1981,” Am. J. Epidemiol., vol. 121, no. 6, pp. 811–822, 1985.

[73] J. Medlock and A. P. Galvani, “Optimizing influenza vaccine distribution,” Science (80-.)., vol. 325, no. 5948, pp. 1705–1708, 2009.

[74] T. A. Reichert, N. Sugaya, D. S. Fedson, W. P. Glezen, L. Simonsen, and M. Tashiro, “The Japanese experience with vaccinating schoolchildren against influenza,” N. Engl. J. Med., vol. 344, no. 12, pp. 889–896, 2001.

[75] M. E. Halloran, I. M. Longini, D. M. Cowart, and A. Nizam, “Community interventions and the epidemic prevention potential,” Vaccine, vol. 20, no. 27–28, pp. 3254–3262, 2002.

[76] I. M. Longini Jr et al., “Estimation of the efficacy of live, attenuated influenza vaccine from a two-year, multi-center vaccine trial: implications for influenza epidemic control,” Vaccine, vol. 18, no. 18, pp. 1902–1909, 2000.

[77] S. Bansal, B. Pourbohloul, and L. A. Meyers, “A comparative analysis of influenza vaccination programs,” PLoS Med, vol. 3, no. 10, p. e387, 2006.

[78] A. N. Hill and I. M. Longini Jr, “The critical vaccination fraction for heterogeneous epidemic models,” Math. Biosci., vol. 181, no. 1, pp. 85–106, 2003.

[79] E. Duijzer, W. van Jaarsveld, J. Wallinga, and R. Dekker, “The most efficient critical vaccination coverage and its equivalence with maximizing the herd effect,” Math. Biosci., vol. 282, pp. 68–81, 2016.

[80] L. E. Duijzer, W. L. van Jaarsveld, J. Wallinga, and R. Dekker, “Dose-optimal vaccine allocation over multiple populations,” Prod. Oper. Manag., vol. 27, no. 1, pp. 143–159, 2018.

[81] S. Enayati and O. Y. Özalt\in, “Optimal influenza vaccine distribution with equity,” Eur. J. Oper. Res., vol. 283, no. 2, pp. 714–725, 2020.

[82] L. Matrajt, J. Eaton, T. Leung, and E. R. Brown, “Vaccine optimization for COVID-19: who to vaccinate first?,” medRxiv, 2020.

[83] M. E. Gallagher et al., “Considering indirect benefits is critical when evaluating sars-cov-2 vaccine candidates,” medRxiv, 2020.

[84] A. H. Briggs, M. C. Weinstein, E. A. L. Fenwick, J. Karnon, M. J. Sculpher, and A. D. Paltiel, “Model parameter estimation and uncertainty analysis: a report of the ISPOR-

SMDM Modeling Good Research Practices Task Force Working Group--6," Med. Decis. Mak., vol. 32, no. 5, pp. 722–732, 2012.

[85] E. Vynnycky and R. White, An introduction to infectious disease modelling. OUP oxford, 2010.

[86] T. Berge, J.-S. Lubuma, G. M. Moremedi, N. Morris, and R. Kondera-Shava, "A simple mathematical model for Ebola in Africa," J. Biol. Dyn., vol. 11, no. 1, pp. 42–74, 2017.

[87] I. Cooper, A. Mondal, and C. G. Antonopoulos, "A SIR model assumption for the spread of COVID-19 in different communities," Chaos, Solitons \& Fractals, vol. 139, p. 110057, 2020.

[88] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, "A methodology for performing global uncertainty and sensitivity analysis in systems biology," J. Theor. Biol., vol. 254, no. 1, pp. 178–196, 2008.

[89] E. Massad, R. H. Behrens, M. N. Burattini, and F. A. B. Coutinho, "Modeling the risk of malaria for travelers to areas with stable malaria transmission," Malar. J., vol. 8, no. 1, pp. 1–7, 2009.

[90] M. Amaku, R. S. Azevedo, R. M. de Castro, E. Massad, and F. A. B. Coutinho, "Relationship among epidemiological parameters of six childhood infections in a non-immunized Brazilian community," Mem. Inst. Oswaldo Cruz, vol. 104, no. 6, pp. 897–900, 2009.

[91] Z. Zhang, R. Gul, and A. Zeb, "Global sensitivity analysis of COVID-19 mathematical model," Alexandria Eng. J., vol. 60, no. 1, pp. 565–572, 2021.

[92] S. H. A. Khoshnaw, R. H. Salih, and S. Sulaimany, "Mathematical modelling for coronavirus disease (COVID-19) in predicting future behaviours and sensitivity analysis," Math. Model. Nat. Phenom., vol. 15, p. 33, 2020.

[93] S. M. Blower and H. Dowlatabadi, "Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example," Int. Stat. Rev. Int. Stat., pp. 229–243, 1994.

[94] J. Wu, R. Dhingra, M. Gambhir, and J. V Remais, "Sensitivity analysis of infectious disease models: methods, advances and their application," J. R. Soc. Interface, vol. 10, no. 86, p. 20121018, 2013.

[95] J. C. Helton and F. J. Davis, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," Reliab. Eng. \& Syst. Saf., vol. 81, no. 1, pp. 23–69, 2003.

[96] I. Ciufolini and A. Paolozzi, "Mathematical prediction of the time evolution of the COVID-19 pandemic in Italy by a Gauss error function and Monte Carlo simulations," Eur. Phys. J. Plus, vol. 135, no. 4, p. 355, 2020.

[97] M. Chinazzi et al., “The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak,” Science (80-.)., vol. 368, no. 6489, pp. 395–400, 2020, doi: 10.1126/science.aba9757.

[98] B. J. Quilty et al., “The effect of travel restrictions on the geographical spread of COVID-19 between large cities in China: a modelling study,” BMC Med., vol. 18, no. 1, pp. 1–10, 2020.

[99] C. J. Worby, S. S. Chaves, J. Wallinga, M. Lipsitch, L. Finelli, and E. Goldstein, “On the relative role of different age groups in influenza epidemics,” Epidemics, vol. 13, pp. 10–16, 2015.

[100] K. A. Auger et al., “Association between statewide school closure and COVID-19 incidence and mortality in the US,” Jama, vol. 324, no. 9, pp. 859–870, 2020.

[101] A. A. C. Burns and A. Gutfraind, “Effectiveness of isolation policies in schools: evidence from a mathematical model of influenza and COVID-19,” PeerJ, vol. 9, 2021.

[102] A. Madrigal and R. Meyer, “The COVID Tracking Project,” The COVID Tracking Project, 2020. [Online]. Available: https://covidtracking.com/.

[103] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track COVID-19 in real time,” Lancet Infect. Dis., vol. 20, no. 5, pp. 533–534, 2020.

[104] M. Roser, “Ortiz-Ospina E.(2019b) Global Rise of Education,” Режим доступа https//ourworldindata. org/financing-education, дата обращения, vol. 7, 2019.

[105] M. Bailey, R. Cao, T. Kuchler, J. Stroebel, and A. Wong, “Social connectedness: Measurement, determinants, and effects,” J. Econ. Perspect., vol. 32, no. 3, pp. 259–280, 2018.

[106] O. Wahltinez and others, “COVID-19 Open-Data: curating a fine-grained, global-scale data repository for SARS-CoV-2,” 2020.

[107] R. Zheng, Y. Xu, W. Wang, G. Ning, and Y. Bi, “Spatial transmission of COVID-19 via public and private transportation in China,” Travel Med. Infect. Dis., 2020.

[108] M. Hu et al., “The risk of COVID-19 transmission in train passengers: an epidemiological and modelling study,” Clin. Infect. Dis., 2020.

[109] “Airline On-Time Statistics,” Bureau of Transportation Statistic, 2020. [Online]. Available: https://www.transtats.bts.gov/ONTIME/.

[110] N. Y. S. D. of H. New York State Task Force on Life & the Law, “VENTILATOR ALLOCATION GUIDELINES,” 2015.

[111] P. Maas et al., “Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery.,” in KDD, 2019, vol. 19, p. 3173.

[112] A. Aktay et al., “Google COVID-19 community mobility reports: Anonymization process description (version 1.0),” arXiv Prepr. arXiv2004.04145, 2020.

[113] “Baidu Migration data,” Baidu Migration, 2020. [Online]. Available: https://qianxi.baidu.com/.

[114] J. Elflein, “Topic: Physicians,” Statista. .

[115] “COVID-19 Capacity Predictor,” Definitive Healthcare, 2020. [Online]. Available: https://www.definitivehc.com/resources/covid-19-capacity-predictor.

[116] “No Title.” .

[117] “Data from PPE Shortage Index shows critical need for personal protective equipment in U.S.,” Get Us PPE. Jan-2021.

[118] Z. Levin, K. Choyke, A. Georgiou, S. Sen, and P. Karaca-Mandic, “Trends in Pediatric Hospitalizations for Coronavirus Disease 2019,” JAMA Pediatr.

[119] “US Hospital Facility Bed Capacity Map,” CovidCareMap. .

[120] R. M. Anderson, J. A. Crombie, and B. T. Grenfell, “The epidemiology of mumps in the UK: a preliminary study of virus transmission, herd immunity and the potential impact of immunization,” Epidemiol. Infect., vol. 99, no. 1, pp. 65–84, 1987.

[121] H. Sjödin et al., “COVID-19 healthcare demand and mortality in Sweden in response to non-pharmaceutical mitigation and suppression scenarios,” Int. J. Epidemiol., vol. 49, no. 5, pp. 1443–1453, 2020.

[122] “COVID-19 Government Measures Dataset,” ACAPS. Dec-2020.

[123] M. Lyu and R. Hall, “Dynamic Modeling of Reported COVID-19 Cases and Deaths with Continuously Varying Case Fatality and Transmission Rate Functions,” medRxiv, 2020.

[124] Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, “The reproductive number of COVID-19 is higher compared to SARS coronavirus,” J. Travel Med., vol. 27, no. 2, pp. 1–4, 2020, doi: 10.1093/jtm/taaa021.

[125] H. P. Gavin, “The Levenburg-Marqurdt Algorithm For Nonlinear Least Squares Curve-Fitting Problems,” Duke Univ., pp. 1–19, 2019.

[126] M. L. Holshue et al., “First case of 2019 novel coronavirus in the United States,” N. Engl. J. Med., vol. 382, no. 10, pp. 929–936, 2020, doi: 10.1056/NEJMoa2001191.

[127] G. Chowell, J. M. Hyman, L. M. A. Bettencourt, and C. Castillo-Chavez, “The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends,” Math. Stat. Estim. Approaches Epidemiol., pp. 1–363, 2009, doi: 10.1007/978-90-481-2313-1.

[128] P. Van Den Driessche, “Reproduction numbers of infectious disease models,” Infect. Dis. Model., vol. 2, no. 3, pp. 288–303, 2017, doi: 10.1016/j.idm.2017.06.002.

[129] Z. Fu, G. Liu, and L. Guo, “Sequential quadratic programming method for nonlinear least squares estimation and its application,” Math. Probl. Eng., vol. 2019, 2019.