Analytics and theoretical studies of complex systems and their applications in epidemic models

by

Kaiming Bi

B.S., Northeastern University, 2015

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Industrial Manufacturing Systems Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

Through human history, infectious diseases are among the top of the unintentional causes of human death worldwide. In the near several centuries, massive efforts have been done to control and prevent the spreading of infectious diseases. However, infectious disease is still one of the top 10 killers that seriously threaten the health of people in the 21st century. On the other hand, the development of computer hardware, big data, and algorithms science provides a new approach to research infectious diseases. Computer-based simulation can validate the epidemic control using mathematical models instead of the real-world experiment. Big data technology can benefit the epidemic modeling by providing more disease information. Algorithms are able to design the control strategies by scientific calculation rather than empiricism. Hence, the goal of this research is to study the epidemic control strategies by using the modeling, analysis, simulation, and optimization technics.

To better discussing the epidemic control strategies, this research studies the modeling of vector-borne diseases. A partial differential equation model with age structure in human infections is introduced to describe the transmission of Zoonotic Visceral Leishmaniasis. A closed population dynamic system is introduced to study the prevention of Zika Virus. An agent-based model is presented to study emotion transmission during the epidemic like the 2009 flu pandemic. In this dissertation, analysis methods like sensitivity, stability and time series analysis are widely applied to further research the established models.

The major contribution of this research is developing the new methodology of numerical epidemic control. The Pontryagin maximum principle-based optimal control algorithm is studied to control the Zika Virus. Moreover, an innovative heuristic algorithm based method is proposed to solve the optimal control problem with the highly nonlinear objective function. This

dissertation also introduces evidence data based optimal control method, which trained the neural network with epidemic data to control the current prevalence. The last but not least, the simulation is used to predict the future epidemic and verify the designed control strategies.

Analytics and theoretical studies of complex systems and their applications in epidemic models

by

Kaiming Bi

B.S., Northeastern University, 2015

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Industrial Manufacturing Systems Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Chih-Hang Wu

Copyright

Abstract

Through human history, infectious diseases are among the top of the unintentional causes of human death worldwide. In the near several centuries, massive efforts have been done to control and prevent the spreading of infectious diseases. However, infectious disease is still one of the top 10 killers that seriously threaten the health of people in the 21st century. On the other hand, the development of computer hardware, big data, and algorithms science provides a new approach to research infectious diseases. Computer-based simulation can validate the epidemic control using mathematical models instead of the real-world experiment. Big data technology can benefit the epidemic modeling by providing more disease information. Algorithms are able to design the control strategies by scientific calculation rather than empiricism. Hence, the goal of this research is to study the epidemic control strategies by using the modeling, analysis, simulation, and optimization technics.

To better discussing the epidemic control strategies, this research studies the modeling of vector-borne diseases. A partial differential equation model with age structure in human infections is introduced to describe the transmission of Zoonotic Visceral Leishmaniasis. A closed population dynamic system is introduced to study the prevention of Zika Virus. An agent-based model is presented to study emotion transmission during the epidemic like the 2009 flu pandemic. In this dissertation, analysis methods like sensitivity, stability and time series analysis are widely applied to further research the established models.

The major contribution of this research is developing the new methodology of numerical epidemic control. The Pontryagin maximum principle-based optimal control algorithm is studied to control the Zika Virus. Moreover, an innovative heuristic algorithm based method is proposed to solve the optimal control problem with the highly nonlinear objective function. This

dissertation also introduces evidence data based optimal control method, which trained the neural network with epidemic data to control the current prevalence. The last but not least, the simulation is used to predict the future epidemic and verify the designed control strategies.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to express my sincerest thanks to my advisor Dr. Chih-Hang Wu for his valuable guidance, understanding, and support throughout my career as a graduate student. Actually, He has not only taught me the knowledge to be a good researcher, but also taught how to be a real man. Dr. Wu, thank you for this incredible mentorship. I will never forget your advices. These advices will inspire me to be more professional in the future career.

I would like to thank my Ph.D. supervisory committee members Dr. David Ben-Arieh, Dr. Ashesh Kumar Sinha, Dr. Roman Reddy Ganta, and Dr. Juan Du for their time and guidance during my graduate studies. Their insightful comments and suggestions strengthened this dissertation's research substantially.

I thank Dr. Bradley Kramer for his constant financial support and opportunity to develop my research and professional skills. I also offer my sincere appreciation to all faculty and staff from the department of Industrial and Manufacturing Systems Engineering at Kansas State University for their support, knowledge, and assistance.

Last but not least, I would like to thank my family, friends, and colleagues who helped me during my academic and professional journey. Their encouragement, support, and willingness also made this work possible.

Chapter 1 - Introduction

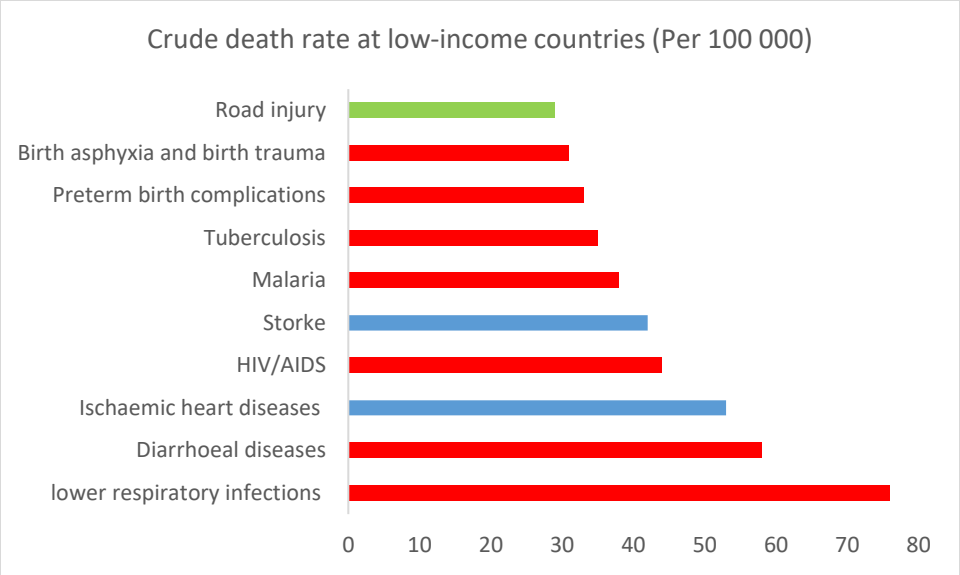
1.1 Research background and significance

Infectious diseases are diseases caused by biological agents, like viruses, bacteria, parasites, fungi, and prions, which can be transmitted between humans [1]. The earliest recorded epidemic can be traced back to the Athens era [2]. Before the middle ages, most of the epidemics are caused by the plague. These include the famous Black Death epidemic, which killed 30%-60% of the total population of Europe [3]. Between the 16th and 18th centuries, the recorded epidemics reflected more types of infectious diseases like measles, smallpox and yellow fever [4]. In the meantime, the locations of these epidemics were not only restricted to Europe and the Mediterranean. More epidemics were founded in North America, Asia, and Australia. In the 19th century and early of the 20th century, cholera and influenza became the two of the most fatal infectious disease. The cholera pandemics, flu pandemics, and famous Spanish flu killed millions of people worldwide[5]. From the 1960s, HIV/AIDS became the most dangerous disease. The complete cure for HIV has not established by 2019 [6].

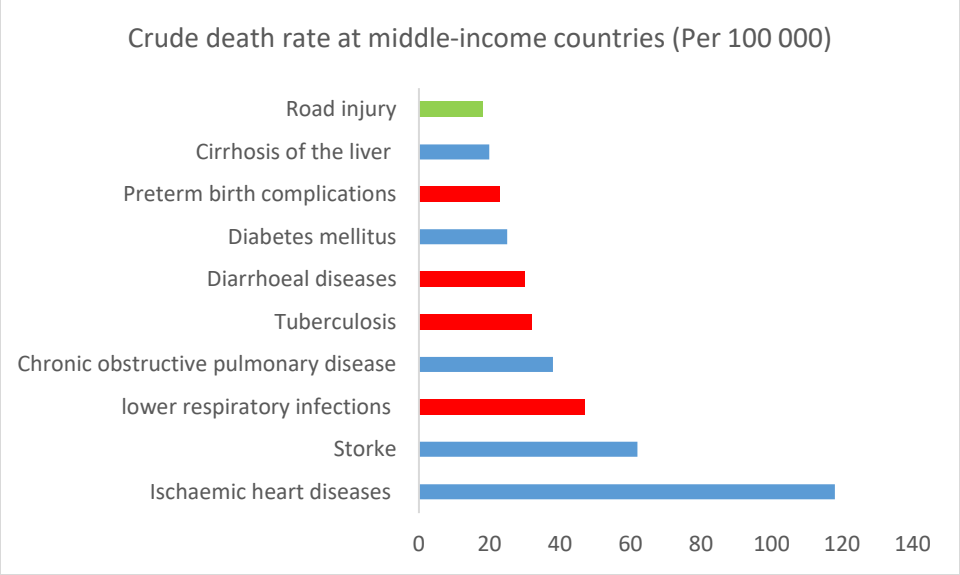
The public health intervention was widely applied to prevent people from infectious disease. As early in the 17th century, several European cities started to adopt the isolation of the ill and quarantine of travelers [7]. In the 18th-19th century, "The great sanitary awakening", the identification of filth as a cause of disease and the ensuing embrace of cleanliness, was spread in Europe and North American [8]. At almost the same time, voluntary general hospitals were established to take care of the infected people. Benefited by the development of bacteriology, in the late 19th century, state and local health departments in the United States began to establish laboratories and institutes. These laboratories researched the detection and control of bacteria in water systems, diagnosis of disease in individuals and vaccination developments [9]. By the

early 20th century, the United States already established complete disease registries and analysis systems. The federal also provided more funds to support the National Institute of Health and Center for Disease Control in infectious disease research.

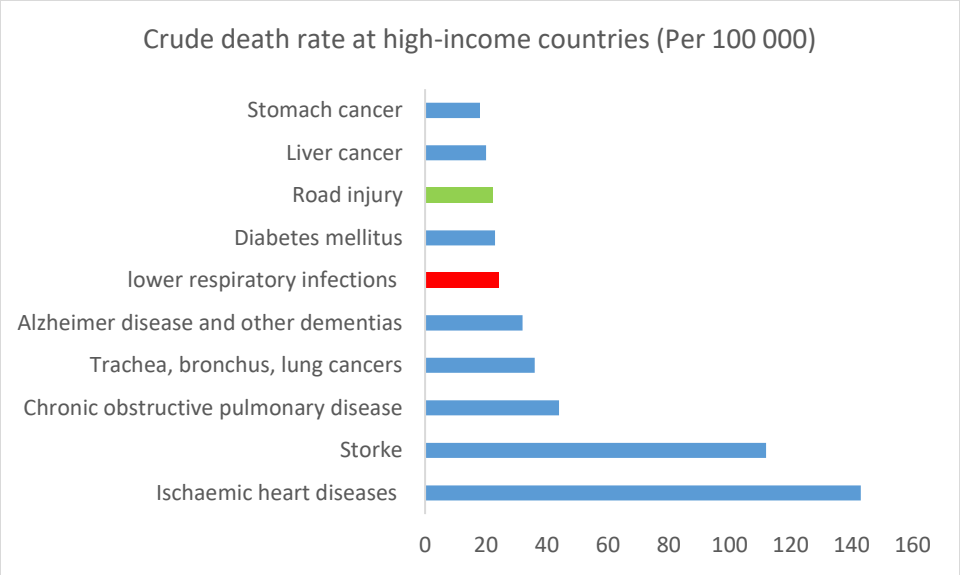
Even using centuries of successful prevention efforts, Infectious disease is still one of the top killers that seriously threaten the health of the global community. Figure 1.1 shows the top 10 causes of deaths in different levels of income countries in 2016. By comparing the situations in these four kinds of countries, we know that the higher income, the less negative effects of infectious diseases. The seven of the top 10 causes of deaths are infectious diseases in low-income countries. However, for the high-income countries, these communicable diseases exist only once in the top 10 causes. Subtracting out the impact of population density, geographical conditions, and living habits, the input level of epidemic control strategies and treatments also determine the performance of the diseases [10]. For developing countries, public investments have to focus more on the basic requirements of food, living, and equipment. For developed countries, governments have more financial ability to control epidemic transmission. Therefore the question is, especially for the developing countries, how to develop the strategies to balance the control performance of the epidemic and the cost of disease control? Alternatively, in other words, how to use a limited budget to minimize the harmful effect of the epidemic. Therefore, the main goal of this dissertation is to study the infectious disease epidemic using the modeling and simulation approaches. Optimization and machine learning analytical technics were used to design the control strategies aimed at the epidemic.


Crude death rate at low-income countries (Per 100 000)
Road injury
Birth asphyxia and birth trauma
Preterm birth complications
Tuberculosis
Malaria
Storke
HIV/AIDS
Ischaemic heart diseases
Diarrhoeal diseases
lower respiratory infections
0
10
20
30
40
50
60
70
80

(a)


Crude death rate at middle-income countries (Per 100 000)
Road injury
Cirrhosis of the liver
Preterm birth complications
Diabetes mellitus
Diarrhoeal diseases
Tuberculosis
Chronic obstructive pulmonary disease
lower respiratory infections
Storke
Ischaemic heart diseases
0
20
40
60
80
100
120
140

(b)



(c)

Figure 1. 1. Top 10 causes of deaths in 2016: (a) at low-income countries; (b) at middle-income countries; (c) at higher-income countries, where red represents the communicable diseases, blue represents the non-communicable diseases, green represents the injuries [11].

In the 21st century, the prevention control of infectious disease is still a worthwhile research issue to be explored. Different from the previous studies, many new technologies are contributed to the epidemic control area in the 21st century. For the last decade, the fast advancements of the internet, data science, and computer hardware enable us to take more advantages from the abundant volume of epidemic data. These data can be used to build accurate models, which can help people to better understand infectious disease transmission and predict the trends of future epidemics. Moreover, the epidemic model can be used to design the prevention or mitigation strategies, which can limit the negative influence of infectious diseases. In the past, most of the epidemic control and intervention strategies are based on policies established from empirical or intuitions. If the current prevalence is different from the historical epidemics, the empirical strategies may not be effective. In addition, the increasing computing

power benefits the computer-based models and sophisticated simulate to predict the behaviors of complicated disease transmission. The epidemic simulation can also be considered as the zero-cost experiments to verify the designed epidemic prevention or mitigation strategies. The epidemic predictions, then, can be used toward future epidemics.

With the above motivations, this dissertation discusses the topics in epidemic modeling, numerical control, and simulation. Four different models are included for different types of infectious diseases. Three types of optimal control strategies are designed to prevent or reduce epidemic transmission based on the established models. Then, the simulation experiments are used to verify the effects of the assumed control strategies and predict future epidemics.

1.2 Research motivation and objective

Several studies have been conducted to contribute to infectious disease modeling, decision making, and epidemic control. Dasbach et al. reviewed recent modeling works in a dynamic system area and summarized the potential benefits of vaccinating the susceptive population [12]. Nianogo and An reviewed recent contributions by using agent-based models to study the epidemic in the large-scale population [13, 14]. Based on the existed contributions, Real and Peterson proposed a spatial dynamic model the epidemic transmission between different locations [15, 16]. Holt et al. applied the cross-species epidemic model based on the general disease model for humans [17]. Castillo-Chavez et al. proposed the theoretical support for using the partial differential equation to study the age-structure for the epidemic model [18]. Frias-Martinez et al. studied the agent-based epidemic model with both disease and information transmission [19].

In addition to epidemic modeling, many researchers used existed models to study the control strategy and limit disease transmission. Zhou et al. studied the control of the SARS

epidemic by adjusting the parameters of the discrete epidemic model [20]. Wang et al. discussed the pest-control as an impulsive control strategy and researched the corresponding periodic conditions [21]. Sen et al. utilized a feedback control model to study the vaccination control strategy for the general SEIR epidemic model [22]. Moreover, the disease optimal control strategy is combined with disease control and optimization with a certain object. Zaman and Kar studied the optimal control for the general SIR epidemic model and applied this method in the vaccination allocation strategy [23, 24]. Chen et al. extended the existed optimal control method into heterogeneous networks [25]. Lemos-Paiao et al. researched optimal control treatment for cholera through quarantine [26]. Lashari et al. studied both the vector-reduction and human protection strategies for vector-borne disease by using optimal control [27]. Okosun et al. analyzed the treatment and vaccination with waning immunity of malaria disease and found the optimal strategies using the determinate objective function [28].

Even though the previous researches provided some contributions in epidemic modeling, the specific compartment models for infectious diseases with multiple species are still less studied. The current compartmental modeling is mainly used for the person-to-person direct contact diseases, like droplet spread and sexually transmitted diseases [29]. However, these models cannot be used for indirect contact diseases, which include animal-to-person contact, animal reservoirs, and insect bites. Also, the model with age-structure for human infections of vector-borne diseases is deficient. The human infections for some vector-borne diseases have a significant difference in the age-structure. For example, there are more young infections than adult infections for the zoonotic visceral leishmaniasis [30].

Analogously, the current contributions in epidemic numerical controls are still limited. There is a large portion of the numerical control studies without using the optimization thoughts.

Some of the studies used the feedback control, which assumed the relationship between the prevention with the prevalence variable [22]. Some others optimized the settings of each system parameters independently, instead of systematically considering the epidemic control [31].

There are several numerical optimal control algorithms, which applied some nonlinear programming methods like Pontryagin's Maximum Principles [32]. However, these methods can only solve the optimal control problem with the convex objective and constraint functions for dynamic models. There are very few works that utilized the historical epidemic data to find patterns by machine learning and artificial intelligence algorithms. No methodologies are using the patterns learned from the historical data to design the disease prevention strategies for the future.

Hence, this research aims to develop models with different characteristics and make decisions to have better controls of future epidemics, which includes three main research tasks as follows:

Task 1, modeling the disease with different characteristics: Using one uniform framework to model different infectious diseases is unreasonable since infectious diseases have different characteristics. To better reflect the epidemics in the real world, different components were added into the general diseases model with different modeling technics. During the task 1, the specific tasks are as follows:

1. Study disease models with different types of carriers and reservoirs.
2. Develop the disease model with age structure to reflect the difference in susceptive abilities.
3. Discuss the disease with information and emotion transmission through the population.

Task 2, decision making by analyzing the model to control the epidemic transmission: There are multiple parameters in the disease model. Some of these parameters are controllable. Finding the appropriate value for these controllable parameters can help us control infectious diseases better. During the task 2, the specific tasks are as follows:

1. Utilize the sensitivity analysis and the value of the basic reproduction number to discuss the appropriate values for the controllable parameters.
2. Study the equilibriums of the disease model after the system reaching the static state.
3. Discuss the potential bifurcation points which may change the stability status of the equilibriums.

Task 3, Optimize the control strategy with using the optimization methods and data: With a certain object, the optimal control strategy can be designed to minimize the negative influence of the epidemic. Unlike other optimal control methods in disease modeling area, this task mainly focuses on the control strategy in time series. During task 3, the specific tasks are as follows:

1. Design the optimal control strategy in time series for the epidemic dynamic system using Pontryagin's maximum principle.
2. Use an appropriate heuristic algorithm to solve the complicated optimal control problem with highly nonlinear objective functions.
3. Study the evidence-based optimal control strategy utilizing the historical epidemic data.

1.3 Research map and contributions

This research proposed several methodologies to control infectious disease epidemics using mathematical models and numerical computations. In addition, this research considers both descriptive and analytical studies in our research plans. The descriptive studies include modeling

and simulation methods, which are utilized to describe and simulate the disease transmission. The analytical studies include model, analyze and optimal control and intervention strategies for balancing the magnitude of the ongoing epidemic and the associated costs for interventions. Figure 1.2 shows a research map that describes the research objectives, research methodologies, and potential contributions.

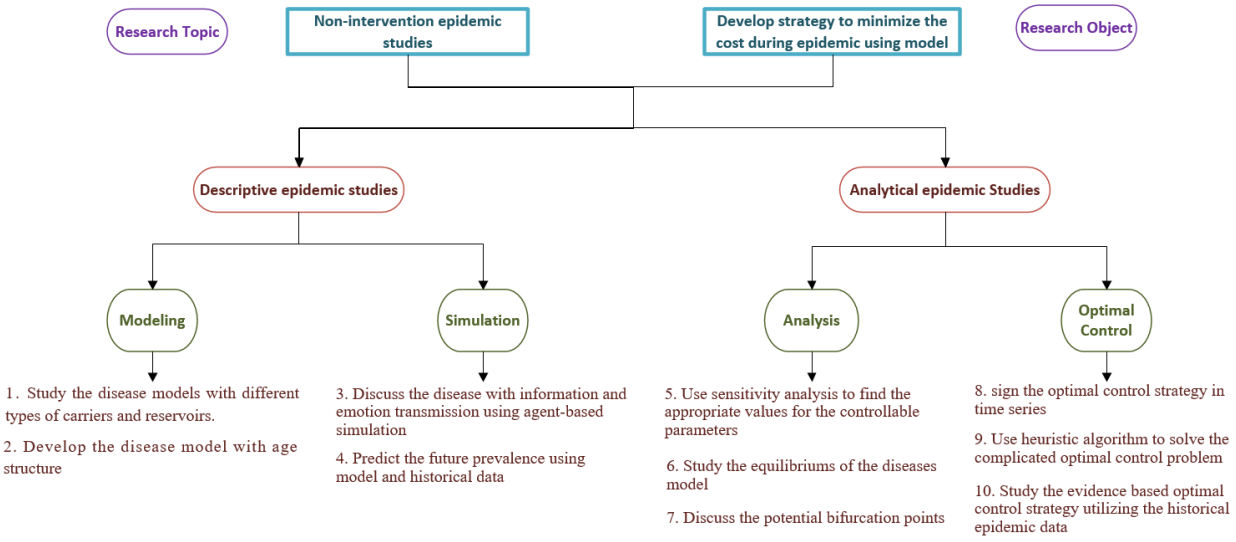


Figure 1. 2 Research Map of the dissertation

In the modeling part, this dissertation plans to develop more disease models with different species evolved, like mosquitos, sandflies, and dogs. Serval tropical diseases like Zika Virus, Zoonotic Visceral Leishmaniasis and malaria are expected to be modeled with the integrated transmission map. This dissertation also contributes deeper to the modeling of Zoonotic Visceral Leishmaniasis, whose infection patterns shown an obvious correlation to the patient's age groups [33]. The children and teenagers have a significantly higher infection rate compare to adults. This phenomenon can be modeled mathematically by using the partial differential equation-based model.

In the simulation part, this dissertation proposed an agent-based simulation to research emotion transmission during the epidemic. The infectious disease can be transmitted through the contact networks of each individual. Also, the information regarding the ongoing epidemic can be disseminated through internets, TV, and newspaper. In the meantime, the awareness of individuals would most likely be affected by the epidemic information. The change of awareness would then result in certain behavior changes. This dissertation is to model and mimic the above process using agent-based simulation. In this dissertation, the agent-based simulation is used to predict the outcomes of future epidemics. This prediction may utilize the disease model built in the modeling part with historical epidemic data.

In the analysis part, our goal is to analyze the established diseases model further. Assumed the model is correct, then what kind of control or interventions can be carried out to mitigate various negative impacts of the diseases. The first type of analysis method is sensitivity analysis, which is expected to find the appropriate values for the parameters of the disease models. The analytical method is the equilibrium analysis, which contributes to finding the possible steady states at the end of the epidemic outbreak. Since several factors may affect the equilibrium points, the equilibrium can lower the infection number at the steady-state by changing these factors. The last type of analysis method is the bifurcation analysis, which is used to determine the stability of the equilibrium points. This analysis method may help us to predict if the current epidemic would eventually vanish or not at the end.

The optimal control part is one of the most significant contributions of this dissertation, which attempts to optimize the control strategies with time series regulated complex systems. By using the dynamic control strategies for infectious disease, both the infection costs and control costs will be minimized. The optimal control is also expected to be triggered only when

necessary. If it is not necessary, the optimal control strategy should be able to shut down the interventions to save the associated costs. However, the traditional optimal control models in the epidemic area are mostly based on Pontryagin's maximum principle, which requires the convexity conditions on the objective and constraint functions to be satisfied. Therefore, this dissertation designs a new memetic algorithm-based method to solve the epidemic optimal control problem with complicated objective functions. This dissertation is also expected to utilize various machine-learning algorithms to take advantage of evidence data, which contributes to the design of the evidence-based optimal controls.

1.4 Dissertations Outlines

The reset of the dissertation is organized into seven characters: Chapter 2 provides a review of the current state-of-the-art methodologies that inspired our studies in this dissertation. These methodologies included modeling, system analysis, control theories, and system optimization. Chapter 3 is a review paper published in the Journal of Biomedicine and Biotechnology [34], which reviewed recent modeling and disease control contributions in the Visceral Leishmaniasis area. Chapter 4 is a published research paper in Chaos Soliton & Fractals [30], which discussed a new Zoonotic Visceral Leishmaniasis dynamic transmission model with age-structure. Chapter 5 is a published journal article in Communications in Nonlinear Science and Numerical Simulation [35], which proposed a new method using a memetic algorithm to solve the optimal control problem of Zika Virus epidemic. Chapter 6 is an ongoing working paper, which researched the methodology using a neural network to analyze the evidence database of historical epidemic control. The trained neural network could be used to design an evidence-based impulse control for the future epidemic. Chapter 7 is from a published research paper in Computers & Industrial Engineering [36], which studied the potential learning and

forgetting processes of the people during the epidemic by agent-based modeling. These human behaviors may affect the performance of disease transmission. Chapter 8 summarizes the main conclusion and contributions of this dissertation and discussed the potential future works.

Chapter 2 - Overview of the methodologies in epidemic modeling

Since the 17th century, infectious disease modeling has been studied extensively [37]. Many methods have been used to understand and predict the epidemic. In the modeling of the epidemic transmission, the compartmental model and agent-based model are two types of general methods that have been used most. The compartmental model is usually investigated through ordinary differential equations to model the epidemic transmission in a given population [38]. Agent-based model is utilized to research each individual and corresponding interacts with their environments during the epidemic [39]. In the meanwhile, network modeling, game theory, and stochastic process modeling have been widely used to describe the epidemic transmission [40]. Also, analysis technics are combined with the mentioned models to deeper understand the epidemic. For instance, the equilibrium and bifurcation analysis has been used to research the stability of the compartment models [41]. Sensitivity and statistic analysis assist the value settings for parameters [42].

Designing the numerical control strategy is another main object of epidemic modeling. The real-world interventions like vaccination, isolation, and hospitalization can be joined into the established model. Optimal control is one of the most popular control designing method, which is introduced by Pontryagin [43]. Optimal control methods utilize the characteristics of the compartment model to design a time series of control strategies, which can minimize the total cost during the epidemic. Other methods used in epidemic control included feedback state control, machine learning and heuristic algorithms [22, 44, 45]. Numerical simulation can be considered as a “what-if” experiment to verify the effectiveness of the control strategies. Since the real-world control may cause a high cost with the irreversible results, the numerical simulation can test the designed control strategy as the experiment before the implementation.

This section is organized as follows: the compartment model and agent-based model are respectively introduced in sections 2.1 and 2.2. Section 2.3 presents the equilibrium and bifurcation analysis methods in epidemic system. Section 2.4 illustrates the optimal control methods and the applications in the infectious disease model. Section 2.5 summarizes the selected machine learning methods and heuristic algorithms used in this dissertation. Section 2.6 describes the numerical simulation and corresponding platforms.

2.1 Compartmental model

The compartmental model is one of the most popular techniques in the modeling area of infectious disease [46]. The compartmental model is usually used to track the epidemic transmission in a population. In a compartmental model, the population is divided into different groups based on the infectious status. The simplest and commonest compartmental model in the infectious disease area is the SIR model.

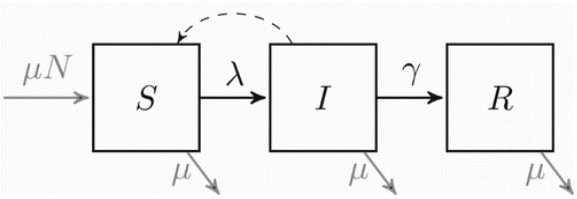


Figure 2. 1 Flow chart of the SIR model

SIR model was introduced by Kermack and McKendrick in 1927 [47]. The population in the SIR model is consisted of there compartments, where $S(t)$ represents the susceptible population, $I(t)$ represents the infected population and $R(t)$ represents the recovered population. Typically, the SIR model is presented by using the ordinary differential equations as follow:

$$\frac{dS}{dt} = \mu N - \frac{\beta IS}{N} - \mu S \tag{2.1}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I \tag{2.2}$$

$$\frac{dR}{dt} = \gamma I - \mu R \tag{2.3}$$

Where $N(t) = S(t) + I(t) + R(t)$ represents the total studied population; β is the overall transmission rate of a susceptive person; γ is the recovery rate of an infected person; μ is both the birth rate and death rate without considering the factors of the current disease. SIR model assumes that the birth rate always equals the death rate, to keep the population as a constant. For some infectious diseases, the epidemic outbreak period is far less than the human lifespan. Meaning the birth and death can be ignored in the SIR model, μ can be set as zero for these specific cases.

Since the SIR model is generally applicable in the epidemic, it can describe the infectious disease transmission in most cases. However, the SIR model cannot cover for all types of diseases. For instance, some infectious disease doesn't exist the immunity in the recover population. In this case, people usually consider the SIRS or SIS model, where the recovery population will be sent back as the susceptive population [48]. For the infectious disease which includes a significant incubation period, the exposed compartment has consisted of the SIR model into the SEIR model [49]. For some vector-borne diseases, revised SIR and SEIR models are used to illustrate the complex transmission routes between humans with other species [50].

2.2 Agent-based model

Agent-based model is a type of model that can simulate the actions and interactions of agents to predict the appearance of complex phenomena [51]. Typically, an agent-based model includes 4 key components: agents, properties, environment and action rules. In most cases, the agent-based model is combined with the Monte Carlo simulation. In recent years, benefited by the hardware upgrade in the computer industry, more and more researchers start to use large-scale simulations to implement the complex agent-based models.

Like the compartmental model, the agent-based models are enormously used to model infectious disease transmission and intervention control. Therefore, researchers usually utilize agent-based models to create general disease dynamics and simulate a specific outbreak. Beyond that, the agent-based model has two general advantages over the compartmental model. First, the agent-based model can provide more details of the interactions between individuals, and individual interactions with local environments [52]. The agent-based simulation also enables us to track the behavior of each individual during the epidemic. Second, the agent-based model doesn't rely too much on the data. It is possible to capture the feature of a dynamic system with only a few data [53].

In the agent-based model of infectious disease, the population can be divided based on the infectious status, like susceptive, infected and recovered. This infectious status can be updated during the simulation. The environment is the transmission map or the contact network in the studied population. Different action rules like isolation, self-protection, and hospitalization can be modeled for each agent to switch their behavior during the epidemic period.

2.3 Stability and bifurcation analysis

One of the most important concerns about infectious disease is the ability to invade a population [54]. For an epidemic system, the stability analysis can prove if the studied infectious disease has a natural demise or not. Almost all epidemic compartmental model has a disease-free equilibrium at which the population remains in the absence of disease. To determine the stability of the system at the disease-free equilibrium, the value of the threshold parameter R_0 is discussed. R_0 is also known as the basic reproduction number, which is defined as 'the expected number of secondary cases produced, in a completely susceptive population, by a typical infective individual' [55]. If $R_0 < 1$, then on average the infected produces of this infected

person is less than its infectious period, meaning the infectious cannot grow. Conversely, if $R_0 > 1$, on average each infected individual produces are more than one, meaning the disease can invade the population.

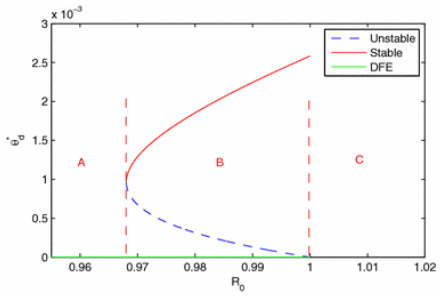


Figure 2. 2 Bifurcation illustration of the Zoonotic visceral leishmaniasis dynamic system [50] A: The disease-free equilibrium is stable; B: The disease-free equilibrium and one endemic equilibrium are stable while the other endemic equilibrium is unstable; C: The disease-free equilibrium is unstable while the endemic equilibrium is stable

Another type of equilibrium is the endemic equilibrium, which is the globally stable point with the disease persists in the population [56]. The question of the uniqueness and global stability of the endemic equilibrium is valuable, especially when the R_0 is greater than one. Because the stability of the endemic equilibrium determines if the disease can persist in the epidemic system for the long term. However, the stability status of the endemic equilibrium varies with some values of the parameters. This phenomenon is called bifurcation, which can lead a disease-free equilibrium to an endemic equilibrium [57]. Consider Figure 2.2 as an instance, there are no endemic equilibriums in section A. However, there are two different endemic equilibriums in section B, one is stable and the other one is unstable. A potential significance to study the bifurcation is to design the intervention on the selected parameters. This intervention can change the stable endemic equilibrium into an unstable status. Thus, the disease invasion can be stopped at the researched population.

2.4 Optimal control strategies

Optimal control theory solves the continuous control problem in a time series, typically for a dynamic system with a certain objective function [52]. The original concept of optimal control was presented by Lev Pontryagin [43]. His maximum principle states the necessity for any optimal control along with the optimal state trajectory to solve the Hamiltonian system. Particularly, when the objective and constraint functions satisfy the convexity conditions, the necessity becomes sufficient [58]. Optimal control theory has been applied to multiple application areas like aerospace, power systems and finance [59-61]. In the epidemic area, optimal control theory has been used to develop intervention strategies for many diseases like HIV, dengue, and malaria [62-64]. Typically, in the epidemic optimal control problem, there is a dynamic system with control variables. The epidemic transmission under the intervention can be described by the dynamic system as follow:

$$\dot{x}_i = f_i(x) + g_i(x)u_i \tag{2.4}$$

$$\dot{x}_j = f_j(x) \tag{2.5}$$

Eq. (2.4) and (2.5) respectively represent the controllable and uncontrollable parts of the dynamic system. Where x_i represents the controllable system variables, u_i represents the continuous control variables, $f_i(x)$ and $g_i(x)$ represent the smooth nonlinear system functions; x_j represents the uncontrollable system variables, $f_j(x)$ represents the smooth nonlinear system function.

$$J_c = \int_0^T C_1\big(x_I(t)\big) + C_2(u_c(t))dt \tag{2.6}$$

Eq. (2.6) defined the objective function over the epidemic period. Where $C_1\big(x_I(t)\big)$ represents the cost due to the infections at time t, $C_2(u_c(t))$ represents the cost due to the interventions at time t. In the view of the epidemic control, the optimal control variable is expected to minimize the objective function.

$$Min\, J_c\big(u_c(t)\big) \tag{2.7}$$

To solve the optimal control problem using the Pontryagin maximum principle, the Hamilton function should be first defined as follow:

$$H = C_1\big(x_I(t)\big) + C_2\big(u_c(t)\big) + \sum_{i=1}^{n} \lambda_i(t)\left(f_i\big(x(t)\big) + g_i\big(x(t)\big)u_c(t)\right) + \sum_{j=n+1}^{n+m} \lambda_j(t) f_j(x(t)) \tag{2.8}$$

where $\lambda(t) = [\lambda_1(t), \lambda_2(t), \dots, \lambda_{n+m}(t)]^T \in \mathbb{R}^n$ denotes the vector of adjoint functions $\lambda_i(t)$, $i = 1,2, \dots, n+m$. The optimal values of adjoint variables are defined by

$$\dot{\lambda}_i^*(t) = -\frac{\partial H}{\partial x_i(t)} \tag{2.9}$$

In addition, at the terminal time T, adjoint functions satisfied the transversality condition

$$\lambda_i^*(T) = 0 \tag{2.10}$$

Based on the Hamiltonian minimizing condition, the continuous optimal control $u_c^*(t)$ can be computed by

$$u_c^*(t) = argmax\{ H\big(x^*(t), \lambda_i^*(t), u_c^*(t)\big)\} \tag{2.11}$$

$$\Rightarrow \frac{\partial H}{\partial u_c(t)} = C_2{}'\big(u_c^*(t)\big) + \sum_{i=1}^{n} \lambda_i(t) g_i(x(t)) = 0$$

The $u_c^*(t)$ can be rearranged as

$$u_c^*(t) = C_2{}'^{-1}(-\sum_{i=1}^{n} \lambda_i(t) g_i(x(t))), u_c^*(t) \in \mathbb{R}_{[0,1]} \tag{2.12}$$

The value of the adjoint variables can be calculated by Eq. (2.9) and (2.10), and the control variables can be calculated by Eq. (2.12) using the value of adjoint variables. In the real world problem, the Pontryagin maximum principle is expected to solve the numerical solution of the optimal control problem. Forward-Backward Sweep (FBS), shooter method and direct optimization process are three classical numerical calculation methods for the maximum principle [65].

2.5 Optimal control strategies

In epidemiology, the mathematical models are used to understand the transmission process and predict future epidemics [66]. A number of technics like R_0 calculation, equilibrium analysis and sensitivity analysis can be used to better understand the disease transmission process. However, it is tough to use these technics to predict the outcome of the future epidemic. Even though the numerical calculations can be used to predict the epidemic outcome for a one-time epoch, the predicted results still lost the epidemic outcomes during the entire transmission period. Therefore, epidemiologists have long used simulation to present the transmission process

with the model. Especially, for the agent-based model, the simulation provides the viewing angle to observe the behavior for each agent.

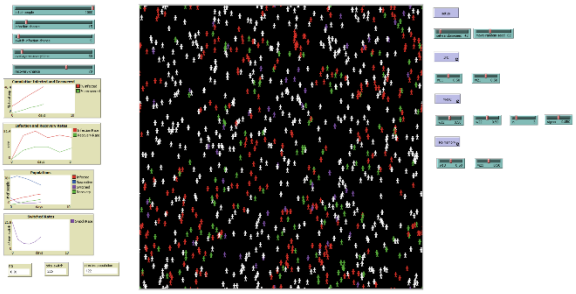


Figure 2. 3 Agent-based model illustration using Netlogo frame [67]

In the simulation for the compartmental models, many types of research utilized the toolbox of R, Matlab, and Python as a numerical platform to simulate the epidemic in a period [68]. In the agent-based simulation, there are multiple platforms have been used to simulate the disease transmission through the agent-based model. Netlogo is one of the most popular platforms in the immune system and infectious disease modeling area [13, 69, 70]. Instead of Netlogo, Mason and the Gama Platform were contributed in the agent-based simulation with disease modeling [71-73].

Chapter 3 - Current Visceral Leishmaniosis Research: A Research Review to Inspire Future Study

Chapter 3 is based on the manuscript "Current Visceral Leishmaniosis Research: A Research Review to Inspire Future Study" Published in BioMed Research International [34].

Abstract

Visceral leishmaniosis (VL), one of the deadliest parasitic diseases in the world, causes more than 50,000 human deaths each year and afflicts millions of people throughout South America, East Africa, South Asia, and Mediterranean Region. In 2015 the World Health Organization classified VL as a neglected tropical disease (NTD), prompting concentrated study of the VL epidemic using mathematical and simulation models. This paper reviews literature related to prevalence and prevention control strategies. More than thirty current research works were reviewed and classified based on VL epidemic study methods, including modeling approaches, control strategies, and simulation techniques since 2013. A summarization of these technical methods, major findings, and contributions from existing works revealed that VL epidemic research efforts must improve in the areas of validating and verifying VL mathematical models with real-world epidemic data. In addition, more dynamic disease control strategies must be explored and advanced simulation techniques must be used to predict VL pandemics.

Key word: Visceral leishmaniosis; control strategy; mathematical modeling; simulation

3.1 Introduction

Visceral leishmaniosis (VL), or kala-azar, is a protozoan disease that, second only to malaria in numbers of fatalities, afflicts millions of people worldwide [74]. VL is primarily distributed in East Africa, South Asia, South America, and Mediterranean Region, with an

estimated 50,000 to 90,000 new VL cases each year. Ninety percent of reported VL cases occur in Brazil, Ethiopia, India, Kenya, Somalia, South Sudan, and Sudan [75]. In 2015 the World Health Organization (WHO) classified VL as a neglected tropical disease (NTD) due to relatively minimal granted attention from the public, resulting high mortality rates (more than 20,000 in 2015), and endemic spreading in poverty-stricken regions around the world [76, 77].

VL is one of the most widespread human diseases, with more than 20 Leishmania species identified worldwide [78]. Unlike other common forms of leishmaniasis, such as cutaneous leishmaniasis (CL), mucocutaneous leishmaniasis (ML), and post-kala-azar dermal leishmaniasis (PKDL), VL symptoms usually occur internally [74], meaning VL is more difficult to detect and cure than other leishmaniasis. Based on different kinds of susceptible species, VL can be classified as anthroponotic visceral leishmaniosis (AVL) or zoonotic visceral leishmaniosis (ZVL). AVL, which is transmitted between humans via vector carriers, is primarily caused by *L.donovani* throughout East Africa and the Middle East, especially Sudan, Somalia, Yemen, and Saidi Arabia [77]. Since most VL vaccinations for humans are still in research or minimally effective [79], AVL preventions include only early detection and treatment or use of long-term insecticide nets. ZVL, however, is transmitted between humans and other mammals, such as dogs. *L.donovani*, *L.infantum*, and *L.archibaldi* can result in ZVL, with specific concentrations in East Africa, South America, the Mediterranean basin, and South Asia [79]. Because dogs are the most common mammal carriers of ZVL, ZVL control strategies such as culling dogs, dog vaccinations, and use of insecticide collars are more prevalent than any AVL control strategies [80-83].

The use of mathematical models to describe and predict epidemic transmissions has become a recent trend in disease research area. Mathematical models intuitively exhibit complex

VL transmission processes, and they measure variables and system parameters to reveal VL spreading dynamics and related dominating factors. Rapid advancements in computer technologies have resulted in computer-aided simulation that helps mathematical models directly predict future VL prevalence. Using results from model analysis, parametric estimation, and simulation experiments, researchers can study and anticipate disease transmission dynamics and identify disease control strategies to fight a VL pandemic. Consequently, an increasing number of studies have focused on mathematical modeling and corresponding analysis for VL disease dynamics. Approaches used in these studies can be generally categorized as system dynamic models, including ordinary differential equation (ODE) or partial differential equation (PDE) models, as well as statistic models, or machine learning models. The main contributions and results are concentrated in precise prediction tested by validation, determining the key parameters by sensitive analysis and analyzing the bifurcation point of the disease reproduction number.

A well-defined mathematical model can be used to develop disease control strategies that are ascertained by solving the mathematical model or using numerical experiments. Numerical control strategies are robust and reliable approaches because potential bias from empirical data is not included. Conversely, implementation of real-life control strategies can be cost prohibitive, irreversible, and difficult to apply in a large-scale format, especially for developing countries that lack public health resources. However, computer-aided simulations that compare possible control strategies derived from a mathematical model can be carried out, and they are relatively inexpensive and can be performed repeatedly to examine system sensitivity and determine optimal control parameters. Almost half of corresponding research used mathematical modeling approaches to study potential disease control strategies, including dog culling, use of insecticidal

dog collars, vector controls, and insecticide spraying strategies. Using optimal control, parametric analysis, or stochastic control methods, research results provided well-developed guidelines for disease control centers to prevent or mitigate a VL pandemic.

The rest of this paper is organized into comprehensive sections. Section 2 introduces current worldwide VL pandemic situations, and Section 3 presents VL control strategies in existing literature and recent breakthroughs in this field. Section 4 reviews papers on VL mathematical modeling and summarizes new developments and significant contributions since 2013. Section 5 reviews papers on control strategies and the use of numerical simulation, and

Section 6 concludes the paper and suggests possible areas for future VL pandemic research. The research tree of this paper is shown in Figure 3.1.

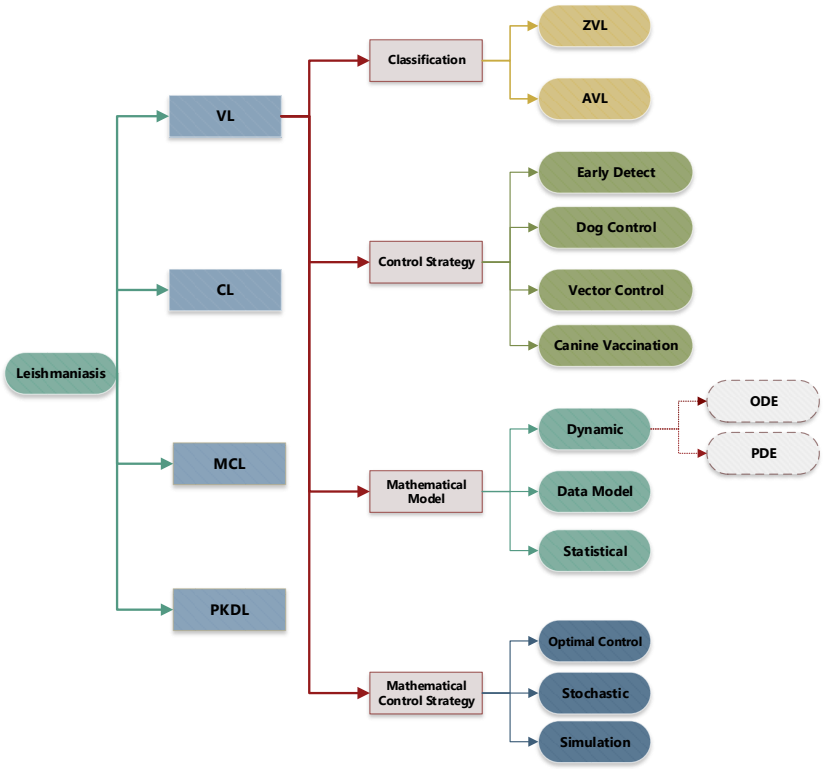


Figure 3. 1 Research tree for chapter 3

3.2 VL Pandemic Recent History

Pigott *et al.* collected and summarized prevalence data reports of CL and VL from 1960 to 2012 [84]. According to their summarized database, the worldwide VL pandemic has affected at least 55 countries and territories (Figure 3.2), and another 45 countries have reported unspecified and borderline VL cases. Affected countries are located primarily in Latin America, the Mediterranean basin, Northeastern Africa, and Asia. Moreover, VL outbreaks historically occur most often in developing or agricultural countries because citizens are more likely to be in contact with disease transmission vectors such as dogs and mosquitoes [85]. Worse still, VL traps many families in these developing countries into vicious circle. Families affected by VL have to spend more directly cost (like treatment) and indirectly cost (like loss of household income) during the VL epidemic.

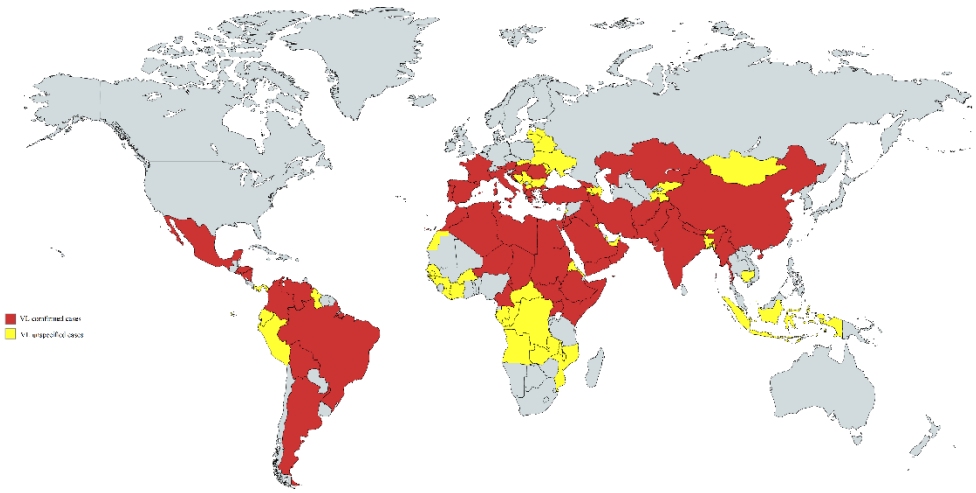


Figure 3. 2 Distributions of confirmed and borderline VL cases from 1960 to 2012

According to the WHO neglected tropical diseases (NTD) report in 2007, VL is identified as one of the NTD's [86]. The primary reason WHO classified VL as an NTD is because confirmed VL cases have decreased worldwide from approximately 60,000 to around 20,000

since around 2010 [11], as shown in Figure 3.3. However, thousands of VL cases may not be covered in the WHO VL estimation report [87] since some countries without public health information systems do not submit their infection report data to WHO (e.g., Chad, the Central African Republic, and the Democratic Republic of the Congo). Even though for the countries with completed public health information system, the reported epidemiological data could be incomplete, and official figures are likely to underestimate grossly [81]. Based on the estimation, there are 500,000 new cases of VL and more than 50,000 deaths from the VL each year [85]. In 2012, another research group from the WHO Leishmaniasis Control Team corrected the number of VL estimation cases into 0.2-0.4 million, and the number of VL deaths into 20,000-40,000 [88]. Therefore, VL continues to be a serious and present threat to public health worldwide.

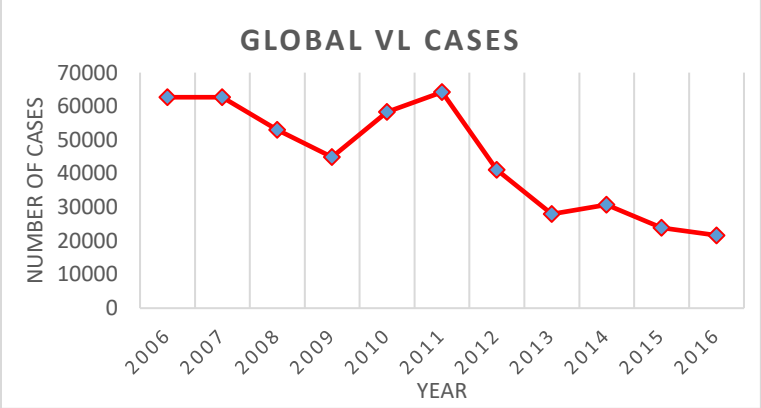


Figure 3. 3 Reported VL cases from 2006 to 2016 [11]

By observing the VL epidemiological situations for each countries in Figure 4 based on the report from WHO [11], the significant drop in reported VL cases can primarily be attributed to a significant decrease in reported VL cases in India and Bangladesh. Between 2006 and 2016 the number of reported VL cases in India decreased from 39,173 to 6,249 and reported cases in Bangladesh declined from 9,379 to 255. The major reason of this decreasing is caused by the widely utilization of insecticide - treated nets [89, 90]. However, VL prevalence did not change

significantly for other countries. For example, the number of annually reported VL cases in Brazil were approximately 2,700 throughout the 10-year reporting period.

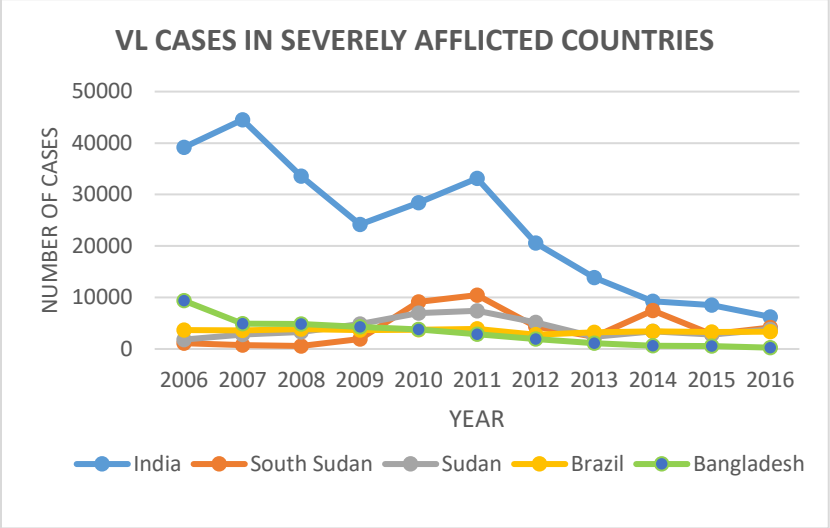


Figure 3. 4 Reported VL cases in severely afflicted countries from 2006 to 2016 [11]

Conversely, reported VL cases in several African countries have shown significant increase since 2006, as shown in Figure 3.5. For example, Somalia reported less than 100 cases in 2006 and 781 cases in 2016. Although the population of Somalia (14.32 million) is much less than the population of India (13.24 billion), the proportions of VL cases reported in these two countries were similar in 2016. If no immediate measures are taken in these selected African countries, large-scale VL outbreaks are imminent.

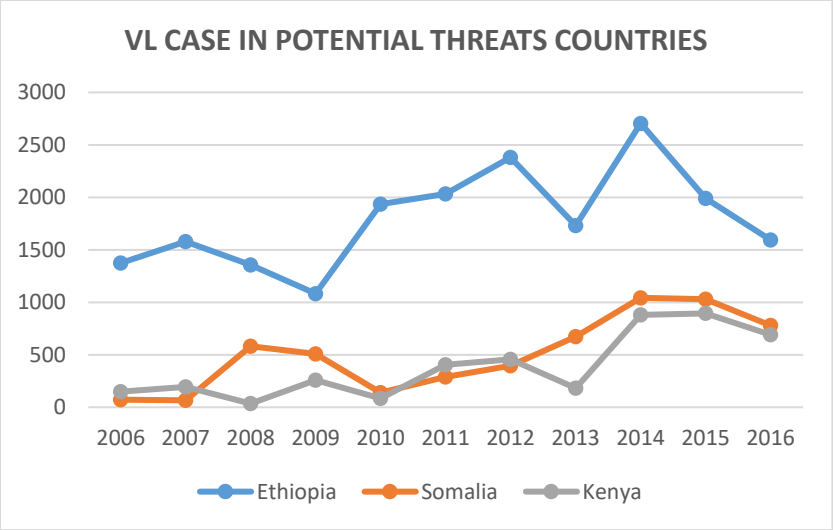


Figure 3. 5 Reported VL cases in vulnerable countries from 2006 to 2016 [11]

Therefore, VL is still a serious disease which threatens people life health especially in the developing countries. To strain the transmission of VL, WHO and health organizations in VL afflicted countries should apply effective prevention and control strategies. In the next section, this paper will review serval existing strategies against VL.

3.3 VL Mitigation and Prevention Methods

Since 1995, researchers have focused on ZVL when investigating the intervention and prevention of VL transmission because many ZVL control strategies are related to animals. Tesh categorized former ZVL control strategies into three main classes: early detection of human cases, destruction or treatment of infected dogs, and vector control [80]. Tesh's paper mentioned treating infected people does not affect the parasite transmission, and the drug resistant of *L.infantum* made the treatment even more difficult. For the dog control, many serologically positive dogs are hard-detectible. And the VL infected dogs need the expensive and continual treatments. Tesh's paper also proposed that preventing the disease in dogs population is the best control strategies, since it can interrupt the transmission cycle of ZVL. In 2002, Guerin *et al.*

asserted that principal interventions of VL via early diagnosis and treatments, dog controls, and vector controls are effective control strategies [81]. However, Guerin *et al.* also pointed out that vector controls such as residual-insecticide house spraying are cost-prohibitive and rarely used. Especially in India, *Phlebotomus argentipes* (the dominated species of sand fly) is becoming resistant to the insecticide. The authors also mentioned some special challenges in VL endemic areas: lack of financial support (India and Bangladesh), remote places (Brail) and civil war (Sudan). In 2006, Dantas-Torres *et al.* introduced the Brazilian Leishmaniosis Control Program (BLCP), which includes diagnosis and early treatment of human cases, immunological screening of seropositive dogs, and insecticide spraying [82]. Dantas-Torres *et al.* did indicated that the culling of seropositive dogs has limited effects, as proven by experiments and mathematical models. Authors summarized several key points of unsuccessful culling dogs: the limitation of the immunological screenings to detect anti-Leishmania antibodies, the opposition of dog owners to the culling of asymptomatic dogs and the lack of evidence that it is an effective intervention strategy. A paper by Quinnell *et al*. suggested that tropical insecticides-collars and pour-ons can be used to reduce VL incidences for dogs by more than 83% [83]. Moreover, the recent research introduced a seconde generation vaccine called Leish-Tec®, which is mixed by saponin and the recombinant protein A2 [91]. This new vaccination was tested in endemic areas of ZVL and demonstrated a protection in 92.9% of vaccinated dogs. However, like the insecticide spraying, the expensive cost limits the feasibility of the tropical insecticides-collars and canine vaccination. The authors offered an alternative method, insecticidal bath, which can protect dogs for at least 3.5 months against *Phlebotomus chinensis*.

In 2014, Werneck considered the effectiveness of control strategies on the basic reproduction number R_0 [92]. The author found that vector controls (e.g., controlling vector

density, the ratio of female vector, and the extrinsic incubation period of *L.infantum* in sand flies) and dog controls (e.g., culling infected dogs, dog vaccinations, and insecticide-releasing dog collars) can cause the disease reproduction rate R_0 to decrease to less than 1, meaning the number of infected individuals will eventually decrease to zero. However, the author did not find the relevant data to support these control strategies. He pointed out some potential implementation difficulties for these strategies, such as costs issues associated with the continual uses of tropical insecticides-collars. Werneck also voiced concern that the effectiveness of using insecticide-collars in the large community (like Brazil) may not work so well, since the insecticide-collars strategy has the relatively short-term effect and consequent need.

Due the high cost of indoor residual spraying, insecticide-treated nets (ITNs) were introduced as an alternative control strategy for ZVL [93]. Experimental trials in sub-Sahara Africa, Latin American, Thailand, Pakistan and Iran show that ITNs could reduce infections with malaria by 17%-62%. In 2015, Picado *et al.* summarized the results of the KALANET project to analyze ITN effectiveness. The KALANET project is a cluster-randomized controlled trial in India and Nepal [94]. The KALANET project was conducted in 26 high-incidence regions (in India and Nepal) with more than 20,000 inhabitants observed over 24 months. Results showed that use of long-lasting insecticidal nets resulted a 50% reduction of *L.donovani* infections. However, this study also shown that some of those nets were untreated, many were damaged, and most families did not have enough nets to protect all family members. The authors suggested to standardize the color and size of insecticide nets, they also want the government can replace the untreated and damaged nets by new treated nets in free for publics. Although the long-lasting insecticidal nets can prevent VL transmissions while people are sleeping, recent entomological

findings in India indicated that vectors are more exophilic than previously thought, meaning that people engaged in outdoor work have more opportunities to become infected [95].

In 2016, Özbel *et al.* analyzed the geographical distribution, ecological aspects, and species habits of VL vectors (sand flies) in Bangladesh [96]. In general, two genera of sand fly (*Phlebotomus* and *Sergentomyia*) total 13 species in Bangladesh, among which *Phlebotomus.argentipes* is the dominant vector species in VL endemic areas of Bangladesh (>94%). Researchers also found that the *Phlebotomus.argentipes* population peaks around monsoon season and reaches the lowest ebb during winter and summer in Bangladesh. They also determined that eight ecological parameters (soil temperature and moisture, rainfall, air temperature, relative humidity, soil pH, soil organic carbon, and wind speed) can influence the *Phlebotomus.argentipes* population. Future research must enact control strategies based on their ecological aspect. The potential application of this research can provide early warning of the incoming VL epidemic, and narrow the range of the insecticide spraying.

As the most effective control strategy for infectious diseases, the successful vaccination on VL is long-awaited for the VL afflicted countries. The experiment on VL vaccination was started in 1990s, researchers tried to utilize the proteins from *L.donovani* to development vaccines [97]. At the beginning of the 21-century, researches investigating the possibility of vaccination based on DNA and genetically attenuated parasites [98, 99]. More recently, Duthie *et al.* studied several different vaccine antigens for VL transmissions using recombinant proteins from *E.coli* [100]. Their research results have shown that several potential vaccine antigen candidates are verified in different platforms. The authors consolidated seven vaccine candidates as recombinant proteins in *E.coli*, and they validated the effectiveness of *E.coli* to *L.donovani* via experiments on mice. However, their research pertained to only nonhuman experiments.

Significant research and advancements are still needed to obtain effective vaccination for humans against VL parasites.

For now, even though many contributions have been done for the VL controls and prevention, an effective, feasible, and economical control strategy is still an ongoing effort. The current VL control strategies and corresponding deficiencies are summarized in Table 1. Since the most of VL severely-afflicted countries are developing countries, how to balance the effectiveness and costs involved in such VL control plan is delicate tradeoff. This paper will discuss more about this particular issue in Section 3.5.

Table 3. 1 Current VL control strategies and corresponding deficiencies

Strategy	Category	Deficiency	Reference
Early detection	Human control	Doesn't affect the parasite transmission	[80]
Culling dogs	Dog control	Opposition of dog owners & Hard to detect infected dogs	[80, 82]
Dog treatments	Dog control	Expensive	[83]
Canine vaccination	Dog control	Expensive & Drug resist	[83]
Spraying insecticide	Vector control	Expensive & Drug resist	[81, 93]
Immunological screening	Dog control	Expensive & Need high level technique support	[80, 83]
Insecticide collars	Dog control	Expensive & May not work in large community	[92]
Insecticide-treated nets	Vector control	Damaged and untreated nets have low effectiveness	[94]
Ecological control	Vector control	Need more time to be applied in the real world	[96]
Vaccination control	Human control	Not available at the current time	[100]

3.4 Mathematical Epidemiology Models for VL

3.4.1 System Dynamic Model

In 1996 Dye first introduced a 4-equation ODE susceptible-latent-infectious-removed (SLIR) model to describe the VL epidemic [101]. SLIR respectively represents the susceptible, latent, infectious, and resistant populations of VL, and the model considered the transitions

between these populations. However, the author did not consider the behaviors of dog and vector in his model. Courtenay *et al.* improved Dye's model by considering the dog population as a key parameter in the SLIR model [102]. Variations of this parameter directly influence the human infection rate. Ribas *et al.* built an ODE model to describe VL transmissions among humans, dogs, and vectors [103]. The model utilized 11 ODE equations, including the susceptible (dog, sand fly, and human), latent (dog, sand fly, and human), infectious (dog, sand fly, and human), and recover (dog and human) populations. Although their model was presented without data, simulation, and analysis, it was the first model to describe behaviors for all species involved in VL.

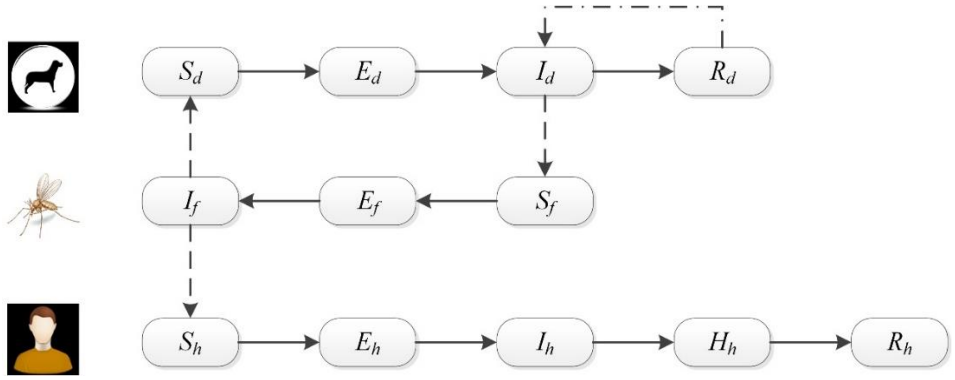


Figure 3. 6 System diagram of ZVL transmission model [50]. Where $\boldsymbol{d}, \boldsymbol{f}, \boldsymbol{h}$ represent the dog, sandflies and human species. Where $\boldsymbol{S}, \boldsymbol{E}, \boldsymbol{I}, \boldsymbol{R}, \boldsymbol{H}$ represent the susceptive, exposed, infected, recovered and hospitalized population for each species.

Since WHO's designation of VL as an NTD in 2015 [104], an abundance of research and studies have focused on developing mathematical models of VL. In 2016 Zhao *et al.* developed a 12-equation ODE model to comprehensively describe a VL epidemic (Figure 3.6) [50]. They improved the model by including a hospitalized population into the ODE system; this population has a higher probability of survival than infections without hospitalization due to the systematic treatment. The authors utilized a backward bifurcation method to analyze VL equilibrium

behavior and the basic reproduction rate R_0. Their research showed that VL equilibrium is highly related to a critical model parameter, R_c, as the epidemic threshold value for R_0. Similarly, Subramanian *et al.* proposed a compartment-based ODE model of VL transmission to explain disease transmissions in symptomatic VL, asymptomatic VL, and PKDL-infection classes [105]. Sensitivity analysis of model parameters found that the biting and birthing rates of sandflies and the recovery rate of symptomatic humans are dominating factors for VL epidemic control.

Biswas simplified the 12-equation ODE model from Zhao *et al.* to an 8-equation ODE model by dividing the non-human populations (dog and vector) into susceptible and infected population groups [106]. With a simplified ODE model, researchers reduced the complexity of system sensitivity analysis and reduced the numbers of assumed or estimated model parameters. This model also successfully reproduced the 2011 VL epidemic in South Sudan. Shimozako *et al.* transferred the ODE model in Ribas *et al.* by considering the dog population as the only source of infection since vectors could not transfer ZVL without dogs [107]. Therefore, their mathematical model contained eight variables corresponding to the susceptible, latent, infectious, and recovery populations for dogs and humans. Le Rutte *et al.* compared three ODE models and corresponding simulation results while considering indoor residual spraying [108]. The primary difference between the three models was how relationships between PKDL and the recovery population were modeled. Their research predicted that, using 60%–80% IRS coverage, VL epidemics could be eliminated within three years in Bihar, India. In addition, other researchers have made incremental contributions using various ODE VL epidemic models [109-112].

Although many ODE models describe VL epidemics and transmission, the development of a novel dynamic model is an active area of research in the investigation of complicated transmission behaviors of VL under various situations and the development of improved

mitigation and control strategies. Bi *et al.* introduced a two-dimensional PDE model based on an existing ODE model [113]. The model presented in their research considered human age structures and time as two dimensions since historical data have strongly suggested that VL infection rates are highly correlated to human age groups. For example, children and teenager (age from 0 to 20) are more likely to become infected compared to other age groups. This research used computer simulation and mathematical analysis to explain this phenomenon and recover the historical VL endemic data published by WHO.

3.4.2 Models Based on Real-World Data

VL attracted significant epidemiology research attention, abundant statistical data were collected and reported on the current VL pandemic worldwide by scholars and researchers. Many researchers realized the importance of data utilization in VL model development. Disease data is generally utilized three ways: use of reported data to build statistical models, use of historical data to predict future prevalence, and use of existing data to calibrate model parameters in mathematical epidemic models.

The primary objective of building a VL statistical model is to statistically identify key parameters in the VL transmission process and determine relationships between the number of parameters and the number of infected population. Werneck *et al.* used consolidated census tracts to analyze VL disease prevalence data from different regions of Brazil [114]. The authors developed a spherical covariance structure model based on census data from 1992 to 1996 in Teresina, Brazil. By exploring spatial correlation structures of the census data, they found a positive correlation between reported VL incidences and residence in areas of green vegetation, especially in favela. In 2007 Werneck *et al*. extended their previous work by analyzing and comparing the results from 21 statistical models [115]. According to human and canine VL data

in Teresina, Brazil, the study found significant correlations between residence in areas with green vegetation and infected dogs and between the human infection rate and urbanization index or socioeconomic status index.

Thompson *et al.* studied relationships between climate and VL epidemics by establishing the statistical regression model [116]. Their research found that rainfall is the most significant parameter statistically correlated to VL incidences. The influence of geographical features of areas of residence (e.g., cities, plains, or foothills) to VL transmission was also considered. The foothill population statistically revealed a higher risk of VL infection than other populations. De Araújo *et al.* considered the statistical model with data and found that spatial data is more reliable and accurate for VL epidemic study and analysis [117]. Therefore, they applied spatial statistical modeling and the Bayesian approach to model and estimate risks of VL using historical data from Belo Horizonte, Brail. Their research showed that the relative risk of VL is correlated to income, education, and the ratio of infected dogs to inhabitants in Belo Horizonte. However, as opposed to their previous research findings, residence in areas with green vegetation did not show significant correlations to the risk of VL.

Ecological niche modeling (ENM), stemming from the genetic algorithm [118], has been widely used to predict VL prevalence since 2006. Nieto *et al.* first used ENM to analyze VL data from northeastern Brazil [119]. Using the geographic information system (GIS), the ENM model can predict prevalence risks in three levels (high, moderate, and low). When validating with historical data from Bahia, Brazil, the predictive model demonstrated high accuracy (more than 90%) on high-level and moderate-level data. Similar approaches were used to predict VL prevalence in North America and Middle Eastern regions [120, 121]. Several other methods were utilized to predict the trend of VL epidemics with various factors. Elnaiem *et al.* summarized

data from eastern Sudan in GIS and then used that data to build predictive models of VL infections based on rainfall and corresponding distance to a river [122]. Oshaghi *et al.* built a predictive degree-day model for VL using the single triangulation method [123]. This model predicted temporal and spatial distribution of VL infection density and generations of sandflies. Karagiannis-Voules *et al.* employed Bayesian geostatistical models to fit the VL incidence data from Brazil, and they identified environmental and socioeconomic predictors using Bayesian variables [124]. Their research results predicted that regions with humid climates and dense vegetation distributions are more vulnerable to VL than other regions.

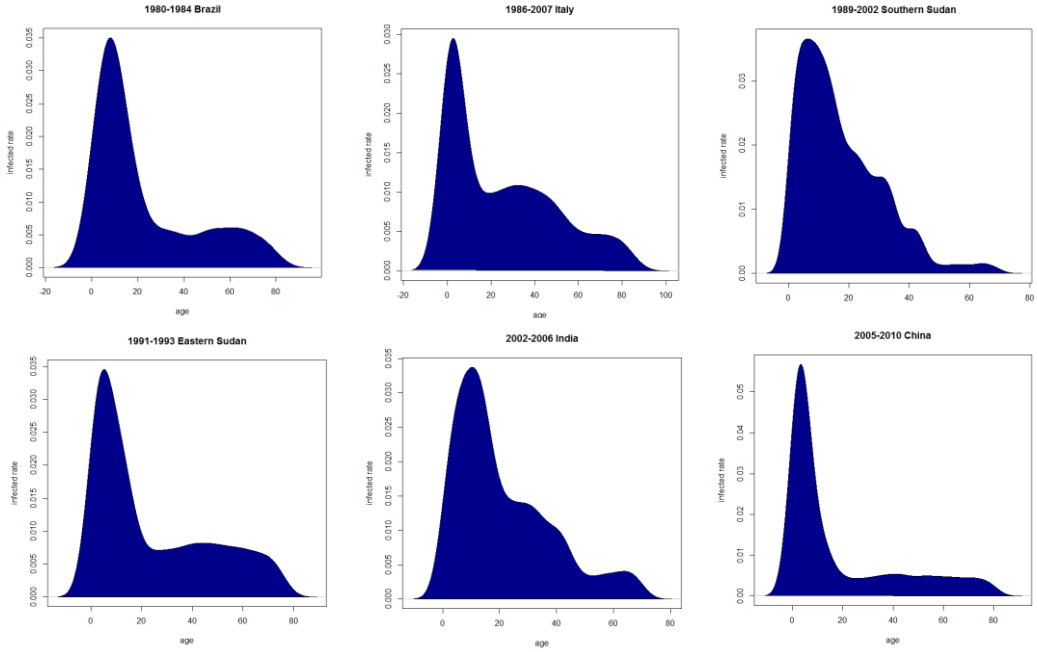


Figure 3. 7 Infection rate distribution based on human age in various countries

Parameter estimation is another essential application when validating VL mathematical models using real-world data since the use of assumed system parameters in the model may reduce model reliability. Our current research summarized age structures of VL infections in various regions. We revised the constant infection rate into age-dependent distribution of infection rates by studying historical VL prevalence data from existing literature from diverse

regions over time (Figure 3.7). Use of this age-dependent distribution function allowed their PDE epidemiology model to reflect VL human prevalence in both age and time periods. Biswas *et al.* established the posterior distribution of different parameters and initial parameters based on observation data [106]. Use of parameter distribution allowed the simulation experiments to reflect more than one result with unique possibilities. Even though most mathematical models still use assumptions or estimations with existing literature as their system parameters [50, 105, 107, 108], an increasing number of studies are utilizing real-world data to more accurately estimate system parameters and validate their models.

3.5 Optimal Control Strategies using VL Models

3.5.1 Parameter Control Strategy

The most generalized control strategy in a VL mathematical model is the parameter control strategy, which assumes that the key parameters in the model are adjustable. When the parameters are adjusted the model outputs become dependent variables; therefore, the parameter adjustment process can be considered a corresponding real-world control strategy. In 2002 Courtenay *et al.* introduced the numerical control strategy to the field of VL mathematical modeling [102]. They assumed that the parameter of dog density could be controlled by culling dogs in specific areas, and their research findings showed that dog culling could effectively reduce the proportion of infectious population. Ribas *et al.* improved the parameter control method by considering additional control parameters [103], including the dog treatment rate, the insecticide collar utilization rate, the dog culling rate (natural mortality rate of dogs), the dog vaccination rate, and the mortality rate of sandflies (vector control). Using the control parameter method, they used simulation experiments to compare the efficiencies of each control strategy, as shown in Figure 9, where vector control proved to be more effective than control strategies such

as culling and vaccine. Shimozako *et al.* incorporated control strategies such as dog treatment, dog vaccination, and use of insecticidal dog collars into an ODE mathematical model of VL [125]. However, they increased the dimension of the model by using control variables to replace the control parameters in their model. A cost-effect analysis and simulation experiments showed that use of insecticidal dog collars should be the most utilized control strategy.

3.5.2 Optimal Control Strategy

Lev Pontryagin established the optimal control strategy in the 1950s [126, 127]. This strategy provides optimality criterion by maximizing or minimizing a given objective function subject to constraints defined in the differential equation mathematical model. Zhao *et al.* introduced optimal control into their 12-equation ODE mathematical model of VL [50], including the susceptible, latent, and infectious populations for sandflies, humans, and dogs; recovered human and dog populations; and hospitalized human population. Their study included three control strategies in the model, including the control dog prevention level (vaccine protection or dog culling), the control insecticide usage level (insecticide sprayed around buildings), and the control personal protection level (long-lasting insecticide). Consideration of exposed human, infected human, and sandfly populations, as well as control cost as the control objective function caused the final control strategy to effectively lower the negative influence of VL by approximately 80% (Figure 9). Agusto *et al.* considered the use of fabrics and insecticidal animal collars as an additional control [112]. Their research was innovative because they used various combinations of controls instead of utilizing control strategies identical to previous research. Their simulated experimental results showed that use of an ODE mathematical model allowed their control strategies to effectively reduce VL and PKDL infected human populations. Biswas *et al.* revised a previous optimal control by suggesting a strategies mechanism and

incorporating various combinations of optimal strategies into the ODE mathematical model with a given objective function [128]. This objective function was a linear combination of VL and PKDL infections with the cost of control strategies. They calculated the corresponding infection averted ratios (IAR) and incremental cost-effectiveness ratios (ICER) for each control strategy, where IAR is the ratio of the number of infections averted to the number of recovered and ICER is the additional cost per additional health outcome. The researchers then selected the strategy with the highest IAR and lowest ICER as the optimal strategy. An optimal control strategy specifically targeting the human VL vaccination was also analyzed, but no evidence revealed the effectiveness of a human VL vaccination [106].

3.5.3 Control Strategy Selection Using Simulation

Simulation comparison is the most common method of VL mathematic control modeling in simulation. Figure 3.8 compares the human infected population with control (dog culling, sandfly control, and human protection) and without control [50]. The solid line in the figure shows epidemic performance (which includes the exposed human E_h in (a), the infected human I_h in (b), and the total population of sand-flies N_f in (c)) of the control strategy, which reduced disease prevalence to 80%. This simulation proved the effectiveness of the combined control strategies.

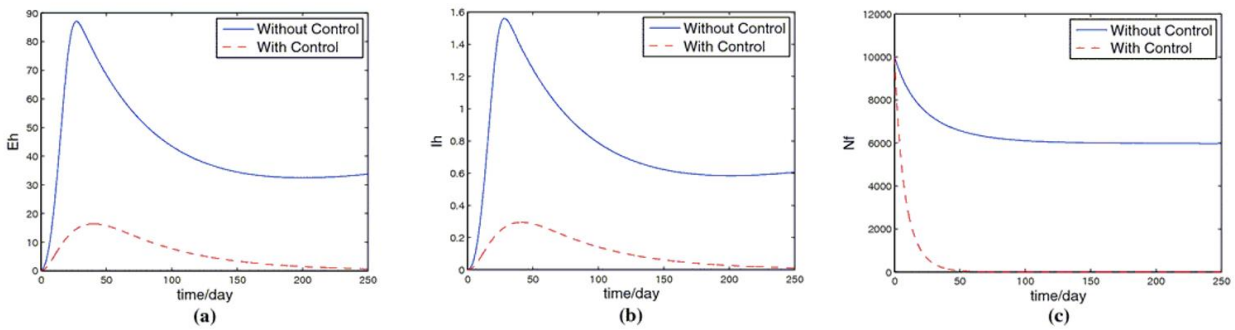


Figure 3. 8 Simulation of dog culling [50]

Efficacy comparison between control strategies is another type of simulation comparison. Ribas *et al.* compared the human prevalence influence using vector control, insecticidal collars, dog culling, dog vaccines, and dog treatment [103], as shown in Figure 3.9. When the parameter was changed to manipulate control levels, the human prevalence by vector control decreased most rapidly, proving that human prevalence is most sensitive to vector controls. Therefore, vector control is the most effective control strategy.

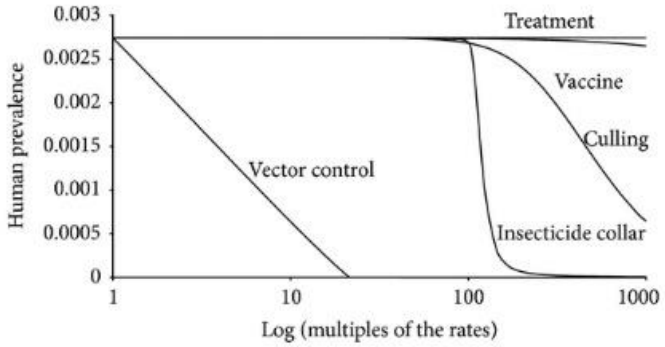


Figure 3. 9 Efficacy comparison of control strategies [103]

Spatial simulation is a simulation estimation method that provides spatial information throughout the model behaviors. Using GIS, spatial simulation can exhibit VL prevalence information from various regions, as shown in Figure 3.10. Karagiannis-Voules *et al.* utilized historical data (2001–2009) from Brazil to build a statistical model [124]. Their simulation used GIS to predict VL prevalence in Brazil in 2010, thereby directly reflecting high infection density in the eastern region of Brazil.

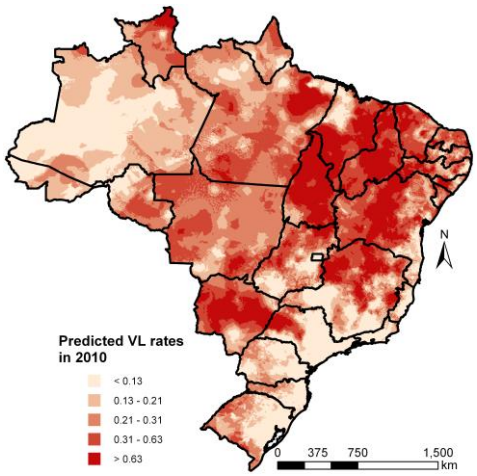


Figure 3. 10 Spatial simulation of predicted VL rates in 2010, Brazil [124]

3.5 Discussion and Conclusion

Although globally reported, VL-confirmed cases have decreased since 2011, VL prevalence has not improved significantly worldwide except in South Asia (i.e., India and Bangladesh). However, VL outbreaks have increased in Ethiopia, Somalia and Kenya since around 2008. Moreover, public health agencies in underdeveloped African countries such as Chad and the Central African Republic do not have resources and capabilities to collect and report VL incidences. However, because these countries are near regions severely afflicted with VL, such as Sudan and South Sudan, the number of global VL cases reported from WHO may be underestimated.

This paper reviewed current VL epidemiological research ranging from VL epidemic control strategies, VL mathematical models, and related optimal control strategies. The research demonstrated how to use numerical methods such as modeling and sensitivity analysis, as well as equilibrium/stability studies and simulation experiments to assist mitigation and prevention strategies for a worldwide VL pandemic. Governments and health organizations can utilize the modeling and simulation results to predict or estimate impacts of various control strategies.

Table 3. 2 Recent papers on mathematical modeling of VL

Paper	Published Year	Real Data Involved	Control Strategies	Transmission Models	Simulation
Ribas *et al.* [103]	2013	No	Yes	Yes	No
Zhao *et al.* [50]	2016	No	Yes	Yes	Yes
Subramanian *et al.* [105]	2015	Testing model	No	Yes	Yes
Biswas *et al.* [106]	2017	Parameter estimated	Yes	Yes	Yes
Shimozako *et al.* [107]	2017	Testing model	No	Yes	Yes
Le Rutte *et al.* [108]	2017	Testing model	No	Yes	Yes
Costa *et al.* [109]	2013	No	No	Yes	No
Sevá *et al.* [110]	2016	No	No	Yes	No
Zou *et al.* [111]	2017	No	No	Yes	Yes
Agusto *et al.* [112]	2017	No	Yes	Yes	Yes
de Araújo *et al.* [117]	2013	Yes	No	No	No
Karagiannis-Voules *et al.* [124]	2013	Yes	No	No	Yes
Miller *et al.* [129]	2014	Yes	Yes	Yes	No
Biswas *et al.* [128]	2017	Testing model	Yes	Yes	Yes
Shimozako *et al.* [125]	2017	Testing model	Yes	Yes	Yes
Stauch *et al.* [130]	2014	Testing model	No	Yes	Yes
Zamir *et al. [131]*	2017	No	No	Yes	Yes
Boukhalfa *et al.* [132]	2017	No	No	Yes	Yes
Gorahava *et al.* [133]	2015	No	Yes	Yes	No
Rock *et al.* [134]	2016	No	Yes	Yes	Yes

Despite significant research efforts using mathematical models for the VL epidemic, research gaps still exist and many areas of study remain unexplored. Table 3.2 summarizes 21 literature works related to mathematical modeling and VL disease control strategies since 2013. Most of the reviewed research used or proposed system dynamic mathematical models or statistical models; approximately half of these research considered real-world data and studied possible control strategies. Only one paper included real-world data, control strategies, mathematical models, and numerical simulation experiments.

Future, thorough VL epidemic research using mathematical or statistical models ought to consider the four following aspects:

1. Building more sophisticated mathematical models to explain underlying infectious disease transmission dynamics
2. Including real-world data to aid model validation and verifications
3. Exploring possible disease control/mitigation strategies to increase understanding of model maneuverability and robustness
4. Using numerical simulation experiments as a predictive tool to verify the feasibility of model and control strategies.

Moreover, future work in these four aspects of VL mathematical modeling must utilize modern analytical tools. The disadvantage of current modeling is the limited diversity of model types. A majority of existing VL mathematical models are ODE models, which are widely used but produce limited predicted results without details. Therefore, more statistical, machine learning, and PDE models are needed to build sophisticated, comprehensive mathematical models of VL. Statistical and machine learning models can more advantageously utilize real-world data to ensure model prediction accuracy, while use of a PDE model can enrich predicted results with age, gender, socioeconomic group, ethics, and spatial information. For the second aspect, the inclusion of real-world data, most test data currently used to validate and verify underlying mathematical models are estimated or assumed, consequently limiting the mathematical model to reflect only data from previous VL epidemic episodes. Future research efforts should utilize recent epidemic data with temporal and spatial data during the modeling phase, making the modeling process increasingly dynamic and reflecting real-time data while predicting possible trends of an ongoing epidemic. The current primary disadvantage of the third

aspect, exploring possible control strategies, is that the control strategies lack of applicability in the real world. In fact, the most effective control strategies suggested by the mathematical models may not be operable or they may be cost prohibitive to implement. Operable control strategies should be carefully quantized, such as specific consideration of the optimal level of canine culling in a particular time frame or the level of insecticide spraying in each area affected by VL. For the fourth aspect, current studies using numerical simulation experiments frequently provide insufficient information from simulation results. Most simulations of VL models can only predict the trend of VL infections. Future research should focus on spatial simulation and agent-based simulation as well as study the interactions between multiple regions or environments.

In conclusion, the use of mathematical models to study, analyze, and predict VL epidemics and to explore effective and implementable control strategies remains an active and study-worthy area of future research. However, research results from more comprehensive studies that use modern analytical tools will help public health organizations understand and prevent the VL disease.

Chapter 4 - A New Zoonotic Visceral Leishmaniasis Dynamic Transmission Model with Age-Structure

Chapter 4 is based on the manuscript "A New Zoonotic Visceral Leishmaniasis Dynamic Transmission Model with Age-Structure" Published in Chaos, Solitons & Fractals [30].

Abstract

Visceral leishmaniasis (VL) is a fatal, neglected tropical disease primarily caused by *Leishmania donovani* (*L. donovani*) and *Leishmania infantum* (*L. infantum*). According to VL infectious data reports from severely affected countries, children and teenagers (ages 0–20) have a significantly higher vulnerability to VL infection than other populations. This paper utilizes an infected function (by age) established from epidemic prevalence data to propose a new partial differential equation (PDE) model for infection transmission patterns for various age groups. This new PDE model can be used to study VL epidemics in time and age dimensions. Disease-free and endemic equilibriums are discussed in relation to theoretical stability of the PDE system. This paper also proposes system output adjustment using historical VL data from the World Health Organization. Statistical methods such as the moving average and the autoregressive methods are used to calibrate estimated prevalence trends, potentially minimizing differences between stochastic stimulation results and reported real-world data. Results from simulation experiments using the PDE model were used to predict the worldwide VL severity of the epidemic in the next four years (from 2017 to 2020).

Key word: zoonotic visceral leishmaniasis, dynamic transmission, age structure, partial differential equation, epidemic

4.1 Introduction

Visceral leishmaniasis (VL), or black fever (kala-azar), is a fatal, parasitic disease that infects a host's internal organs. According to the World Health Organization (WHO), VL is common in more than 80 countries, with approximately 50,000–90,000 new cases reported each year, primarily in Brazil, Ethiopia, India, Somalia, South Sudan, and Sudan. War and population displacement in Africa have promoted the explosion of VL. Consider the existence of the asymptomatic VL (much more than symptomatic VL), the total number of VL infections may be much higher than the number of reported cases [135]. Visceral leishmaniasis is designated as a neglected tropical disease (NTD) by WHO [75, 86]. Contrary to well-known NTDs such as; endemic treponematoses, dracunculiasis, and trypanosomiasis, VL is usually fatal within two years if left untreated. Even with treatment, however, VL infections can still become post-kala-azar dermal leishmaniasis (PKDL) [136]. VL is transmitted to humans through infected female phlebotomine sand flies, with more than 90 sand fly species identified as carriers of Leishmania parasites. Depending on the geographical transmission area, VL is primarily caused by two leishmanial species: *Leishmania donovani* (*L. donovani*) and *Leishmania infantum* (*L. infantum*) [85]. The transmission cycle designates VL as zoonotic visceral leishmaniasis (ZVL) or anthroponotic visceral leishmaniasis (AVL). ZVL occurs in the transmission area of *L. infantum*, and AVL occurs in the transmission area of *L. donovani*. ZVL, which is transmitted to humans and dogs via phlebotomine sand flies, primarily occurs in Mediterranean countries, central Asia, and developing areas in Africa and Latin America [137]. Dogs are the primary reservoirs of ZVL infection during transmission [83], hares and foxes can also be considered as the reservoirs (also known as the blood mule) of ZVL infection cycle [138, 139]. Moreover, ZVL has recently occurred in some nonendemic areas (Americas and Europe) due to dogs were transferred or

immigrated from endemic areas [140, 141]. The first confirmed case of ZVL was on the Greek island of Crete in 1907, and before the 1990s, a majority of cases were documented in the northeast region of Brazil [142].

Researchers have implemented mathematical modeling of VL epidemics to increase understanding and control the transmission of VL epidemics. In 1996 Dye introduced an ordinary differential equation (ODE) model to describe a VL epidemic, but this model only considered a susceptible, latent, infectious, recovered human population [101]. Courtenay *et al.* improved Dye's model by considering the dog population as an infectious source, determining that human infection rate is dramatically affected by the size of the dog population [102]. In 2016 Zhao *et al.* developed a 12-equation ODE model that considered VL transmission between susceptible (dog, sand fly, and human), latent (dog, sand fly, and human), infectious (dog, sand fly, and human), recovered (dog and human), and hospitalized (human) population [50]. Werneck *et al.* utilized statistical methods to analyze VL data from Brazil [114]. Their results revealed a positive association between reported VL incidences and residential areas with green vegetation. Also, the urbanization index is a key parameter of the human VL infection rate [115]. In addition, several research efforts studied numerical optimal control strategies based on established mathematical models, affirming that strategies such as the utilization of vector controls and insecticides are effective for reducing disease transmission [50, 108, 110, 112, 125, 128, 143].

Previous research, however, has not demonstrated the correlation between ZVL transmission and categorical age structures of infected human populations. In 2007 Chappuis *et al.* proposed a potential relationship between VL transmission and human age, stating that, although *L. donovani* and *L. infantum* can cause VL, *L. donovani* infects people of all ages, but

L. infantum primarily infects children [85]. Other papers have stated that ZVL is an epidemic primarily occurring in young children [80, 83, 144, 145] and, according to data reports, ZVL patients are often young children, with only a small percentage of diseases occurring in adults or senior citizens [146-151]. Moreover, the higher percentage of diseases occurring in children, even though they are less likely to be exposed to sandflies than adults [152]. The possible explanation is that VL has life-long immunity [85, 153]. Most adults have obtained this immunity since they recovered from the asymptomatic VL infections.

Although these papers highlighted the relationship of ZVL to human age, they did not demonstrate this assertion or investigate how human age directly affects ZVL transmission. Therefore, this paper applies a partial differential equation (PDE) to account for human age structure in a new ZVL transmission model to determine the potential relationship between ZVL transmission and human age. This paper also examines the basic reproduction number and model stability, with a specific focus on increasing accuracy of the new model for predicting future ZVL transmission using methods such as the moving average (MA) and autoregressive (AR) methods.

The main contribution of this paper is a proposed PDE model of ZVL with age structure to reveal dynamic transmission using WHO reported data for model revision. Section 2 introduces the PDE model for ZVL transmission with age structure. This section also simulates the model by considering Sudan as an example. Section 3 determines the equilibrium points of the system and discusses the stability of those points. Section 4 revises the model by considering the spatial features by using WHO cases report data from 22 VL infected countries and predict the ZVL transmission patterns in these countries from 2017-2020. Section 5 includes conclusions.

4.2 ZVL Model with Age Structure

4.2.1 Basic ZVL transmission dynamic model

This section will review the model from our previous work [50]. Reservoir hosts, dogs, can be divided into four states: susceptible (S_d), exposed (infected but not infectious) (E_d), infectious (I_d), and recovered (R_d). New dogs become susceptible with an assumed constant birthrate, λ_d. Susceptible dogs are exposed when they encounter infectious sand flies with an average biting rate, b_{fd}, and an infection probability, β_{fd}. The overall instantaneous rate of infection for susceptible dogs can be calculated as $\frac{b_{fd}\beta_{fd}I_f}{N_d}$, where N_d denotes the total dog population. Exposed dogs become infectious dogs (I_d) after the incubation period, with an instantaneous transfer rate, τ_d, and infectious dogs recover with a recovery rate r_d. Since the study shows that the dog population had a high rate of VL contraction which leads the dogs to develop a T- cell compartment [154].This model assumed the recovered dogs (R_d) would not become susceptible again. Infectious dogs are removed from the system with the death rate d_d. Exposed dogs, although they are seropositive, have no additional death rate. For each state, dogs could exit the system with a natural death rate μ_d. In this model, we ignore the migration behavior of dogs.

Sandflies can be divided into three states: susceptible (S_f), exposed (E_f), and infectious (I_f). In this paper, the seasonal patterns of the sand fly population are not researched. New sand flies become susceptible with a constant birthrate λ_f. Susceptible sand flies become exposed when they bite infectious dogs with a biting rate, b_{fd}, and an infection probability of β_{df}. This paper ignores the contagion from an infected human to susceptive sand files because the infection rate is low [155]. The overall instantaneous rate of infection for susceptible sand flies is

$\frac{b_{fd}\beta_{df}I_d}{N_f}$, where N_f is the total sand fly population. Exposed sand flies become infectious after an incubation period, with an instantaneous transfer rate, τ_f. For each state, they could exit the system with a death rate μ_f. In this model, we ignore the migration behavior of sand flies.

The human population can be divided into five states: susceptible (S_h), exposed (E_h), infected (I_h), hospitalized (H_h), and recovered (R_h). In this paper, lifelong immunity from VL is assumed. New individuals become susceptible with a constant birthrate λ_h. Susceptible individuals become exposed when bitten by infectious sand flies with a biting rate of b_{hf} moreover, a probability of infections with each bite β_{hf}. The overall instantaneous infection rate for susceptible human can be expressed as $\frac{b_{fh}\beta_{fh}I_f}{N_h}$, where N_h is the total human population in the system. Following the incubation period, exposed individuals turn into infected individuals with a transfer rate τ_h, and infected individuals become hospitalized with a rate of δ_h. Due to disease, infected individuals can leave the system with the death rate d_I before they become hospitalized, or infected individuals can become recovered individuals with the natural recovery rate r_I without being hospitalized. Hospitalized individuals can exit the system with the death rate d_H due to disease, or hospitalized individuals can recover with the treatment recovery rate r_I during hospitalization. For each state, individuals could leave the system with the natural death rate μ_h. In this model, we ignore the migration behavior of humans and the asymptomatic infected human population (due to the lack of data).

The entire population system for dogs, sand flies, and humans can be modeled using an ODE system as follows:

$$\frac{dS_d}{dt} = \lambda_d - \frac{b_{fd}\beta_{fd}I_f S_d}{N_d} - \mu_d S_d \tag{4.1}$$

$$\frac{dE_d}{dt} = \frac{b_{fd}\beta_{fd}I_f S_d}{N_d} - \tau_d E_d - \mu_d E_d \tag{4.2}$$

$$\frac{dI_d}{dt} = \tau_d E_d - r_d I_d - d_d I_d - \mu_d I_d \quad (4.3)$$

$$\frac{dR_d}{dt} = r_d I_d - \mu_d R_d \quad (4.4)$$

$$\frac{dS_f}{dt} = \lambda_f - \frac{b_{fd}\beta_{df} I_d S_f}{N_f} - \mu_f S_f \quad (4.5)$$

$$\frac{dE_f}{dt} = \frac{b_{fd}\beta_{df} I_d S_f}{N_f} - \tau_f E_f - \mu_f E_f \quad (4.6)$$

$$\frac{dI_f}{dt} = \tau_f E_f - \mu_f I_f \quad (4.7)$$

$$\frac{dS_h}{dt} = \lambda_h - \frac{b_{fh}\beta_{fh} I_f S_h}{N_h} - \mu_h S_h \quad (4.8)$$

$$\frac{dE_h}{dt} = \frac{b_{fh}\beta_{fh} I_f S_h}{N_h} - \tau_h E_h - \mu_h E_h \quad (4.9)$$

$$\frac{dI_h}{dt} = \tau_h E_h - \delta_h I_h - d_I I_h - r_I I_h - \mu_h I_h \quad (4.10)$$

$$\frac{dH_h}{dt} = \delta_h I_h - d_h H_h - r_h H_h - \mu_h H_h \quad (4.11)$$

$$\frac{dR_h}{dt} = r_h H_h + r_I I_h - \mu_h R_h \quad (4.12)$$

Definitions of system parameters are summarized in Table 4.1.

Table 4. 1 Table of parameters

Parameters	Description	Values	Source
λ_d	The new birth of susceptible dogs per day	0.02% of the total population of dogs	Reckoned from [156]
λ_f	The new birth of susceptible sand flies per day	0.6% of the total population of sandflies	[157]
λ_h	The new birth of susceptible humans per day	0.004% of the total population of human	Reckoned from [158]
b_{fd}	Average biting rate to dogs	0.1	[156]
b_{fh}	Average biting rate to humans	0.1	[156]
β_{fd}	Transmission probability from infectious sand fly to a susceptible dog	0.5	[156]
β_{df}	Transmission probability from infectious dog to a susceptible sand fly	0.7	[156]
β_{fh}	Transmission probability frominfectious sand fly to a susceptible human	0.5	[156]

μ_d	Natural death rate of dogs	0.0002 per day	[156]
μ_f	Natural death rate of sand flies	0.006 per day	[157]
μ_h	Natural death rate of individuals	0.000034 per day	Reckoned from [158]
$1/\tau_d$	Incubation period in dogs	10 days	[156]
$1/\tau_f$	Incubation period in sand flies	7 days	[101]
$1/\tau_h$	Incubation period in humans	90 days	[157]
r_d	Natural recovery rate of dogs	0.01 per day	[156]
r_I	Natural recovery rate of humans before hospitalization	0.00181 per day	Reckoned from [157]
r_h	Treatment recovery rate of humans after hospitalization	0.05 per day	Reckoned from [144]
d_d	Death rate of dogs due to disease	0.01 per day	[156]
d_I	Death rate of humans due to disease before hospitalization	0.006 per day	Assumed based on [159]
d_h	Death rate of humans due to disease after hospitalization	0.003 per day	Reckoned from [159, 160]
δ_h	Transform rate of humans from infected to hospitalized	0.8 per day	Assumed based on [50]

4.2.2 ZVL age data

Previous research and data reports have shown a difference in the distribution of human ZVL diseases for each age group. Research has shown that the children group (ages 0–20) tends to have a higher number of ZVL disease cases for various years and regions. Researchers have ascertained that ZVL transmission is correlated to age groups of individuals [18]. In 2006 Alborzi *et al.* reported that a majority of infected with VL caused by *L. infantum* were infants (0-2 years) in Iran [161], and in 2009 Quinnell *et al.* found that ZVL affects primarily young children [83]. Chappuis *et al.* have found that children and immunosuppressed individuals are most vulnerable to *L. infantum* infections [85].

In this paper, we started to collect and summarize ZVL disease data from existing literature. Our research goal is to thoroughly investigate the relationships between the age structure and ZVL infections, only data that described the number of infected cases according to the age group were considered. Toward this goal, human VL disease data are summarized respect to different age groups from reports [146-151] from Brazil, Italy, South Sudan, Sudan,

Bihar (India), and China, as shown in Table 2.2. VL disease data was used because no ZVL disease data exists. Because VL in Brazil, western Europe, central Asia, and China is zoonotic and both ZVL and AVL have been reported in East Africa, VL disease data was a reasonable option for determining rules of age distribution in ZVL infections [162, 163].

Table 4.2 shows the numbers of infection cases by age group for each region. For example, the number of infected humans in a group comprised of children younger than 16 years old accounted for more than 50% of total infection cases in Brazil from 1980 to 1984. In Italy from 1986 to 2007, a total of 515 infected humans were older than age 16, while 514 infected cases were children. In South Sudan from 1989 to 2002, approximately 57% (1817/3200) of infection cases occurred in children group (ages 0–20). From 1991 to 1993, Sudan reported 1283 (2288 in total) reported cases in a group comprised of children under 15 years old. In Bihar, India, from 2002 to 2006, the number of cases in children group (ages 0–20) comprised approximately 50% of total infection cases, and from 2005 to 2010, more than 1684 infectious cases were recorded in China.

Table 4. 2 Reported VL cases by age group for six regions

Years	Regions	*Leishmania* species and sources	Types of sand flies	Number of cases in age groups						
1980-1984	Brazil	*L. chagasi* [146]	*Lutzomyia longipalpis*	**Age**	**0-3**	**4-7**	**8-11**	**12-15**	**16-19**	**≥20**
				Cases	31	81	86	57	42	145
1986-2007	Italy	*L. donovani & L. infantum* [147]	*P. perniciosus P. perfiliewi P. neglectus*	**Age**	**0-2**	**3-6**	**7-16**	**17-50**	**51-70**	**>70**
				Cases	329	95	90	375	105	35
1989-2002	South Sudan	*L. donovani* [148]	*Phlebotomus orientalis*	**Age**	**0-9**	**10-15**	**16-24**	**25-34**	**35-44**	**≥45**
				Cases	1200	617	567	487	219	110
2002-2006	Bihar, India	*L. major & L.donovani* [150, 164]	*Phlebotomus*	**Age**	**0-9**	**10-15**	**16-24**	**25-34**	**35-44**	**≥45**
				Cases	467	346	217	214	165	144
2005-2010	China	*L. donovani & L. infantum* [151]	*P. longiductus P. chinensis P. wui*	**Age**	**0-6**		**7-14**		**≥15**	
				Cases	1368		316		766	
1991-1993	Sudan	*L. major & L.donovani & L. infantum* [149]	*Phlebotomus orientalis*	**Age**	**0-5**		**6-15**		**≥16**	
				Cases	714		569		1005	

Although many studies have investigated ZVL epidemics, most have not considered the age structure of ZVL infections. Also, most existing mathematical models are based on the ODE system, meaning the models only reflect ZVL transmission tracks in one dimension (time dimension). Therefore, the primary objective of this research was to study the spread of ZVL while considering age structures in the population model. A mathematical model based on a PDE was built to study ZVL epidemics over time and with consideration of human age groups. An infection density function based on age structure was developed using kernel density estimation. The infection density function was then incorporated into the mathematical model to accommodate various infection rates throughout age groups in the PDE system. The next section details the development of the infection density function.

4.2.3 Data process using kernel density estimation (KDE)

Section 2.2 summarized ZVL cases from six regions over time to demonstrate strong correlations between disease prevalence and age groups. However, the summarized data in Table 2.2 were not based on a consistent standard because the reported data were collected at different time periods from each region and the quantization of age groups differed by region, preventing consistent studies of the infection rates in age groups using this discrete dataset. Therefore, the main objective in this section is to develop a uniform infection density function from several discrete datasets using KDE.

The main objective of KDE is to approximate real, underlying probability distribution using a smooth kernel function to best-match the sampled data [165]. KDE considers all observation data and width as parameters of a chosen kernel function. (Width is the smoothing parameter that controls the size of the neighborhood around the target point.) Then the estimation function for kernel density can be obtained using linear superposition, and the final probability

density function can be obtained after normalization. Therefore, for observation dataset $\{x_1, x_2, \dots, x_n)$, the target point x_0, the estimated density for any target point x_0 can be defined as

$$\widehat{p(x_0)} = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x_0 - x_i}{h}), \tag{4.13}$$

where K is the kernel function with $\int K(x)dx = 1$, $h > 0$ is the bandwidth, which determines the smoothness of K.

$$K_{Gaussian} = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}x^2) \tag{4.14}$$

The popular Gaussian kernel estimator is a smoothing case with a set of essential smoothing properties [166] (as shown in equation (4.14)). Use of the Gaussian kernel estimator to process data from the six regions presented in Table 4.2 resulted in six infected population density distributions throughout the various age groups, as shown in Figure 4.1.

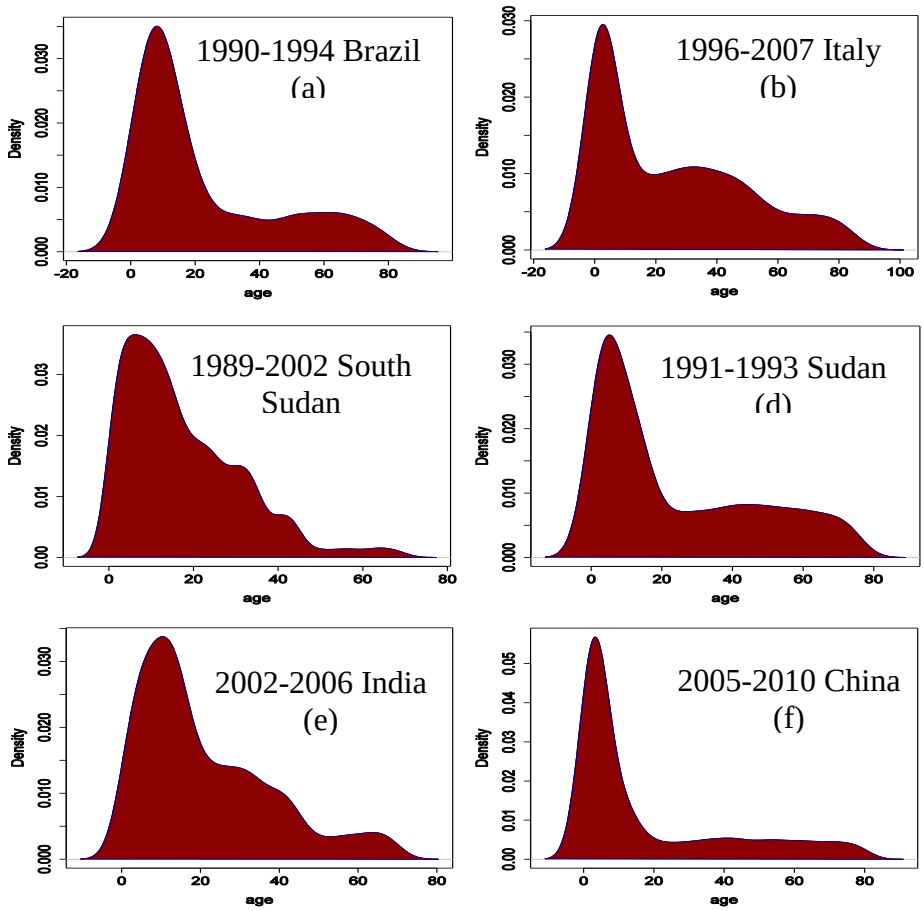


Figure 4. 1 Infected population density distributions by age in selected countries

As shown in Figure 4.1, the age group of 0 to 20 years old demonstrated a much higher density than other age ranges for all six regions, with density peaking near age 10. Throughout all six regions, the figure consistently shows that the numbers of infected individuals were significantly different for different age groups. The younger population under 20-year-old shown to be much more vulnerable to VL infection than the older age groups, thereby affirming conclusions from previous literature [83, 85].

Also, Figure 4.1 also contains ages less than 0 because density values of negative age were artificially produced by the Gaussian kernel estimator during the process of smoothing out the discrete data points. In Table 4.2, the density information started at 0 years of age, but the

KDE method smooths the left tail of age 0 in the horizontal axis. To eliminate the density information for negative age values in Figure 4.1, the density for each age was reprocessed using the following equation:

$$p'(a_i) = p(a_i) / \int_0^{100} p(a_j) da_j, \tag{4.15}$$

where $p(a_i)$ is the infected population density of age a_i estimated using the Gaussian kernel estimator.

In this paper, the calculated infected population concerning different age groups will be used to calibrate the corresponding infectious rates. Since ZVL infected population data are not available for all nations around global until 2018, this research utilized transmission rates of VL according to the following scheme during the simulation. When the investigated region was one of the six regions identified for data collection, the region's historical transmission rate will be used as the current transmission rate. If the region under study was not one of the six identified regions, transmission rates from one of the six regions closest to that region were used since disease transmission characteristics were similar.

4.2.4 ZVL transmission dynamic model with age structure

Section 4.2.3 established the relationship between infected rates and human ages using the KDE method. This section proposes a mathematical model that incorporates age structure in human populations over time to identify the dynamics of the ZVL epidemic for different age groups. Modeling assumptions are summarized in Table 4.3.

Table 4. 3 Basic Assumptions for ZVL Age-Structure Modeling

Assumption 4.1:	Average biting rates per infectious sand fly to children, adults, and seniors are equal throughout the epidemic time; the infected rate for humans is only related to age.
Assumption 4.2:	The ages of dogs and sand flies remain the same throughout the disease epidemic.
Assumption 4.3:	Except for the infected rate, all humans have the same set of parameters (e.g., natural death rate, hospitalization rate, and recovery rate).
Assumption 4.4:	Initial age distributions of susceptible, exposed, infected, hospitalized and recovered humans are known.

Assumption 4.5: The birth rate for susceptible humans is constant.

The system dynamic transmission model is illustrated in Figure 4.2.

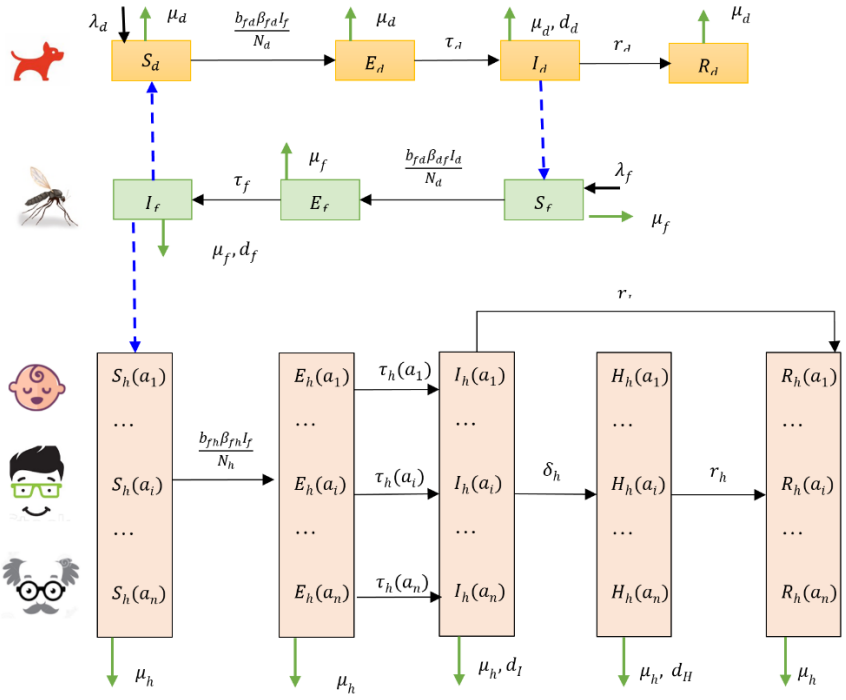


Figure 4. 2 System dynamic diagram of ZVL transmission model

The dashed lines in Figure 4.2 represent ZVL transmission pathways between dogs, sand flies, and humans. ZVL can be transmitted from infectious dogs to susceptible sand flies during blood meals, and then it can be transmitted from infectious sand flies to susceptible dogs and susceptible humans. However, if a susceptible sand fly bites an infected human, this sand fly cannot infect other susceptible humans or susceptible dogs. The solid arrow represents the population transfer direction within the species. Only age structure within the human population was considered. The terms $S_h(a,t)$, $E_h(a,t)$, $I_h(a,t)$, $H_h(a,t)$, $R_h(a,t)$ and $N_h(a,t)$ represent the number of susceptible, exposed, infected, hospitalized, recovered, and total human population in age group a at time t. $\tau_h(a,t)$ is the function of infected rate for humans in age group a at time t. Susceptible humans can become exposed at a rate $\frac{b_{fh}\beta_{fh}I_f}{N_h}$ for all age groups.

Exposed humans of age a become infected at a rate of $\tau_h(a)$. Infected humans can become hospitalized humans at a rate δ_h and recover at a rate r_I for all ages. Hospitalized humans become recovered at a rate r_H for all ages.

Based on the assumptions in Table 4.3 and the described ZVL disease transmission, the population model for dogs and sand flies remained unchanged in equations (4.1)–(4.7). ZVL cannot transmit vertically, so the human population model in equations (4.8)–(4.12) can be modified to incorporate age structure without vertical transmission (mother-to-fetus transmission) as following:

$$\frac{dS_h(a,t)}{dt} + \alpha\frac{dS_h(a,t)}{da} = -\frac{b_{fh}\beta_{fh}I_f S_h(a,t)}{N_h(a,t)} - \mu_h S_h(a,t) \quad (4.16)$$

$$\frac{dE_h(a,t)}{dt} + \alpha\frac{dE_h(a,t)}{da} = \frac{b_{fh}\beta_{fh}I_f S_h(a,t)}{N_h(a,t)} - \tau_h(a,t)E_h(a,t) - \mu_h E_h(a,t) \quad (4.17)$$

$$\frac{dI_h(a,t)}{dt} + \alpha\frac{dI_h(a,t)}{da} = \tau_h(a,t)E_h(a,t) - \delta_h I_h(a,t) - r_I I_h(a,t) - \mu_h I_h(a,t) \quad (4.18)$$

$$\frac{dH_h(a,t)}{dt} + \alpha\frac{dH_h(a,t)}{da} = \delta_h I_h(a,t) - d_H H_h(a,t) - r_H H_h(a,t) - \mu_h H_h(a,t) \quad (4.19)$$

$$\frac{dR_h(a,t)}{dt} + \alpha\frac{dR_h(a,t)}{da} = r_I I_h(a,t) + r_H H_h(a,t) - \mu_h R_h(a,t) \quad (4.20)$$

The left sides of equations (4.16)–(4.20) denote population change as the change of time and age for each state. Total susceptible humans at time t between age a_1 and age a_2 is defined as $\int_{a_1}^{a_2} S_h(a,t)\,da$, and total susceptible humans at time t for all ages is $\int_0^A S_h(a,t)\,da$, where A is the maximum age. $E_h(a,t)$, $I_h(a,t)$, $H_h(a,t)$, and $R_h(a,t)$ can be computed similarly. The right-hand sides of equations (4.16)–(4.20) describe population change according to disease transmission for each state. Because the age of a human is a variable related to time, a scalar coefficient α is to maintain age and time in an equal scale unit. For example, if the unit of age is year and unit of time is a week, then α is computed as $\alpha = \frac{1}{52}$ (Year/Week).

Because ZVL cannot spread from mother to fetus, newborns were regarded as a susceptible population [167]. Thus, when $a = 0$, the birth rate of exposed, infected, hospitalized, and recovered class were all equal to 0. Based on assumption 4.5, boundary conditions are

$$S_h(0,t) = \sigma_h \tag{4.21}$$

$$E_h(0,t) = 0 \tag{4.22}$$

$$I_h(0,t) = 0 \tag{4.23}$$

$$H_h(0,t) = 0 \tag{4.24}$$

$$R_h(0,t) = 0 \tag{4.25}$$

Initial conditions for the model with age structure can be denoted as

$$S_h(a,0) = S_h(a) \tag{4.26}$$

$$E_h(a,0) = E_h(a) \tag{4.27}$$

$$I_h(a,0) = I_h(a) \tag{4.28}$$

$$H_h(a,0) = H_h(a) \tag{4.29}$$

$$R_h(a,0) = R_h(a) \tag{4.30}$$

If $t = 0$ in equation (4.21), the boundary condition can be computed using equation (4.31):

$$\sigma_h = S_h(0) \tag{4.31}$$

4.2.5 Numerical Simulation

This section demonstrates disease transmission simulation for Sudan, a country with the most serious and typical ZVL outbreaks of the six studied regions. The moisture environment in Sudan causes thousands of ZVL cases each year, accounting for approximately two-thirds of the total cases of VL in Sudan [74]. In addition, ZVL caused a fatal influence in Sudan [168]. Therefore, we utilized Sudan as the research example to simulate the ZVL transmission.

Simulation parameter values are shown in Table 2.1. The birthrate was assumed to be equal to the natural death rate for dog and sand fly populations, so estimated birthrates of dog and sand fly populations were based on natural death rates from references [50, 103, 156, 169,

170]. For the sack of solution accuracies, all the computational experiments in this research utilized the Runge-Kutta method to numerical solve the model in equations (4.1)-(4.12), (4.16)-(4.20) [171].

Initial values of populations for dog, sand fly, and human were set using the following methods. Because Africa has approximately 148 million dogs [172] and the areas of Africa and Sudan are roughly 30.37 million km^2 and 1.886 million km^2, respectively, the total number of dogs in Sudan was estimated to be approximately 148*1.886/30.37 = 9.19 million. A sand fly survey from Sudan showed that 596 sand flies were collected in a study area with a total land area of approximately 7.3328 km^2; therefore, the total number of sand flies in Sudan was calculated to be approximately 596*1,886,000/7.3328 = 153 million [173]. The human population of Sudan is 39.58 million, so initial dog populations were $S_d(0) = 9{,}181{,}000$, $E_d(0) = 100{,}000$, $I_d(0) = 5{,}000$, $R_d(0) = 0$, and initial sand fly populations were $S_f(0) = 152{,}400{,}000$, $E_f(0) = 500{,}000$, $I_f(0) = 50{,}000$. The age groups in susceptible human population were assumed to be divided equally. The unit of time in simulation experiments was set to be one week and the unit of age was one year, so $\alpha = \frac{1}{52}$. In this simulation, the time epoch was $t_f = 52$ weeks and the life span of a human is $A = 80$ years old.

All simulation runs were carried out with MATLAB programs on a Windows-based. System parameters such as infection rates or specific model parameters could be preset manually prior to each simulation run.

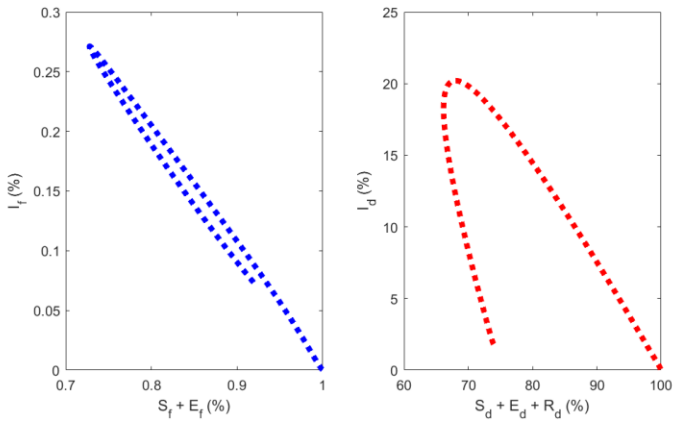


Figure 4. 3 Infectious phases portraits of dog and sand fly population

Figure 4.3 shows the phases portraits of dog and sand fly population calculated by the simulation. The horizontal axis represents the uninfected population divided by the initial population; the vertical axis represents the infected population divided by the initial population. By observing the peak of the epidemic outbreak, about 28% of the total sandfly population and 21% of the total dog population were infected by ZVL.

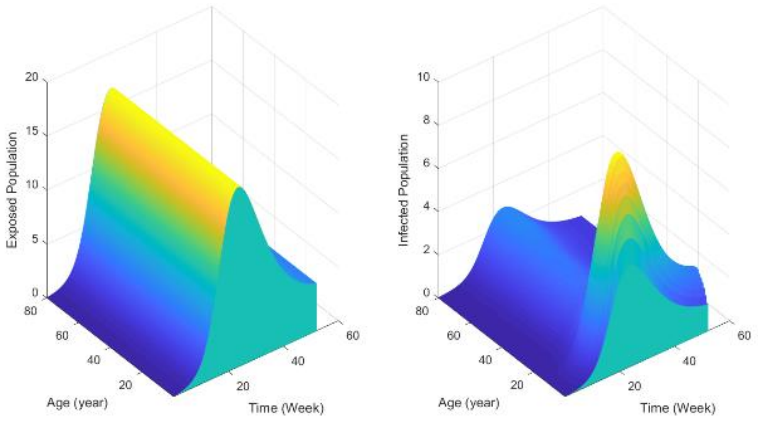


Figure 4. 4 Exposed and infected human population for each age over time

Figure 4.4 illustrates the ZVL transmission through the human population for each age group at each time epoch. The first graph reveals almost no difference in the exposed population between age groups, but the exposed population is shown to be sensitive to time throughout the

epidemic. The peak of the exposed population was around the 20th week. The second graph shows that the children group (ages 0 to 10 years), had a significantly higher infected population than other age groups at other time epochs. The peak of the infected population occurred around the 20th week, which means the peak the ZVL epidemic outbreaks happened at around the 20th week.

4.3 Stability and Sensitivity Analysis

4.3.1 Disease-free equilibrium analysis

The equilibrium point of a disease-free situation can be referred to as disease-free equilibrium (DFE), meaning the PDE system established in Section 2 can be stable under some specific conditions. To find the DFE, first, the PDE system must be disease free, that is at least one set of positive real number solutions exists for the equation set (4.32). Also, $E_d = I_d = E_f = I_f = E_h = I_h = 0$. The homogeneous system of equations (4.32) was obtained by letting the right-hand sides of equations (4.1)–(4.7) and equations (4.16)–(4.20) equal zero.

$$\begin{aligned}
&\lambda_d - \frac{b_{fd}\beta_{fd}I_f S_d}{N_d} - \mu_d S_d = 0\\
&\frac{b_{fd}\beta_{fd}I_f S_d}{N_d} - \tau_d E_d - \mu_d E_d = 0\\
&\tau_d E_d - r_d I_d - d_d I_d - \mu_d I_d = 0\\
&r_d I_d - \mu_d R_d = 0\\
&\lambda_f - \frac{b_{fd}\beta_{df}I_d S_f}{N_d} - \mu_f S_f = 0\\
&\frac{b_{fd}\beta_{df}I_d S_f}{N_d} - \tau_f E_f - \mu_f E_f = 0\\
&\tau_f E_f - d_f I_f - \mu_f I_f = 0\\
&\lambda_h - \frac{b_{fh}\beta_{fh}I_f S_h}{N_h} - \mu_h S_h = 0\\
&\frac{b_{fh}\beta_{fh}I_f S_h}{N_h} - \tau_h E_h - \mu_h E_h = 0\\
&\tau_h E_h - \delta_h I_h - d_I I_h - r_I I_h - \mu_h I_h = 0\\
&\delta_h I_h - d_H H_h - r_H H_h - \mu_h H_h = 0\\
&r_H H_h + r_I I_h - \mu_h R_h = 0
\end{aligned} \tag{4.32}$$

where

$$S_h = \int_0^A S_h(a,t)\,da,\ E_h = \int_0^A E_h(a,t)\,da,\ I_h = \int_0^A I_h(a,t)\,da,\ H_h = \int_0^A H_h(a,t)\,da,\ R_h = \int_0^A R_h(a,t)\,da.$$

Two equilibriums were obtained after solving the homogeneous equation (3.1); consequently, the first equilibrium was a theoretical DFE rather than a realistic one:

$$E_1^0 = \left(S_d, E_d, I_d, R_d, S_f, E_f, I_f, S_h, E_h, I_h, H_h, R_h\right) = (N_d, 0, 0, 0, 0, 0, 0, N_h, 0, 0, 0, 0) \tag{4.33}$$

This way, the population of sand fly is zero. Thus, ZVL cannot be spread over time; it is a theoretical DFE.

The second equilibrium is an approachable DFE, which can be assigned as:

$$E_2^0 = \left(S_d, E_d, I_d, R_d, S_f, E_f, I_f, S_h, E_h, I_h, H_h, R_h\right) = (\frac{\lambda_d}{\mu_d}, 0, 0, 0, \frac{\lambda_f}{\mu_f}, 0, 0, \frac{\lambda_h}{\mu_h}, 0, 0, 0, 0) \tag{4.34}$$

Since the DFE exists, the researched PDE system can be stable under the condition of equation (4.34).

4.3.2 Stability of equilibrium points

The previous section proved the existence of DFE in the researched PDE system. This section utilizes the basic reproduction number (R_0) to analyze the stability of two equilibrium points. A basic reproduction number is defined as the average number of cases increased by the primary case during the infection period, defined as the dominant eigenvalue of the next generation operator [174].

First, the realistic equilibrium point E_2^0 must be analyzed. Let $x = (E_d, I_d, E_f, I_f, E_h, I_h, S_d, R_d, S_f, S_h, H_h, R_h)$, the system also can be rewritten as

$$\dot{x} = \mathcal{F}(x) - \mathcal{V}(x), \tag{4.35}$$

where

$$\mathcal{F} = [\frac{b_{fd}\beta_{fd}I_f S_d}{N_d}, 0, \frac{b_{fd}\beta_{df}I_d S_f}{N_d}, 0, \frac{b_{fh}\beta_{fh}I_f S_h}{N_h}, 0,0,0,0,0,0,0]^T \tag{4.36}$$

$$\mathcal{V} = \begin{pmatrix} (\tau_d + \mu_d)E_d \\ -\tau_d E_d + (r_d + d_d + \mu_d)I_d \\ (\tau_f + \mu_f)E_f \\ -\tau_f E_f + (d_f + \mu_f)I_f \\ (\tau_h + \mu_h)E_h \\ -\tau_h E_h + (\delta_h + d_I + r_I + \mu_h)I_h \\ -\lambda_d + \frac{b_{fd}\beta_{fd}I_f S_d}{N_d} + \mu_d S_d \\ -r_d I_d + \mu_d R_d \\ -\lambda_f + \frac{b_{fd}\beta_{df}I_d S_f}{N_d} + \mu_f S_f \\ -\lambda_h + \frac{b_{fh}\beta_{fh}I_f S_h}{N_h} + \mu_h S_h \\ -\delta_h I_h + (d_H + r_H + \mu_h)H_h \\ -r_H H_h - r_I I_h + \mu_h R_h \end{pmatrix} \tag{4.37}$$

Next, the derivatives of $\mathcal{F}$ and $\mathcal{V}$ at DFE E_0^2 must be considered. The last six dimensions of $\mathcal{F}$ are 0, which correspond to $S_d, R_d, S_f, S_h, H_h, R_h$ in x. To simplify, the first six dimensions of $\mathcal{F}$ must correspond to $E_d, I_d, E_f, I_f, E_h, I_h$ in x. Thus, the derivatives of $\mathcal{F}$ and $\mathcal{V}$ can be simplified:

$$D\mathcal{F}(E_2^0) = \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.38}$$

$$D\mathcal{V}(E_2^0) = \begin{pmatrix} V & 0 \\ J_1 & J_2 \end{pmatrix}, \tag{4.39}$$

where F and V are 6×6 matrices assigned by

$$F = \frac{D\mathcal{F}_i}{Dx_j}(E_2^0) \qquad (1 \le i, j \le 6) \tag{4.40}$$

$$V = \frac{D\mathcal{V}_i}{Dx_j}(E_2^0) \qquad (1 \le i, j \le 6) \tag{4.41}$$

At DFE E_2^0, $N_d = S_d = \frac{\lambda_d}{\mu_d}$, $N_f = S_f = \frac{\lambda_f}{m_f + \mu_f}$, $N_h = S_h = \frac{\lambda_h}{\mu_h}$. Thus,

$$F = \begin{pmatrix} 0 & 0 & 0 & b_{fd}\beta_{fd} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{b_{fd}\beta_{df}\lambda_f\mu_d}{\lambda_d\mu_f} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{fh}\beta_{fh} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.42}$$

$$V = \begin{pmatrix} \tau_d + \mu_d & 0 & 0 & 0 & 0 & 0 \\ -\tau_d & r_d + d_d + \mu_d & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau_f + \mu_f & 0 & 0 & 0 \\ 0 & 0 & -\tau_f & d_f + \mu_f & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_h + \mu_h & 0 \\ 0 & 0 & 0 & 0 & -\tau_h & \delta_h + d_I + r_I + \mu_h \end{pmatrix}$$

(4.43)

The next objective is to calculate the basic reproduction number, which is the spectral radius of the next-generation matrix FV^{-1}. Moreover, the eigenvalues of J_2 have positive real parts, resulting in

$$V^{-1} = \begin{pmatrix} k_1 & 0 & 0 & 0 & 0 & 0 \\ \tau_d k_1 k_2 & k_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_3 & 0 & 0 & 0 \\ 0 & 0 & \tau_f k_3 k_4 & k_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_5 & 0 \\ 0 & 0 & 0 & 0 & \tau_h k_5 k_6 & k_6 \end{pmatrix} \tag{4.44}$$

where

$k_1 = \frac{1}{\tau_d+\mu_d}$, $k_2 = \frac{1}{r_d+d_d+\mu_d}$, $k_3 = \frac{1}{\tau_f+\mu_f}$, $k_4 = \frac{1}{d_f+\mu_f}$, $k_5 = \frac{1}{\tau_h+\mu_h}$, $k_6 = \frac{1}{\delta_h+d_I+r_I+\mu_h}$.

Then

$$FV^{-1} = \begin{pmatrix} 0 & 0 & b_{fd}\beta_{fd}\tau_f k_3 k_4 & b_{fd}\beta_{fd}k_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{b_{fd}\beta_{df}\lambda_f\mu_d\tau_d k_1 k_2}{\lambda_d\mu_f} & \frac{b_{fd}\beta_{df}\lambda_f\mu_d k_2}{\lambda_d\mu_f} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{fh}\beta_{fh}\tau_f k_3 k_4 & b_{fh}\beta_{fh}k_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.45}$$

The eigenvalues of matrix FV^{-1} are can be calculated by

$$\lambda^3\left(\lambda^3-\frac{\lambda b_{fd}^2\beta_{fd}\beta_{df}\lambda_f\mu_d\tau_f\tau_d k_1k_2k_3k_4}{\lambda_d\mu_f}\right)=0 \tag{4.46}$$

Therefore, the basic reproduction number R_0 can be calculated using

$$R_0=\rho(FV^{-1})=b_{fd}\sqrt{\frac{\beta_{fd}\beta_{df}\lambda_f\mu_d\tau_f\tau_d k_1k_2k_3k_4}{\lambda_d\mu_f}} \tag{4.47}$$

According to Theorem 2 regarding locally asymptotically stability in [54], when E_2^0 is the DFE of the system, if $R_0\leq 1$, the DFE E_2^0 is locally asymptotically stable; if $R_0>1$, the DFE E_2^0 is unstable. Similarly, if $R_0=0$ for the first DFE point E_1^0, then E_1^0 is locally asymptotically stable.

4.3.3 Sensitivity analysis for Basic Reproduction Number and Attack Rate

In epidemiology, the disease reproduction number R_0 is defined as the expected number of secondary cases, in a completely susceptible population, by a typical infected individual during the entire period of the epidemic [175, 176]. The value of R_0 can reflect the transmissibility of the epidemic [177]. The higher value of R_0 means the epidemic is harder to be controlled. Typically, if $R_0>1$, meaning the disease can be spread in a certain population; otherwise if $R_0<1$, meaning the transmission cannot be sustained in the long run.

1. Based on Eq. (4.47), the R_0 of the researched ZVL epidemic is affected by multiple system parameters. Utilized the data from Table 4.1, we used the pseudo-random number algorithm to generate 1000 samples from the value range of the parameters, and calculated the corresponding value of R_0. The mean and confidence interval of the distribution is 4.6499 (95% CI: 2.4521-5.4514)for R_0. Meaning, under the condition of non-intervention, the ZVL epidemic has relatively high transmissibility. Moreover, to identify the key factors which determine more of the magnitude of R_0, this paper compared the R_0 by modifying each

parameter using 200 random samples from the value range and maintaining other parameters as the suggested values from Table 4.1. By comparing the R_0 change intervals, we know that the average biting rate to dogs (b_{fd}), the death rate of dogs (μ_d) and the death rate of sand flies (μ_f) are most sensitive to the basic reproduction number. The incubation period in dogs (τ_d) and incubation period in sand files (τ_f) are most insensitive to the basic reproduction number. The sensitivities of other parameters belong to the states of in-betweenness.

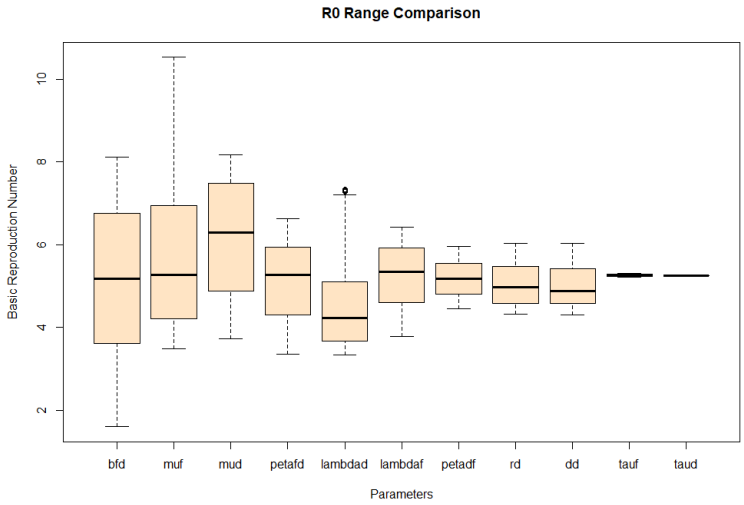


Figure 4. 5 R_0 sensitivity comparison

To further analysis the correlation between R_0 with parameters, this paper performed 2-dimensions sensitivity heat map for three most sensitive parameters in their corresponding value range (as shown in Figure 4.6). From the figure, we know that b_{fd} and μ_d are positive related to R_0, μ_f is negative related to R_0. To limit the transmission of R_0, the effective prevention strategy should be able to reduce b_{fd} and μ_d, or increase μ_f if possible.

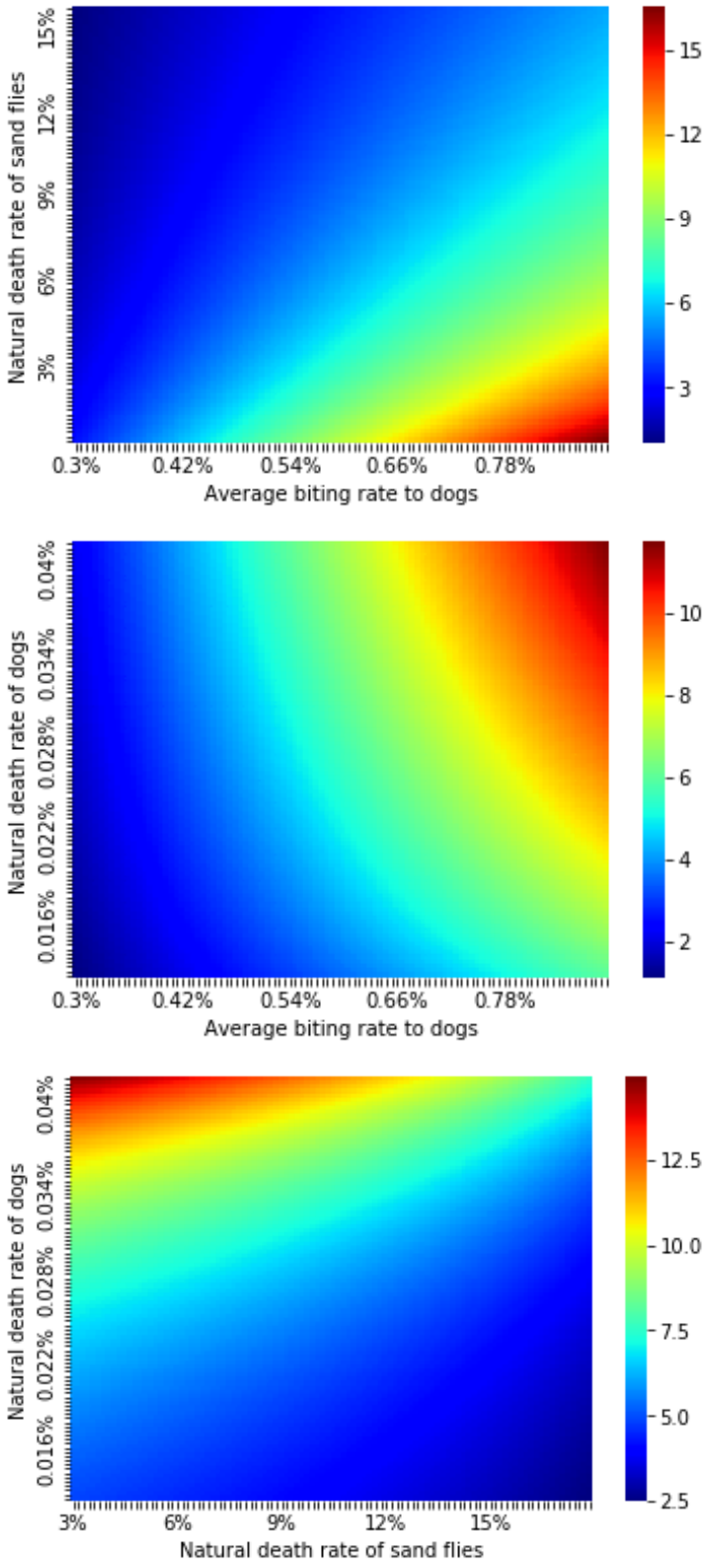


Figure 4. 6 R_0 sensitivity heat map

Instead of R_0, the attack rate is another important indicator, which can reflect the transmissibility of the researched epidemic. The attack rate is defined as the epidemic spread speed in a risk population. In this paper, the attack rate can be calculated by the number of new infectious divided by the current exposed population. Under the same parameter setting of the simulation in section 4.2.5, Figure 4.7 shows the attacks rates of researched ZVL epidemic through different ages. From the figure, we knew that the attack rates in the age group 0-15 years old are significantly higher than other age groups. The peak of the attack rates is around 50%, which is located at the age of 8-9. The average attack rates for all age groups is around 18.81%. This number is close to the real world observed range of ZVL attack rates, which is 17.5%-20.9% [178-180].

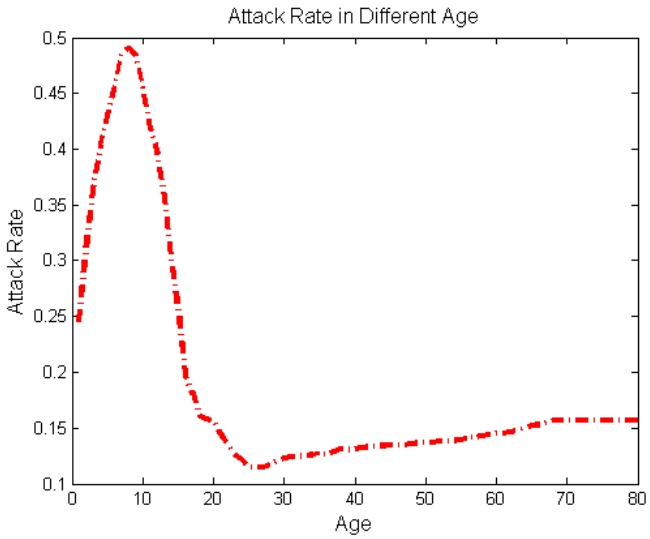


Figure 4. 7 ZVL attack rates with different age groups

To further understand the relationship between the attack rate with other model parameters, sensitivity analysis of the attack rates were set up by using different settings for each parameter. From Figure 4.8, we know that the attack rates are more sensitive to the average biting rate to dogs (b_{fd}), the death rate of dogs (μ_d), the death rate of sand flies (μ_f) and the natural recovery rate of dogs (r_d). Among these, μ_d has a positive correlation with the attack

rate; μ_f and r_d have a negative correlation with the attack rate. The relation between b_{fd} and the attack rate is complicated and irregular, since there may exist a bifurcation. To obtain the lowest attack rate by using the parameter from the value range, b_{fd} can be set as the lower bound ($b_{fd} = 0.03$).

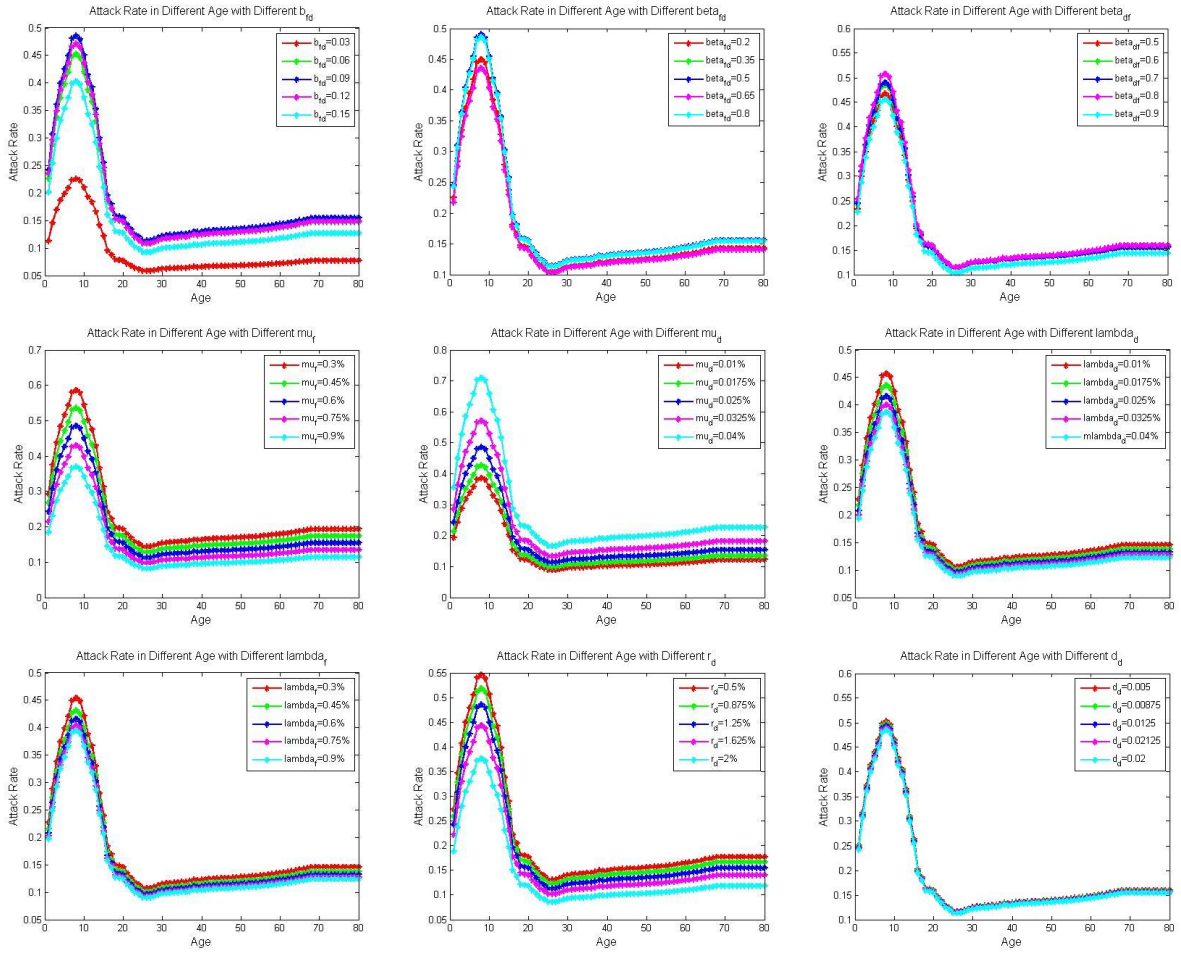


Figure 4. 8 Sensitivity analysis ZVL attack rates with different age groups

From the sensitivity analysis of R_0 and attack rates, we know that: (1) b_{fd}, μ_d and μ_f have significant effects on both R_0 and attack rates; (2) λ_f, β_{df} and β_{fd} can significant influence R_0, r_d can significant influence the attack rates; (3): the effects of other system parameters to R_0 and attack rates are relatively limited. To control the transmission of the ZVL, policymakers should use intervention strategies to limit the value of R_0 and attack rates. Therefore, a good

intervention strategy of ZVL should maintain b_{fd}, μ_d at a low level; maintain μ_f, λ_d, r_d and λ_f at a high level.

4.3.4 Endemic equilibrium and bifurcation analysis

In addition to DFE, endemic equilibrium, a dynamic equilibrium, can also occur at $R_0 > 1$. This equilibrium contains a specified point $x^* = (E_d{}^*, I_d{}^*, E_f{}^*, I_f{}^*, E_h{}^*, I_h{}^*, S_d{}^*, R_d{}^*, S_f{}^*, S_h{}^*, H_h{}^*, R_h{}^*)$ satisfied by $\frac{dS_d}{dt} = \frac{dE_d}{dt} = \frac{dI_d}{dt} = \frac{dR_d}{dt} = 0$; $\frac{dS_f}{dt} = \frac{dE_f}{dt} = \frac{dI_f}{dt} = 0$; $\frac{dS_h(a,t)}{dt} + \alpha\frac{dS_h(a,t)}{da} = \frac{dE_h(a,t)}{dt} + \alpha\frac{dE_h(a,t)}{da} = \frac{dI_h(a,t)}{dt} + \alpha\frac{dI_h(a,t)}{da} = \frac{dH_h(a,t)}{dt} + \alpha\frac{dH_h(a,t)}{da} = \frac{dR_h(a,t)}{dt} + \alpha\frac{dR_h(a,t)}{da} = 0$. Let $\theta_d{}^* = \frac{b_{fd}\beta_{fd}I_f{}^*}{N_d{}^*}$, $\theta_f{}^* = \frac{b_{fd}\beta_{df}I_d{}^*}{N_d{}^*}$, $\theta_h{}^* = \frac{b_{fh}\beta_{fh}I_f{}^*}{N_h{}^*}$, where $N_d{}^* = S_d{}^* + E_d{}^* + I_d{}^* + R_d{}^*$, $N_h{}^* = S_h{}^* + E_h{}^* + I_h{}^* + H_h{}^* + R_h{}^*$.

Solve the equations (4.1)-(4.7) in steady states, then

$$S_d{}^* = \frac{\lambda_d}{\theta_d{}^* + \mu_d} \tag{4.48}$$

$$E_d{}^* = \frac{k_1\theta_d{}^*\lambda_d}{\theta_d{}^* + \mu_d} \tag{4.49}$$

$$R_d{}^* = \frac{k_1k_2\theta_d{}^*\tau_d\lambda_d r_d}{\theta_d{}^* + \mu_d} \tag{4.50}$$

$$I_d{}^* = \frac{k_1k_2\theta_d{}^*\tau_d\lambda_d}{\theta_d{}^* + \mu_d} \tag{4.51}$$

$$S_f{}^* = \frac{\lambda_f}{\theta_f{}^* + \mu_f} \tag{4.52}$$

$$E_f{}^* = \frac{k_3\lambda_f\theta_f{}^*}{\theta_f{}^* + \mu_f} \tag{4.53}$$

$$I_f{}^* = \frac{k_3k_4\lambda_f\tau_f\theta_f{}^*}{\theta_f{}^* + \mu_f} \tag{4.54}$$

$$S_h{}^* = \frac{\lambda_h}{\theta_h{}^* + \mu_h} \tag{4.55}$$

$$E_h{}^* = \frac{k_5\theta_h{}^*\lambda_h}{\theta_h{}^*+\mu_h} \tag{4.56}$$

$$H_h{}^* = \frac{k_5k_6\theta_h{}^*\lambda_h\tau_h\delta_h}{(d_H+r_H+\mu_h)(\theta_h{}^*+\mu_h)} \tag{4.57}$$

$$R_h{}^* = \frac{k_5k_6\theta_h{}^*\lambda_h\tau_h[\delta_h r_H+r_I(d_H+r_H+\mu_h)]}{(d_H+r_H+\mu_h)(\theta_h{}^*+\mu_h)} \tag{4.58}$$

$$I_h{}^* = \frac{k_5k_6\theta_h{}^*\lambda_h\tau_h}{\theta_h{}^*+\mu_h} \tag{4.59}$$

By substituting Eqs. (4.48)-(4.59) into a quadratic equation in terms of $\theta_d{}^*$.

$$a(\theta_d{}^*)^2 + b\,\theta_d{}^* + c = 0, \tag{4.60}$$

where $a = \lambda_d\left[b_{fd}\beta_{df}\tau_d\mu_d k_2k_3(\mu_d + k_2(r_d+\mu_d)\tau_d) + \mu_d\mu_f k_3(\mu_d + 2k_2(r_d+\mu_d)\tau_d) + \tau_d{}^2k_2{}^2(r_d+\mu_d)^2\right]/k_2{}^2k_3{}^2k_4$, $b = \{\left[b_{fd}\beta_{df}\tau_d\mu_d + 2(r_d+\mu_d)\tau_d\mu_f\right]k_1k_2 + 2\mu_d\mu_f k_1 - \mu_f R_0{}^2\}/k_1{}^2k_2{}^2k_3k_4$, and $c = [\mu_d{}^2\mu_f\lambda_d(1-R_0{}^2)]/k_1{}^2k_2{}^2k_3k_4$.

Since there may have more than one solutions that satisfied the equation (4.60), meaning this epidemic system may exist more than one equilibriums with the same value of R_0. However, $\theta_d{}^*$ can only take positive values, since a negative value of $\theta_d{}^*$ represents the negative value of Infected sandfly populations. Therefore, this paper analyzed the potential situation of multiple equilibriums and corresponding bifurcations in Figure 4.9. The bifurcations are calculated by varying the value of b_{fd}, which is one of the component parameters of R_0. The rest of the parameters are fixed. For the generality of the simulation, the bifurcation analysis considered the R_0 as the horizontal axis, instead of b_{fd}.

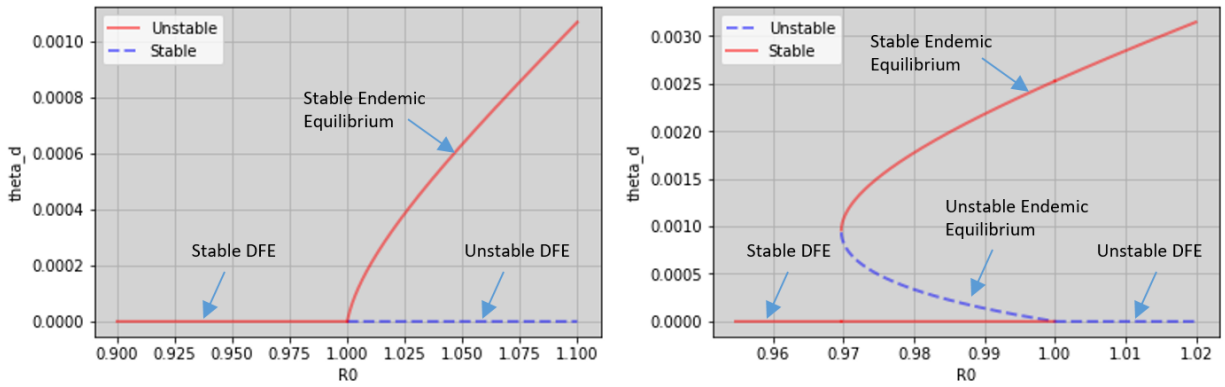


Figure 4. 9 (a) Forward bifurcation; (b) Backward bifurcation.

Since a is always greater than zero, the forward bifurcation existed if $b < 0$ (as shown in Figure 4.9 (a)). For that case, the bifurcation point is located at $R_0 = 1$, and there may have three potential situations: (1) when $R_0 < 1$, there is no positive solution of ${\theta_d}^*$, therefore the system has one stable DFE with no endemic equilibriums; (2) when $R_0 > 1$, there is one positive solution of ${\theta_d}^*$, therefore the system has one unstable DFE with one stable endemic equilibrium; (3) when $R_0 = 1$, there is one solution of ${\theta_d}^*$ located at zero, which means the DFE and endemic equilibrium are coincident.

In addition, the backward bifurcation exists when $b > 0$ (as shown in Figure 4.9 (b)). For that case, the bifurcation point is located at $R_0 = R_c$, where R_c satisfies $b^2 - 4ac = 0$. There have five potential situations: (1) when $R_0 < R_c$, the system has one stable DFE with no endemic equilibriums; (2) when $R_0 = R_c$, the system has one stable DFE with one endemic equilibrium, the stability of the endemic equilibrium is unknown based on the center manifold theory [181]; (3) when $R_c < R_0 < 1$, the system has one stable DFE with two endemic equilibriums, one of the endemic equilibriums is stable, another one is unstable; (4) when $R_0 = 1$, the DFE and unstable endemic equilibrium are coincident, and there existed another stable

equilibrium; (5) when $R_0 > 1$, the system has one unstable DFE with one stable endemic equilibrium.

4.4 Data Revised Methods and Prediction

4.4.1 Historical data for VL

Although the PDE model in Section 4.2.4 represents infection situations for various age groups, discrepancies occurred between the simulation results and real-world situations for the investigated countries. To ensure the PDE model accurate insights to the reality, this study revised and verified the model using real historical VL data. Because the PDE model does not reflect spatial characteristics in different regions and is only sensitive to the initial values of each simulation run, the numerical simulation results cannot accurately restore historical infectious data reported from various literature. Although some countries have similar initial values, disease transmissions vary widely; therefore, a feasible approach must be found to calibrate real data and simulation results.

Historical data was obtained from the data repository of WHO [11], a reliable database that includes reported VL cases each year from 1998 to 2018 for a variety of countries and regions. This database was used to calibrate the prediction from the PDE model according to spatial characteristics of various VL affected areas. In Table 4.4, we summarized VL infection data from the WHO database for eight most severe countries: India, South Sudan, Sudan, Brazil, Ethiopia, Bangladesh, Iraq, and Ethiopia.

Table 4. 4 VL infection cases for eight countries from 1998 to 2017

	India	South Sudan	Sudan	Brazil	Bangladesh	Iraq	Nepal	Ethiopia
2018	4360	1867	2584	3460	124	259	208	1828
2017	5758	3541	3894	4297	210	172	244	1490
2016	6249	4175	3810	3336	255	183	242	1593
2015	8500	2840	2829	3289	544	427	223	1990
2014	9241	7472	3415	3453	650	362	311	2705
2013	13851	2364	2389	3253	1103	575	325	1732
2012	20572	4353	5153	2770	1902	1045	575	2381
2011	33155	10468	7418	3894	2874	1167	886	2032

2010	28382	9166	6957	3716	3800	1843	708	1936
2009	24213	1907	4880	3693	4293	1549	824	1083
2008	33598	582	3310	3852	4840	1041	1371	1356
2007	44533	758	2788	3604	4932	836	1433	1579
2006	39173	1117	1827	3651	9379	1434	1531	2375
2005	32803	3141	3713	3597	6892	2028	1463	2585
2004	24479	3777	2619	3580	5920	3218	1588	1403
2003	18214	7722	7304	2971	6113	2526	2229	No data
2002	12140	4761	4066	2450	8110	3218	2029	No data
2001	12239	2596	1911	2549	4283	2893	1736	No data
2000	14753	1175	3667	4858	7640	2611	2090	No data
1999	12886	1612	4323	3624	5799	744	1794	No data
1998	13627	322	4922	1997	No data	874	1409	No data

Epidemic prediction results were calculated using the numerical simulation based on PDE model. The discrepancies were observed by comparing the difference between simulation data and WHO historical data for each country and each year. Using statistical methods, these discrepancies will be used as a calibration prediction the appraisal of future epidemics. In this research, the revised model can be denoted as following:

$$I_{predict}(t) = I_{model}(t) + \varepsilon_{predicted}(t), \tag{4.61}$$

where I_{model} is the simulation result for country in t year, $I_{predict}$ is the predicted number of infection cases for the country in t year, and $\varepsilon_{predicted}(t)$ is the predicted calibration through the statistical method, which is calculated using the model error, $\bar{\varepsilon}(t)$, in year t.

The differences between the model results to historically observed data is defined as the model error, $\bar{\varepsilon}(t)$, was considered to follow a time series pattern. Furthermore, $\varepsilon_{predicted}(t)$ represents the predicted deviation for the upcoming years, and it can be calculated by analyzing the time series of $\bar{\varepsilon}(t)$ using various statistic methods. Three types of moving average statistical methods were used in this research, including simple moving average (SMA), cumulated weight moving average (CWMA), and exponential moving average (EMA); and two types of autoregressive methods were also tested including autoregressive model with one and two input variables ($AR(1)$ and $AR(2)$) [182-184]. The method with the highest prediction accuracy rate was adopted to calculate the future model errors $\varepsilon_{predicted}(t)$.

4.4.2 Estimating calibration ε through validation

This section utilizes a validation process to determine the most accurate method to analysis time series for each country [185]. The criteria of the prediction accuracy depend on the predicted deviation. In this section, the predicted deviation $e_{predicted}(t)$ and predicted deviation ratio $r_{predicted}(t)$ are defined as follow:

$$e_{predicted}(t) = |I_{model}(t) + \varepsilon_{revised}(t) - I_{real}(t)| \quad (4.62)$$

$$r_{predicted}(t) = \left|\frac{I_{model}(t)+\varepsilon_{revised}(t)}{I_{real}(t)} - 1\right| \quad (4.63)$$

To calculate the $I_{model}(t)$, this study utilized real-world data for the number of infection cases from eight countries and simulated data calculated using the PDE model. The human population data is collected from a public data source [186]. The dog and the mosquito population are collected from several databases and literature [187-191].

Also, this study conducted following assumptions in simulation. The initial number of infected cases each year was assumed to be seven twelfth of a total number of infected cases last year since the infection period of ZVL is around seven months [192, 193]. Because ZVL occurs primarily in poor and neglected areas [92], the number of hospitalized humans was assumed to be 5% of the total number of infected cases. Also, most of the infected humans were assumed to recover after treatments, so the number of recovered humans primarily was from previous hospitalized individuals.

The validation process was used to determine which is the most accurate method for calculating calibrations. This study applied two-thirds to four-fifths of the datasets as training sets. Historical data in recent years (from 2014 to 2016) was used as the validation sets; data from earlier years were used as the training sets for all countries. The results of the average

predicted deviations and predicted deviation ratios for all eight countries are summarized in Table 4.5 and 4.6.

Table 4. 5 Average predicted deviation ratio $\bar{r}_{predicted}$ for the validation set in each country

	$\bar{r}_{model}$	$\bar{r}_{SMA}$	$\bar{r}_{CWMA}$	$\bar{r}_{EMA}$	$\bar{r}_{AR(1)}$	$\bar{r}_{AR(2)}$
Ethiopia	27.9%	30.8%	23.2%	24.7%	20.1%	32.1%
South Sudan	30.7%	32.1%	35.7%	46.2%	58.5%	92.47%
Sudan	30.5%	33.9%	35.4%	36.6%	58.5%	65.1%
Bangladesh	1111.6%	1295%	1049.7%	818.2%	729.3%	471.9%
India	207.2%	227.6%	233.2%	174.6%	92.6%	107.5%
Iraq	284.1%	304%	255.9%	202.7%	174.1%	184.1%
Nepal	376.9%	390.5%	291.0%	194.2%	93.7%	75.0%
Brazil	4.0%	4.8%	4.9%	3.3%	6.7%	10.0%

Table 4. 6 Average predicted deviation $\bar{r}_{predicted}$ for the validation set in each country

	$\bar{r}_{model}$	$\bar{r}_{SMA}$	$\bar{r}_{CWMA}$	$\bar{r}_{EMA}$	$\bar{r}_{AR(1)}$	$\bar{r}_{AR(2)}$
Ethiopia	664.7	738.4	549.5	552.2	514.1	727.5
South Sudan	1727.7	1767.6	1725	1970.2	2613.6	2589.3
Sudan	919.3	993.7	1047.5	1109.7	1930.6	1990.1
Bangladesh	4384	5237.7	4227.3	3340.9	3212	2297.3
India	15728.7	17431.1	17985.7	13325.8	7003.8	7745.5
Iraq	1241.7	1225.2	1129.6	895	819.4	833.6
Nepal	948.7	987.4	737.2	493.4	229.1	191
Brazil	119.3	123	161.8	106.2	201	337.7

Tables 4.5 and 4.6 show that the revised prediction methods did not get better results for areas in Sudan and South Sudan since results from the unrevised PDE model had the lowest average predicted deviation ratio and the near-lowest average predicted calibrations. The revised prediction methods performed better than the unrevised PDE model for the other countries. Based on these validation results, the most accurate methods of prediction for each country are obtained. Future prediction simulation should utilize the CWMA-revised method for Ethiopia, the unrevised PDE model for South Sudan and Sudan, the $AR(2)$-revised method for Bangladesh

and Nepal, the $AR(1)$-revised method for India and Iraq, and the EMA-revised method for Nepal.

4.4.3 Simulation of infection information

Using the results summarized in Section 4.4.2, the most suitable revised methods were selected for each country to predict a future VL prevalence trend. This section conducts a simulation using simulation results from the mathematical epidemic model and revised methods for each country. The simulation period was from 2019 to 2022 (data of 2019 was not public at the time of this writing. Therefore we predicted it by using the predicted calibration), and initial values of 2018 (as shown in Table 4.4) were used as initial values for the simulation. This section also counts and summarizes annual VL infection tallies for each researched country. Final revised simulation results are shown in Figure 4.10.

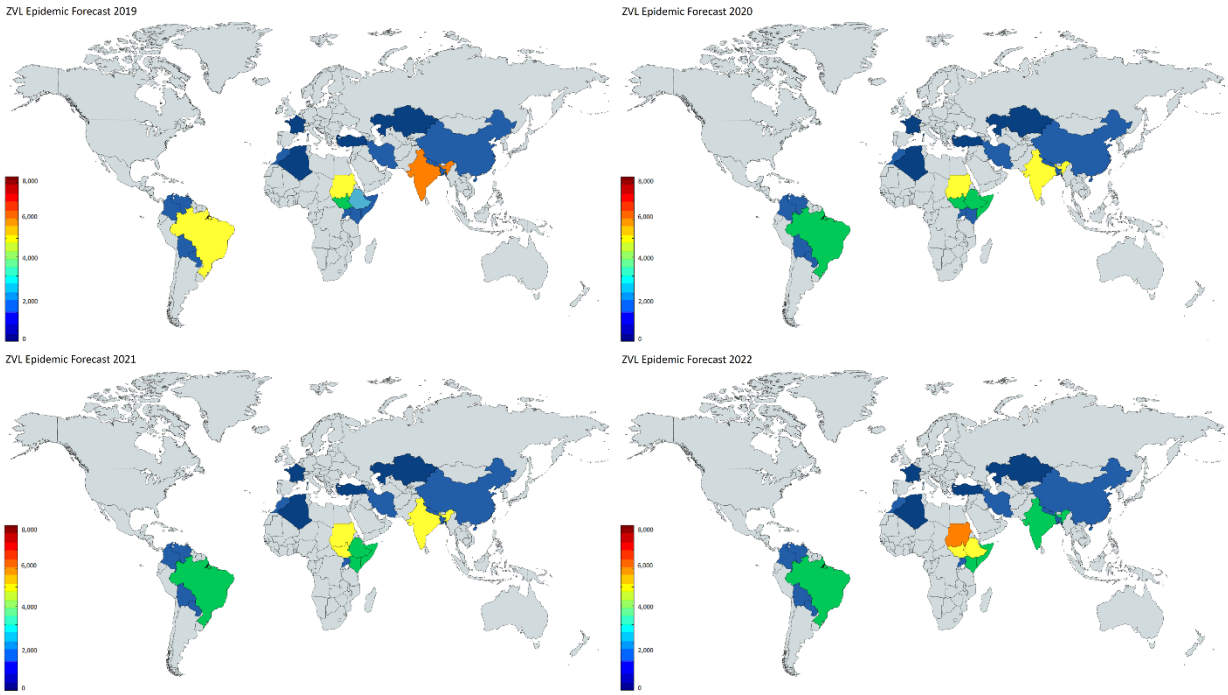
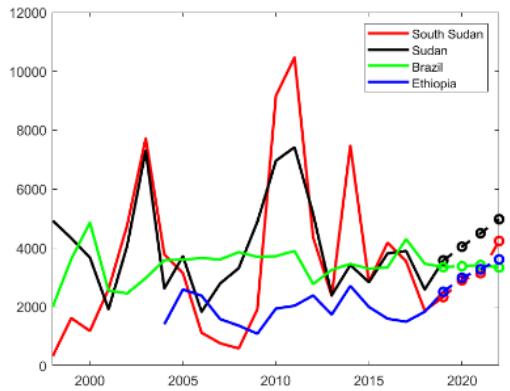


Figure 4. 10 Revised simulation results of VL infection population map from 2019 to 2022

The gray areas in Figure 4.10 represent countries without VL prevalence or countries that did not report VL infection data. Countries colored navy blue (e.g., Iran 2020) had very

sporadically reported VL cases, and countries colored orange (e.g., India 2019) had significant VL prevalence. We also did the monthly simulation of VL infections in eight of the most severe countries; the result was shown in supplement animation.

Worldwide VL prevalence is expected to decrease in general, especially in southern Asia, where the number of VL-predicted cases decreased significantly in India, Bangladesh, and Nepal in the last 10 years. India has previously been the most severely affected area in the world [11]. According to simulation results, Sudan is expected to become the most severely affected area around the year 2022. In the meantime, the VL epidemic is expected to spread to East Africa, specifically South Sudan, Ethiopia, Somalia, and Kenya. Although these countries have had 2,000 or less VL cases reported in the past ten years, reported VL cases are expected to increase to 4,000–5,000 by 2022. Moreover, countries such as Libya, Chad, Central African Republic, and the Democratic Republic of Congo may lack the government resources necessary to collect and report VL cases to the WHO, thereby preventing this model from estimating potential existing VL cases in these countries. Because these countries border Sudan or South Sudan, areas severely affected by VL, VL infection cases are assumed to be more than 1,000 in countries like Libya. Consequently, health organizations should consider more VL prevention or control strategies in eastern and northern Africa.



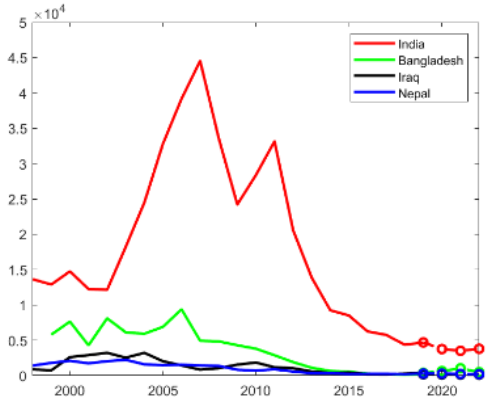


Figure 4. 11 (a) 1998-2022 VL infections in severely affected countries in Africa and Brazil; (b) 1998-2022 VL infections in severely affected countries in Asia.

In Figure 4.11, we combined historical VL infection data (from 1998 to 2018 in solid line) with the predicted VL infection data (from 2019 to 2022 in dotted line) of 8 the most severe countries in the world. In Brazil, the VL epidemic is relatively stable; the number of annual VL reported infected cases is around 3,000 to 4,000. Even though there are no signs of intensifying the VL outbreak in the next four years, Brazil is still one of the top 5 most severely countries in the world. For countries in Eastern African like Sudan, the VL prevalence is relatively fluctuant. There were three VL prevalence outbreak peaks in this area, which happen in around 2003, 2011 and 2014. Also, the predicted VL infections data shows that VL prevalence at Eastern African is in a resurgence in the next four years. The VL prevalence in Asia, especially in India, is the most serious in the world from around the last 20 years. However, this situation should be different in 2022; Eastern African may replace Asia to become the most severe area of VL. Since 2011, the number of VL reported infected population in Asia was decreasing persistently; our simulation forecasted that the VL infections number in Sudan would pass this number in Indian in 2022.

4.5 Discussion and Conclusions

This paper developed a two-dimensional PDE epidemic model to describe ZVL epidemic dynamics. This model considered ZVL transmissions among dogs, sand flies, and humans, as well as the age composition of human infections and exposed population. The analysis of equilibrium and stability showed that the researched ZVL system could reach the situation of disease-free with some conditions. Also, three statistical revised methods were considered to improve the accuracy of ZVL prevalence predictions. Finally, simulation experiments were

carried out for multiple regions and countries to attempt to predict the ZVL epidemic from 2019 to 2022.

Researchers have previously determined that *L. infantum*, a ZVL transmission mediator, has increasingly negative impacts on children [2]. This paper considered distributions of ZVL infection throughout various age groups. A summarization of reports from literature regarding ZVL historical prevalence in Brazil, Italy, Sudan, India, and China [21, 22, 23, 24, 25, 6] showed the infection density of children and teenagers (18 years old and under) to be consistently 5–10 times higher than corresponding adult groups.

Using KDE, a continuous infection rate function was built to reflect the variation of infection possibilities for various age groups. This calculated infection rate function replaced the constant infection rate used in a previous model [12]. To more accurately describe the age distribution for humans in the established ODE model, the age dimension was added to transfer the previous ODE model into the PDE model. Therefore, the new PDE model explains the specific kinds of the human population (susceptible, exposed, infected, hospitalized, and recovered) at the time axis and age axis. The new PDE model also retained the transmission models of dogs and sand flies from the ODE dynamic system, making the new PDE model more detailed and realistic for calibrations using real-world data.

Analysis of PDE system stability revealed two DFE points in the PDE model. Reproduction numbers of these two equilibrium points were discussed in this paper; endemic equilibrium points were also discovered and explored. Prerequisites were established to allow the researched system to achieve stability.

However, the PDE model is a general framework, meaning it only depends on the initial number of system variables without considering the specific circumstances of VL in each

country. Therefore, revised correlation methods are necessary to conduct predictions with fewer calibrations. Reported VL cases from WHO annual reports (1998–2016) were used as the real dataset to train and validate prediction results using model simulation results. The cross-validation method was carried out by categorizing the reported data from 1998 to 2013 as the training set and data from 2014 to 2016 as the validation set. Forecasting methods such as MA and AR were used to revise the prediction data. Final revised results could improve the accuracy of predictions in the validation set from 16.8% to 80.1% in each researched country. Brazil demonstrated the most accurate prediction results, with a revised average predicted deviation ratio of only 3.3% and an average predicted calibration of infection numbers of only 106.2. The final prediction was accomplished using both the PDE model simulation and revised model to predict VL prevalence for current VL-affected countries in the next four years (i.e., 2019–2022). Simulation results predicted the global decline of VL prevalence due to the decreasing infected population in Asia, specifically in India, Bangladesh, and Nepal. VL prevalence in South America (mainly Brazil) is expected to remain stable; however, northern and eastern Africa demonstrated an increasing trend of VL infection cases.

According to this research, the following suggestions were made concerning VL prevention:

1. Specified VL prevention and control strategies should be emphasized for children and teenage age groups (age 0-20).

2. The WHO and other health agencies should prioritize prevention in southern Asia, specifically the most severely affected areas in eastern and northern Africa.

3. Successful VL prevention in India and Bangladesh should be generalized and publicized into Africa and Brazil.

4. More complete and standardized collecting and reporting of confirmed VL infection cases should be established, including consideration of age, gender, and geographical regions.

5. The complete report system of VL cases should be completed in developing countries, especially Libya, Chad, Central African Republic, and the Democratic Republic of Congo.

Future work should include two considerations. First, the PDE dynamic model and data-revised methods were described as two processes in the current paper; however, the revised method should be considered a part of the PDE system, and collection between the dynamic model and real data should be reinforced. Second, the numerical control strategy should be discussed in extension research. Use of optimal control strategies or parameter control to reflect real-world VL prevention could help WHO estimate the effects of intervention strategies or actions.

Chapter 5 - A Memetic Algorithm for Solving Optimal Control Problems of Zika Virus Epidemic with Equilibriums and Backward Bifurcation Analysis

Chapter 5 is based on the manuscript "A Memetic Algorithm for Solving Optimal Control Problems of Zika Virus Epidemic with Equilibriums and Backward Bifurcation Analysis" Published in Communications in Nonlinear Science and Numerical Simulation [35].

Abstract

The large-scale outbreak of Zika virus in 2015–2016 attracted global attention. As of January 2018, 137,515 cases of Zika virus were confirmed in the United States and Brazil, and 223,477 cases were confirmed in the world by PAHO and WHO. This paper utilizes an existed mathematics model in the Zika virus, then analyzes the stability and bifurcation by changing the closed population system to an open system. Moreover, this paper establishes an optimal control problem associated with the open population system based on several popular disease intervention strategies frequently used by public health agencies to mitigate the Zika epidemics. Comparisons of traditional Pontryagin's maximum principle and a new memetic algorithm are conducted for different intervention strategies. Also, our computational results suggest that continuous optimal control strategies may not be practical in real-world applications. Instead, the memetic algorithm-based discrete is relatively easy to be implemented.

Keywords: Zika virus, memetic algorithm, optimal control, bifurcation analysis, equilibrium.

5.1. Introduction

The tropical Zika virus is a type of flavivirus similar to the yellow fever virus, dengue virus, and West Nile virus [194]. The Zika virus is primarily transmitted via bites of infected

female *Aedes* genus mosquitoes, but researchers have also discovered that the Zika virus can be transmitted sexually within the human population [195]. The first occurrence of the Zika virus was in Uganda in 1947, followed by reported epidemics in West Africa and Asia from the 1960s to the 1980s. Then, the Zika virus was transmitted to India, Indonesia, Malaysia, and Pakistan (1969-1983); the island of Yap (2007); French Polynesia (2013 -2014); Brazil (2015-2016) [196]. Millions of human cases of Zika virus infections have been reported from more than 20 countries around the world. For the epidemic in the Americas from 2015 to 2016, there are 137,515 cases of Zika virus confirmed in the United States and Brazil, and 223,477 cases were confirmed in the world [197].

Symptoms of Zika virus infection include rash, conjunctivitis, subjective fever, arthralgia, and arthritis [198]. Zika virus transmission also increases the prevalence of microcephaly in newborn infants of infected mothers [199]. Although health organizations and disease control centers spent significant human and financial resources to combat the global Zika epidemics, confirmed Zika vaccines were not in clinical use until 2017 [200]. For example, in the United States, the total economic impacts of the Zika virus in six southern states (i.e., Alabama, Florida, Georgia, Louisiana, Mississippi, and Texas) exceeded 5 billion dollars in 2016, the direct cost of Zika virus prevention surpassed 1.1 billion dollars that same year [201]. Disease control centers are continually seeking to efficiently utilize limited funds to reduce the impacts of Zika virus epidemics [202].

The outbreak at Brazil in 2015–2016 prompted researchers to focus more on modeling and numerical control of the Zika virus. Several Ordinary Differential Equation (ODE) dynamic systems were built to analyze both 2013–2014 French Polynesian Zika outbreak and 2015–2016 South American Zika outbreak [203, 204]. Isea and Lonngren considered a model that can

simultaneously analyze dengue, chikungunya, and Zika virus [205]. Agusto *et al.* built a dynamic system to investigate the sexual transmission route of Zika [206], and other researchers tested various optimal control strategies to limit the transmission of the Zika virus [207, 208]. More widely, combinations of dynamic modeling and optimal control have also been applied to control various kinds of infectious diseases [41, 44, 209, 210].

This research focuses on the equilibrium and bifurcation analysis of a nonlinear dynamic system. Also, this research explored the optimal controls and related possible mitigating interventions regarding the epidemics caused by Zika virus. Firstly, the disease-free and endemic equilibriums are discussed for the given system and equilibrium conditions of the disease reproductive rate, R_0, and backward bifurcation were investigated. Secondly, three classes of commonly control strategies were introduced, including the release of insects with dominant lethality (RIDL), the use of endosymbiotic bacteria (Wolbachia) to prevent arboviruses replicating within the mosquitoes, and infection rate reduction among the human population via sexual contacts. Then, this paper compares three control strategies to mitigate different patterns of Zika virus epidemics. In this research, the Pontryagin's maximum principle was used initially to solve the optimal disease control problem while minimizing the negative influence of Zika virus in a convex objective function. Then, considering possible non-convex objective functions may exist in other specific cases, a new memetic algorithm (MA) was designed to compare with the conventional optimal control method and limit the transmission of Zika virus. Computational results from simulation experiments shown that the proposed Memetic Algorithm is more effective for the discrete step control strategies with additional implementation constraints.

The rest of this chapter is organized as follows. Section 5.2 introduces the underlying mathematical model of Zika virus with an open population system. Section 5.3 analyzes

corresponding disease equilibrium and bifurcation for the mathematical model. The main results of finding effective optimal Zika epidemics control/intervention strategies are presented in Section 5.4. Finally, conclusions and future research work are summarized in Section 5.5.

5.2. Mathematical Model of Zika Virus

This section extends the ODE Zika dynamic system proposed by Gao *et al.* [204] from a closed population model to an open population model. Since an extended period of study is considered in this research, to ensure the simulation meaningful and accurate insights of the reality, we changed the closed population model to the open population model. Figure 1. explains the relationship between system variables, which includes $S_h(t)$ (susceptible human), $E_h(t)$ (exposed human), $I_{h1}(t)$ (symptomatically infected human), $I_{h2}(t)$ (convalescent human), $A_h(t)$ (asymptomatically infected human), $R_h(t)$ (recovered human), $S_v(t)$ (susceptible vectors), $E_v(t)$ (exposed vectors) and $I_v(t)$ (infectious vectors). The blue solid lines represent the infection and recovery processes for both humans and vectors; blue dash lines represent the sexual transmissions among the human population; red dash lines represent the Zika Virus transmissions between humans and vectors.

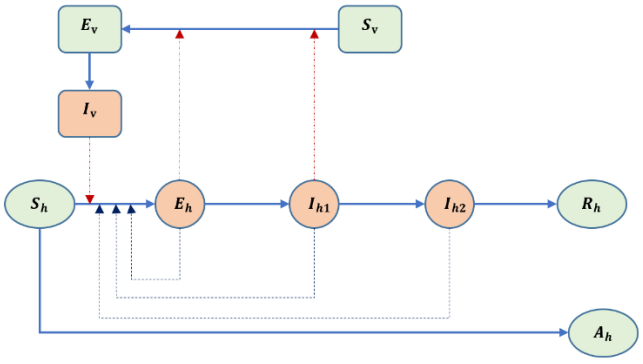


Figure 5. 1 System diagram of Zika virus transmission model

The underlying assumptions of modeling include the following: (1) Symptomatically infected populations are contagious to humans (via sexual transmission) and mosquitoes (via biting) during the incubation period. (2) Symptomatic humans are more contagious to the susceptible population and mosquitoes after the incubation periods. (3) An asymptomatically infected population presumably cannot infect humans and susceptible populations. (4) The sexual ratio of humans is 1:1, and the difference of epidemiological factors between genders is negligible. (5) Consider the short lifespan of the mosquitoes; this model doesn't include the recovered state for infected mosquitoes.

According to preceding assumptions [204, 206-208], the ODE dynamic system in this model is in the form of SEIR (susceptible, exposed, infectious, and recovered), as shown in Eqs. (5.1)–(5.9), where the human population can be classified as $S_h(t)$, $E_h(t)$, $I_{h1}(t)$, $I_{h2}(t)$, $A_h(t)$, and $R_h(t)$. The total number of humans at time t is defined as $N_h(t) = S_h(t) + E_h(t) + I_{h1}(t) + I_{h2}(t) + A_h(t) + R_h(t)$; where the mosquito population is classified as $S_v(t)$, $E_v(t)$, and $I_v(t)$. The total number of mosquitoes at time t is defined as $N_v(t) = S_v(t) + E_v(t) + I_v(t)$. Paramter descriptions are listed in Table 5.1.

$$\frac{dS_h}{dt} = \lambda_h - ab\frac{I_v}{N_h}S_h - \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h - \mu_h S_h \tag{5.1}$$

$$\frac{dE_h}{dt} = \theta\left(ab\frac{I_v}{N_h}S_h + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h\right) - v_h E_h - \mu_h E_h \tag{5.2}$$

$$\frac{dI_{h1}}{dt} = v_h E_h - \gamma_{h1} I_{h1} - \mu_h I_{h1} \tag{5.3}$$

$$\frac{dI_{h2}}{dt} = \gamma_{h1} I_{h1} - \gamma_{h2} I_{h2} - \mu_h I_{h2} \tag{5.4}$$

$$\frac{dA_h}{dt} = (1-\theta)\left(ab\frac{I_v}{N_h}S_h + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h\right) - \gamma_h A_h - \mu_h A_h \tag{5.5}$$

$$\frac{dR_h}{dt} = \gamma_{h2} I_{h2} + \gamma_h A_h - \mu_h R_h \tag{5.6}$$

$$\frac{dS_v}{dt} = \mu_v N_v - ac\frac{\eta E_h + I_{h1}}{N_h}S_v - \mu_v S_v \tag{5.7}$$

$$\frac{dE_v}{dt} = ac\frac{\eta E_h + I_{h1}}{N_h}S_v - (v_v + \mu_v)E_v \tag{5.8}$$

$$\frac{dI_v}{dt} = v_v E_v - \mu_v I_v \tag{5.9}$$

Table 5. 1 Parameter descriptions and values

Parameter	Description	Range	Value	Reference

λ_h	Recruitment rate of susceptible humans	Depends on places	2000	[211]
a	Mosquito biting rate (bites per mosquito per day)	0.3–1	0.5	[212]
b	Transmission probability from an infected mosquito to a susceptible human per bite	0.1–0.75	0.4	[212]
c	Transmission probability from an asymptomatically infected human to a susceptible mosquito per bite	0.3–0.75	0.5	[213]
η	Relative human-to-mosquito transmission probability of exposed humans to symptomatically infected humans	0–0.3	0.1	[204, 214]
β	Transmission probability from susceptible humans to symptomatically infected humans	0.001–0.1	0.004	[204, 214]
κ	Relative human-to-human transmissibility of exposed humans to symptomatic humans	0–1	0.6	[204, 214]
τ	Relative human-to-human transmission probability of convalescent to symptomatic humans	0–1	0.3	[204, 214]
θ	Proportion of symptomatic infections	0.10–0.27	0.18	[198]
m	Average ratio of mosquitoes to humans (mosquitoes per human)	1–10	5	[215]
$1/\nu_h$	Intrinsic incubation period in humans (days)	2–7	5	[216]
$1/\nu_v$	Extrinsic incubation period in mosquitoes (days)	8–12	10	[212]
$1/\gamma_{h1}$	Duration of acute phase (days)	3–7	5	[216]
$1/\gamma_{h2}$	Duration of convalescent phase (days)	14–30	20	[195, 217]
$1/\gamma_h$	Duration of asymptomatic infection (days)	5–10	7	[204]
$1/\mu_v$	Mosquito lifespan (days)	4–50	20	[212, 213]
$1/\mu_h$	Human lifespan (years)	30-122	79	[218, 219]

5.3. Stability and Bifurcation Analysis

5.3.1 Disease-Free Equilibrium Analysis

The Disease-Free Equilibrium (DFE) point in the given ODE dynamic system (Eqs. (5.1)–(5.9)) is located at the system stable points, where $E_h = I_{h1} = I_{h2} = E_v = I_v = 0$.

Assumed $E^* = ({S_h}^*, {E_h}^*, {I_{h1}}^*, {I_{h2}}^*, {A_h}^*, {R_h}^*, {S_v}^*, {E_v}^*, {I_v}^*)$ as the DFE point, then it reached the disease-free condition $E^* = ({S_h}^*, 0,0,0, {A_h}^*, {R_h}^*, {S_v}^*, 0,0)$. Since E^* is in a stable status, it can satisfy $\frac{dS_h}{dt} = \frac{dE_h}{dt} = \frac{dI_{h1}}{dt} = \frac{dI_{h2}}{dt} = \frac{dA_h}{dt} = \frac{dR_h}{dt} = \frac{dS_v}{dt} = \frac{dE_v}{dt} = \frac{dI_v}{dt} = 0$. Therefore, the generalized DFE point is

$$E^* = (\lambda_h/\mu_h\,, 0,0,0,0,0, \lambda_v/\mu_v\,, 0,0). \tag{5.10}$$

E^* can also include extreme situations such as no mosquitoes in the research area ($\lambda_v = 0$) or no newly born humans in the research time ($\lambda_h = 0$), although these extreme situations may never occur in the real world.

To calculate R_0, the ODE dynamic system (Eqs. (5.1)–(5.9)) can be divided into nonlinear part F and linear part V. For general point $x = (S_h(t), E_h(t), I_{h1}(t), I_{h2}(t), A_h(t), S_v(t), E_v(t), I_v(t), R_h(t))$, the system can be transformed as follows:

$$\dot{x} = F(x) - V(x), \tag{5.11}$$

where

$$F = \begin{pmatrix} -(ab\frac{I_v}{N_h} + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h})S_h \\ \theta(ab\frac{I_v}{N_h} + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h})S_h \\ 0 \\ 0 \\ (1-\theta)(ab\frac{I_v}{N_h} + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h})S_h \\ -ac\frac{\eta E_h + I_{h1}}{N_h} \\ ac\frac{\eta E_h + I_{h1}}{N_h} \\ 0 \\ 0 \end{pmatrix} \tag{5.12}$$

$$V = \begin{pmatrix} -\lambda_h + \mu_h S_h \\ v_h E_h + \mu_h E_h \\ -v_h E_h + \gamma_{h1} I_{h1} + \mu_h I_{h1} \\ -\gamma_{h1} I_{h1} + \gamma_{h2} I_{h2} + \mu_h I_{h2} \\ \gamma_h A_h + \mu_h A_h \\ -\lambda_v + \mu_v S_v \\ v_v E_v + \mu_v E_v \\ -v_v E_v + \mu_v I_v \\ -\gamma_{h2} I_{h2} - \gamma_h A_h + \mu_h R_h \end{pmatrix} \tag{5.13}$$

The derivatives of F and V should also be considered. The remaining dimensions of F and V are zero, which corresponds to $R_h(t)$ in x. To simplify, only the first eight dimensions of F and V is considered: $S_h(t), E_h(t), A_h(t), S_v(t), E_v(t), I_{h1}(t), I_{h2}(t), I_v(t)$. Therefore, the derivatives of F and V can be simplified using the following matrix:

$$\dot{F} = \frac{dF}{dx}(E^*) = \begin{pmatrix} F_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{5.14}$$

$$\dot{V} = \frac{dV}{dx}(E^*) = \begin{pmatrix} V_1 & 0 \\ V_2 & V_3 \end{pmatrix} \tag{5.15}$$

To calculate R_0, V_2 and V_3 are not considered since the corresponding position of $\dot{F}$ matrix is zero. At the E^*, F_1 and V_1 are

$$F_1 = \begin{bmatrix} 0 & -\beta\kappa & -\beta & -\beta\tau & 0 & 0 & 0 & -ab \\ 0 & \theta\beta\kappa & \theta\beta & \theta\beta\tau & 0 & 0 & 0 & \theta ab \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\theta)\beta\kappa & (1-\theta)\beta & (1-\theta)\beta\tau & 0 & 0 & 0 & (1-\theta)a \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -ac\eta\mu_h/\lambda_h & -ac\mu_h/\lambda_h & 0 & 0 & 0 & 0 & 0 \\ 0 & ac\eta\mu_h/\lambda_h & ac\mu_h/\lambda_h & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.16}$$

$$V_1 = \begin{bmatrix} \mu_h & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & v_h+\mu_h & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -v_h & \gamma_{h1}+\mu_h & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_{h1} & \gamma_{h2}+\mu_h & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_h+\mu_h & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_v & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v_v+\mu_v & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -v_v & \mu_v \end{bmatrix} \tag{5.17}$$

Then

$$V_1^{-1} = \begin{bmatrix} k_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & v_h k_2 k_3 & k_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & y_{h1} k_3 k_4 & k_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & k_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v_v k_6 k_7 & k_6 \end{bmatrix} \tag{5.18}$$

where $k_1 = 1/\mu_h$; $k_2 = 1/(v_h + \mu_h)$; $k_3 = 1/(\gamma_{h1} + \mu_h)$; $k_4 = 1/(\gamma_{h2} + \mu_h)$; $k_5 = 1/(\gamma_h + \mu_h)$; $k_6 = 1/\mu_v$; $k_7 = 1/(v_v + \mu_v)$.

R_0 is defined as the nonnegative eigenvalue of $F_1 {V_1}^{-1}$:

$$R_0 = \frac{\theta\beta\kappa k_2 + \sqrt{(\theta\beta\kappa\, k_2)^2 + 4\theta a^2 bc\mu_h k_2 k_6 (\eta + v_h k_3)(1 - v_v k_7)/\lambda_h}}{2} \tag{5.19}$$

According to the theorem of regrading local asymptotical stability [54]. For DFE point E^*, if $R_0 \leq 1$, then E^* is locally asymptotically stable; if $R_0 > 1$, then E^* is unstable.

5.3.2 Endemic Equilibrium and Bifurcation

Previous research analyzed backward bifurcation in a variety of epidemic dynamic systems [50, 174]. In order to discuss the backward bifurcation for the ODE dynamic system in Eqs. (5.1)–(5.9), the endemic equilibrium points must be identified and analyzed. Assume an endemic equilibrium exists in $E^{**} = ({S_h}^{**}, {E_h}^{**}, {I_{h1}}^{**}, {I_{h2}}^{**}, {A_h}^{**}, {R_h}^{**}, {S_v}^{**}, {E_v}^{**}, {I_v}^{**})$. Since E^{**} is stable, it can satisfy $\frac{dS_h}{dt} = \frac{dE_h}{dt} = \frac{dI_{h1}}{dt} = \frac{dI_{h2}}{dt} = \frac{dA_h}{dt} = \frac{dR_h}{dt} = \frac{dS_v}{dt} = \frac{dE_v}{dt} = \frac{dI_v}{dt} = 0$. Let $\theta_{mh} = ab\frac{I_v}{N_h}$ represents the infection rate of human by mosquito; then $\theta_{hh} = \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}$ represents the infection rate of human by human and $\theta_{hm} = ac\frac{\eta E_h + I_{h1}}{N_h}$ represents the infection rate of mosquito by human.

Solving the system of equations at its steady-state, results can be summarized as follows:

$${S_h}^{**} = \frac{\lambda_h}{\theta_{mh}+\theta_{hh}+\mu_h} \tag{5.20}$$

$${E_h}^{**} = \frac{\theta\lambda_h(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_2 \tag{5.21}$$

$${I_{h1}}^{**} = \frac{\theta\lambda_h v_h(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_2k_3 \tag{5.22}$$

$${I_{h2}}^{**} = \frac{\theta\lambda_h v_h\gamma_{h1}(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_2k_3k_4 \tag{5.23}$$

$${A_h}^{**} = \frac{(1-\theta)\lambda_h(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_5 \tag{5.24}$$

$${R_h}^{**} = (\frac{\theta\lambda_h v_h\gamma_{h1}\gamma_{h2}(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_2k_3k_4 + \frac{(1-\theta)\lambda_h\gamma_h(\theta_{mh}+\theta_{hh})}{(\theta_{mh}+\theta_{hh}+\mu_h)}k_5)k_1 \tag{5.25}$$

$${S_v}^{**} = \frac{\lambda_v}{\theta_{hm}+\mu_v} \tag{5.26}$$

$${E_v}^{**} = \frac{\theta_{hm}\lambda_v}{\theta_{hm}+\mu_v}k_7 \tag{5.27}$$

$${I_v}^{**} = \frac{\theta_{hm}\lambda_v v_v}{\theta_{hm}+\mu_v}k_6k_7 \tag{5.28}$$

$${N_h}^{**} = \lambda_h k_1 \tag{5.29}$$

$$N_v^{**} = \lambda_v k_6 \tag{5.30}$$

Eqs. (5.20)–(5.30) show that the endemic equilibrium E^{**} is identical to DFE E^* if $\theta_{mh} = \theta_{hh} = \theta_{hm} = 0$. If $\theta_{mh}, \theta_{hh}, \theta_{hm}$ are not equal to zero, then Eqs. (5.20)–(5.30) can be substituted into $\theta_{mh}, \theta_{hh},$ and θ_{hm}:

$$\theta_{mh} = \frac{k_1 {\theta_{hh}}^2 + [1 - \beta\theta k_2(\kappa + v_h k_3 + \tau v_h \gamma_{h1} k_3 k_4)]\theta_{hh}}{\beta\theta k_2(\kappa + v_h k_3 + \tau v_h \gamma_{h1} k_3 k_4) - k_1 \theta_{hh}} \tag{5.31}$$

$$\theta_{hh} = \frac{\lambda_v k_6(\theta_{mh} + \mu_h)\theta_{hm} - ac\theta\lambda_h k_2(\eta + v_h k_3)\theta_{mh}}{ac\theta\lambda_h k_2(\eta + v_h k_3) - \theta_{hm}} \tag{5.32}$$

$$\theta_{hm} = \frac{\lambda_h k_1 \mu_v \theta_{mh}}{ab\lambda_v v_v k_7 - \lambda_h k_1 \theta_{mh}} \tag{5.33}$$

If Eqs. (5.31) and (5.33) are substituted into Eq. (5.32) and the zero solution is abbreviated, then

$$a_0 {\theta_{hh}}^2 + b_0 \theta_{hh} + c_0 = 0, \tag{5.34}$$

where $a_0 = {k_1}^2 k_6 \lambda_v - {k_1}^3 \lambda_h \mu_v$, $b_0 = \theta\beta k_2(\kappa + v_h k_3 + \tau v_h \gamma_{h1} k_3 k_4)\big(\lambda_h k_1 \mu_v + 2{k_1}^2 \lambda_h \mu_v - 2\lambda_v k_1 k_6\big) - \lambda_h k_1 \mu_v + 2\lambda_v k_1 k_6 + {k_1}^2 \lambda_h ac\theta\lambda_h k_2(\eta + v_h k_3)$, $c_0 = \theta\beta k_2(\kappa + v_h k_3 + \tau v_h \gamma_{h1} k_3 k_4)[(\lambda_v k_6 - \lambda_h k_1 \mu_v)\theta\beta k_2(\kappa + v_h k_3 + \tau v_h \gamma_{h1} k_3 k_4) + 1]$.

Based on the quadratic Eq. (5.34), θ_{hh} may have multiple solutions, and when the solutions are applied to Eqs. (5.31)–(5.33) the given system may have multiple endemic equilibriums.

Lemma 5.1:

The system has

(i): One unique endemic equilibrium at $\theta_{hh} = -b_0/a_0$ if $b_0 < 0$ and $c_0 = 0$;

(ii): One unique endemic equilibrium at $\theta_{hh} = -b_0/2a_0$ if ${b_0}^2 - 4a_0 c_0 = 0$;

(iii): Two unique endemic equilibriums at $\theta_{hh} = (-b_0 \pm \sqrt{{b_0}^2 - 4a_0c_0})/2a_0$ if $b_0 < 0$ and $c_0 > 0$ and ${b_0}^2 - 4a_0c_0 > 0$;

(iv): No endemic equilibrium.

For verification, we use the parameters in their possible ranges and consider $a = 0.4, b = 0.5, c = 0.4, \eta = 0.2, \beta = 0.05, \kappa = 0.4, \tau = 0.5, m = 6, v_h = 0.3, v_v = 0.1, \gamma_{h1} = 0.25, \gamma_{h2} = 0.05, \gamma_h = 0.01, \mu_v = 0.07, \mu_h = 0.00003,\ \lambda_v = 1$. If θ is the free variable, then endemic equilibrium bifurcation appears at point θ_1 around $\theta = 6.6$, satisfying ${b_0}^2 - 4a_0c_0 = 0$ (as shown in Figure 5.2). One endemic equilibrium (bottom red line in Figure 5.2) $\theta_{hh} = (-b_0 - \sqrt{{b_0}^2 - 4a_0c_0})/2a_0$ merges with the DFE (blue line in Figure 5.2) at point θ_2 around $\theta = 11.8$, another endemic equilibrium (top red line in Figure 5.2) is $\theta_{hh} = (-b_0 + \sqrt{{b_0}^2 - 4a_0c_0})/2a_0$. Resulting in only one equilibrium (DFE) when $\theta < \theta_1$, two equilibrium points (one DFE and one endemic equilibrium) when $\theta = \theta_1$, three equilibrium points (one DFE and two endemic equilibriums) when $\theta_1 < \theta < \theta_2$, and two equilibrium points (one DFE and one endemic equilibrium) when $\theta \geq \theta_2$.

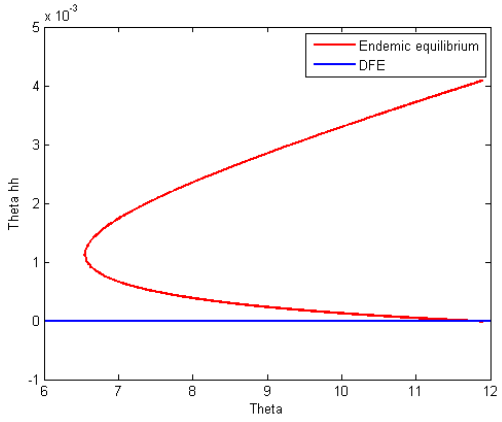


Figure 5. 2 Endemic equilibrium bifurcation

5.4. Control Strategies and Optimization

5.4.1 Control Strategies for Zika Virus and Corresponding Efficacies

Although the vaccination has proven to be one of the most useful strategies to restrain epidemic transmission, no confirmed vaccines of the Zika virus are in clinical use so far [200]. Several vaccination research teams have finished preclinical studies in animals and begun research on phase 1 trial studies in humans [220-222]. A team from the National Institute of Allergy and Infectious Diseases (NIAID) launched phase 2 clinical trials in March 2017, with an estimated completion date in 2019 [223].

Lack of effective Zika virus vaccines has led to the utilization of traditional disease control strategies, such as insecticide spraying and destruction of larval breeding [224, 225]. One general strategy, the release of insects with dominant (RIDL), involves the release of insects carrying dominant lethality genes to introduce a repressible lethal gene into the mosquito population [226]. Another novel control strategy uses endosymbiotic bacteria (Wolbachia) to prevent arboviruses replicating within the mosquitoes [227]. Both strategies can be used to control the population of infectious mosquitoes. In 2011, Foy discovered that the Zika virus is also sexually transmitted among human beings [228]; this evidence was confirmed in 2015 [195]. Therefore, the population of infect human can also be reduced by decreasing the frequency of sexual contact.

Based on previous research, this paper considers three generalized classes of controls: u_1 is the control (in percentage) that can improve the death rate of the mosquito population (e.g., spraying and RIDL [226]); u_2 is the control (in percentage) that can reduce the infectious rate of a susceptible mosquito (e.g., Wolbachia [227]); and u_3 is the control (in percentage) that can reduce the infection rate among the human population via sexual contacts, such as uses of

protection [228]. Therefore, system parameters $\mu_v, c,$ and β in Eqs. (5.1)–(5.9) are transformed into $\mu_v{}^*, c^*\ and\ \beta^*$, where $\mu_v{}^* = \mu_v(1+u_1), c^* = c(1-u_2), \beta^* = \beta(1-u_3)$.

The disease reproductive rate, R_0, is frequently used as a metric of control efficiencies because R_0 is an essential measurement of infectious disease transmissions vs. recoveries. Suggested by the existing literature, R_0 is positively correlated to several control strategies or interventions [229]. The higher reproductive rate reflects the underlying diaseas has higher ability to spread if uncontrolled . To observe the disease transmission ability under control strategies, R_0 in Eqs. (5.19) is updated to ${R_0}^*$ by adding control variables (i.e., $u_1, u_2,$ and u_3) as following.

$$ {R_0}^* = \frac{\theta\beta^*\kappa k_2 + \sqrt{(\theta\beta^*\kappa\, k_2)^2 + 4\theta a^2 b c^* \mu_h k_2 k_6 (\eta + v_h k_3)(1-(v_v/(v_v+\mu_v{}^*)))/\lambda_h}}{2} \tag{5.35} $$

With fixed system parameters in Eq. (5.35), ${R_0}^*$ is a function of the three control variables (i.e., $u_1, u_2,$ and u_3). Therefore, the variation of ${R_0}^*$ can be observed by adjusting each control to demonstrate the efficacies of each control strategy.

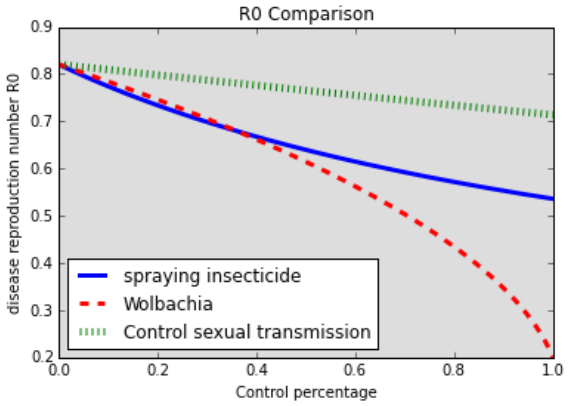


Figure 5. 3 The efficacies of each control strategy

Figure 5.3. shows that reducing the Zika infection rate of the susceptive mosquito (u_2) is the most effective control strategy because maximizing u_2 results in the lowest ${R_0}^*$. On the contrary, controlling human sexual contacts shown to have little impacts on Zika transmission

since $R_0{}^*$ is not sensitive to u_3. The efficacy of u_1 remains limited despite the reasonable control strategy of increasing the death rate of the mosquito population.

5.4.2 Optimal Control Problem to Limit Zika Transmission

Although controlling the infection rate of the susceptive mosquito (u_2) is known to be the most effective control strategy for Zika, disease control centers and health organizations cannot only execute the control strategy to maximize u_2 because an efficient control strategy should be able to mitigate the underlying reproductive rate of Zika and minimize corresponding costs at the same time. The costs of executing control u_2 frequently are prohibitively high when comparing to the subsequent benefits. Therefore, in this paper, an optimal control problem is proposed to determine a trade-off between the spread of a Zika epidemic while minimizing the cost of prevention.

The costs of controls for $u_1, u_2,$ and u_3 are defined as $c_1, c_2,$ and c_3, respectively. Also, the clinical costs to cure symptomatically infections I_{h1} and convalescent infections I_{h2} are defined as c_4 and c_5, respectively, the objective function of the optimal control problem can be written as:

$$\text{Minimize } J(u_1, u_2, u_3) = \int_{t_0}^{t_f} c_1 {u_1}^2(t) + c_2 {u_2}^2(t) + c_3 {u_3}^2(t) + c_4 I_{h1}(t) + c_5 I_{h2}(t) dt \quad (5.36)$$

Subject to the set of constraints, $g_i = 0$, i = 1, …, 9, defined by right-hand-side of Eqs. (5.1)–(5.9), but c^* *and* β^* replace c and β and $\mu_v{}^*$ replaces the death rate of the mosquito population, μ_v.

5.4.3 Pontryagin's maximum principle to solve the revised Optimal Control Problem

Pontryagin's maximum principle (PMP) is traditionally used to solve the optimal control problem mathematically [230], and then forward and backward iterations can be utilized to solve the

approximate numerical optimal control solution [41, 50]. Utilizing the PMP, the Hamilton–Jacobi–Bellman equation H of J (from Eq. (36)) is defined as:

$$H = c_1 {u_1}^2(t) + c_2 {u_2}^2(t) + c_3 {u_3}^2(t) + c_4 I_{h1}(t) + c_5 I_{h2}(t) + \sum_{i=1}^{9} \lambda_i \, g_i \tag{5.37}$$

where g_i is the right-hand-side of the system Eqs. (5.1)-(5.9) with controllable parameters $\mu_v{}^*, c^* \; and \; \beta^*$, where $\mu_v{}^* = \mu_v(1 + u_1), c^* = c(1 - u_2), \beta^* = \beta(1 - u_3)$; λ_i is an adjoint function corresponds to each g_i. Each λ_i satisfies the following constricted trajectory:

$$\frac{d\lambda_i}{dt} = -\frac{dH}{dx_i} \tag{5.38}$$

where x_i represents the ith system variable from Eqs. (5.1)-(5.9). For example, x_1 represents the susceptive human S_h. Also, at the terminal time t_f, λ_i satisfies the transversality condition:

$$\lambda_i \, (t_f) = 0 \tag{5.39}$$

Because of H is a quadratic function to the control variables u_i, when $c_i > 0$ the global optimization point exists at $\frac{dH}{du_i} = 0$. Therefore, the optimal functions for each control variables are shown as follow:

$${u_1}^*(t) = {\mu_v * (\lambda_7 * S_v + \lambda_8 * E_v + \lambda_9 * I_v)}/{2c_1} \tag{5.40}$$

$${u_2}^*(t) = {(\lambda_8 - \lambda_7) * a * c * N_h * S_h * (\eta * E_h + I_{h2})}\Big/{2c_2} \tag{5.41}$$

$${u_3}^*(t) = {\beta * N_h * S_h * (\kappa * E_h + I_{h1} + \tau I_{h2}) * (\lambda_2 * \theta + \lambda_5 * (1 - \theta) - \lambda_1)}\Big/{2c_3} \tag{5.42}$$

In Eqs. (5.40)–(5.42), $u_i^*(t)$ are the control variables of time $0 \le t \le t_f$. To obtain the stable results of the optimal control problem, the forward-and-backward iterations are required. The details of the forward-and-backward iteration algorithm are shown as follow:

Table 5. 2 Forward and backward iteration algorithm

Procedure code for solving epidemic optimal control problem using Pontryagin's maximum principle

Begin

Initialize: Set an initial control strategy for each time $t_0 \le t \le t_f$, usually set $u_i(t)$ as a constant c

While (termination conditions are not satisfied)

// forward iteration for calculating state variables

For each time period $t_0 \le t \le t_f$

Using the Runge-Kutta method to solve the ODE Eqs. (1)-(9) with parameters $\mu_v{}^*, c^* \; and \; \beta^*$, obtain the system variable values $(S_h(t), E_h(t), I_{h1}(t), I_{h1}(t), A_h(t), R_h(t), S_v(t), E_v(t), I_v(t))$ at each discrete time period

End

// backward iteration for calculating the adjoint function and control variables

For each time period $t_f \ge t \ge t_0$

Set the initial value of λ_i based on the transversality condition from Eq. (39)

Using the Runge-Kutta method to solve the trajectory constricts from Eq. (38), obtain the adjoint function $\lambda_i(t)$ at each discrete time period

Get revised control variables $u_i{}^*(t)$ by Eqs. (40)-(42)

End

//update control strategies

For each time period $t_0 \le t \le t_f$

Set the new control variables as $\left(1 - \frac{t}{t_f}\right) u_i(t) + \frac{t}{t_f} u_i{}^*(t)$

End

End

End

The termination condition can be defined as when the values of control variables are stable. Meaning, the values of control variables do not change much from the previous iteration. Alternatively, we can use the objective function value as a criterion. The algorithm stops when the objective function values change from the last iteration is smaller than a predefined tolerance.

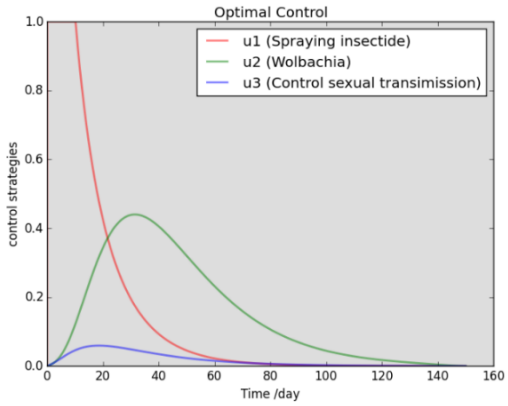


Figure 5. 4 Solved control strategies using Pontryagin's maximum principle

Figure 5.4 shows the values of control strategies which are solved by using PMP (the details of the simulation setting is described in section 4.6). The simulation results showed that control u_1 was at a high level in the first 30 days of the epidemic and then moved to the low level, while the control u_2 was at a high level between day 20 to 70 of the epidemic and control u_3 remained at a low level during the entire epidemic episode.

5.4.4 Memetic Algorithm to Solve the Optimal Control Problem

Even though the PMP provides a straight-forward approach to study the mitigation strategies for the Zika virus, this method is unable to obtain the convergent numerical control strategy when objective functions are highly nonlinear or non-convex [231-233]. In other words, the optimal control results obtained from the PMP must be the unique trajectory that satisfies the boundary and transversality conditions [234]. Therefore, a more robust global optimization algorithm may be a more suitable approach to solve the optimal control problem with multiple local minima [234]. To solve more complex optimal control problems that may have the non-convex objective functions, in this paper, we propose the first to use a revised Memetic Algorithm (MA) for the epidemic optimal control problem. As a companion to the Genetic Algorithm (GA), MA was inspired by Darwinian principles of natural evolution and Dawkins' notion of a meme

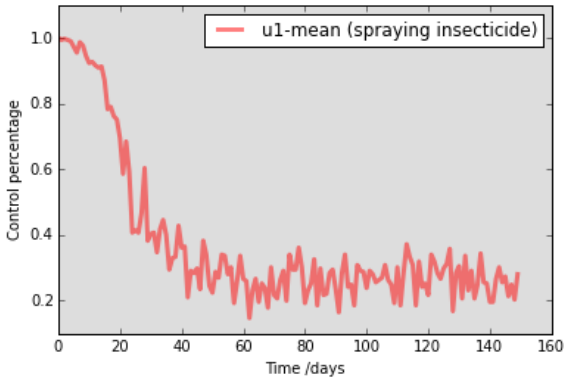
[235]. The primary difference between MA and GA is that every individual wants to improve themselves under the assumption of MA: parents transmit improved genes to their children. Therefore, MA can potentially reach optima faster than GA.

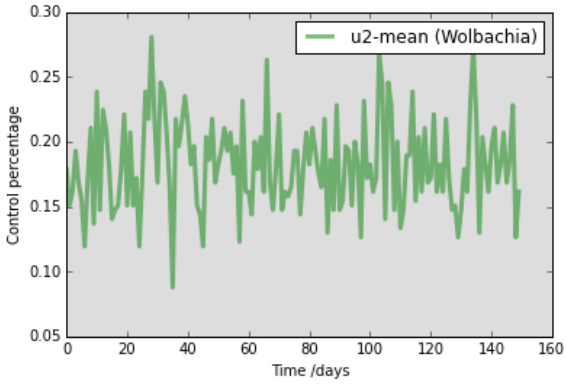
The given problem in Eq. (5.36) contains three control functions: ($u_1(t), u_2(t)$, and $u_3(t)$). The Runge-Kutta method was used as a numerical algorithm to solve the optimal control problem, where $3 * m$ control variables were used, and each control variable is transformed into n-bits binary variables to simplify calculation using the MA (i.e., considering 4-bits binary variables, binary variable '1010' represents $\frac{2^3+2^1}{2^4-1} \approx 66.7\%$ in the value of the control variable). Therefore, each candidate solution in the given optimal control problem has $3 * m * n$ binary variables.

Table 5. 3 Memetic Algorithm (MA)Pseudocode

Procedure code for solving epidemic optimal control problem using MA
Begin
Initialize: Generate an initial population of *k* control strategies, where each control strategy contains 3**m***n* random binary variables (i.e., a set of initial trial controls of u_1, u_2, and u_3)
While (termination conditions are not satisfied)
Evaluate all candidate control strategies based on the objective function in Eq. (36).
For all *k* candidate control strategies within the population
While (threshold of memetic local search number is not reached)
Random select the neighbors by variable neighborhood search algorithm [236] and evaluate them
If one of the neighbors is better than the current candidate strategy
Substitute the current candidate strategy by the best neighbor
End If
End While
End
Apply the standard GA algorithm to generate a new generation via selection, mutation, and crossover.
End While
End

The procedure of MA is presented in Table 5.3 in pseudocode format. MA can be simplified into two general steps: self-improvement of an individual within the generation and reproduce of the new generation. Since selection in both steps are based on evaluating the objective values of individuals, MA can approach the optimization control as what GA can do. Moreover, the self-improvement step can stochastically search the local neighborhood for better control, meaning the MA could potentially be more efficient than GA in searching for global optima. Therefore, MA could be an appropriate algorithm to solve the given epidemic optimal control problem.





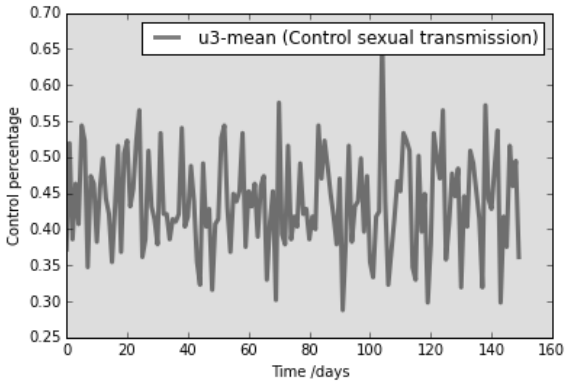


Figure 5. 5 Solved continuous control strategies using MA with 10000 iterations ($m = 3000, n = 4$)

Figure 5.5 shows the values of continuous control strategies, which are solved by using the MA (the detail of the simulation setting is described in section 5.4.6). There are 3000 segments in the entire epidemic period (150 days). Since the step length of the control strategies is 0.05 days, the solved control strategies can be considered as continuous. The simulation results show that control u_1 was at a high level in the first 20 days of the epidemic and then moved to the low level, while the control u_2 and control u_3 remained at the mid-level during the entire epidemic episode.

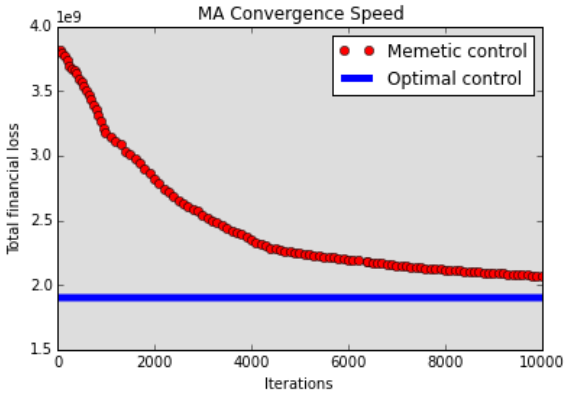


Figure 5. 6 MA-based optimal control convergence

With more iterations, the MA can reach a solution that is closer to the optimal level calculated by the PMP algorithm. To observe the convergence speed of MA, in Figure 5.6, the

theoretical optimal control is compared with the MA-based continuous control strategies with different steps. This comparison verifies that MA has the potential to approach the result of the real optimum.

5.4.5 Discrete control strategies to solve the revised optimal control problem

In real-world applications, regulating and adjusting the values of control strategies could be challenging to implement, since it may take time to have the changes taking effects or require substantial overhead costs during these changes. Therefore, merely considering the cost of control strategies based on the values of the control variables may not be practical. The regulating times must be included when considering the true implementation cost for a control strategy. Therefore, in this section, we revised the total cost during the entire epidemic in Eq. (5.36) to $J_{revised}$ as defined below:

$$J_{revised} = J + c_6 k_1 + c_7 k_2 + c_8 k_3, \tag{5.43}$$

where J is defined in Eq. (5.36), which represents the total cost of clinical and control. k_i are the numbers of times adjusted corresponding to controls u_i, where i = 1, 2, and 3, respectively. c_6, c_7 and c_8 are cost parameters involved for adjusting the controls u_i (i =1, 2 and 3).

By considering the revised total cost $J_{revised}$, the solution obtained by the PMP based algorithm may not be the optimal control anymore. Since the continuous control strategies could be cost-prohibitive due to the high control regulating costs. Moreover, the PMP based optimal control strategies may not be practical for implementation, since it may be impossible to implement and adjust different serval strategies in a short period of the time. For the same reasons, utilizing MA with a short time segment to design the continuous control strategies is not practical. To reduce the involved costs due to the control regulating, we used MA with a longer time segment to design the discrete control, which is executable in the real world with reasonable expenditure.

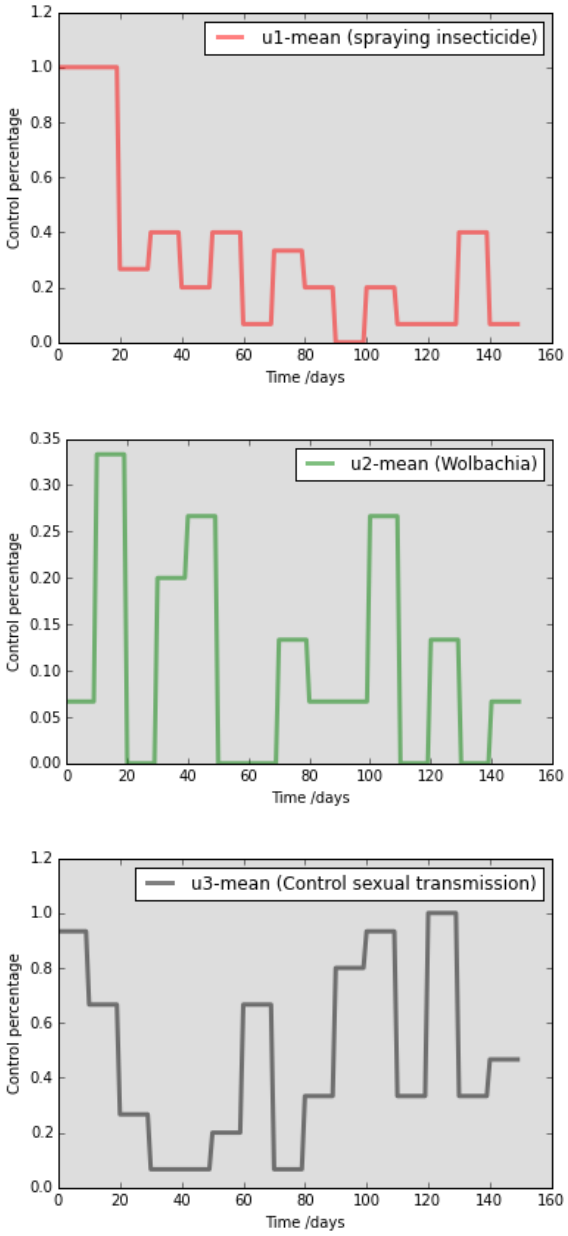
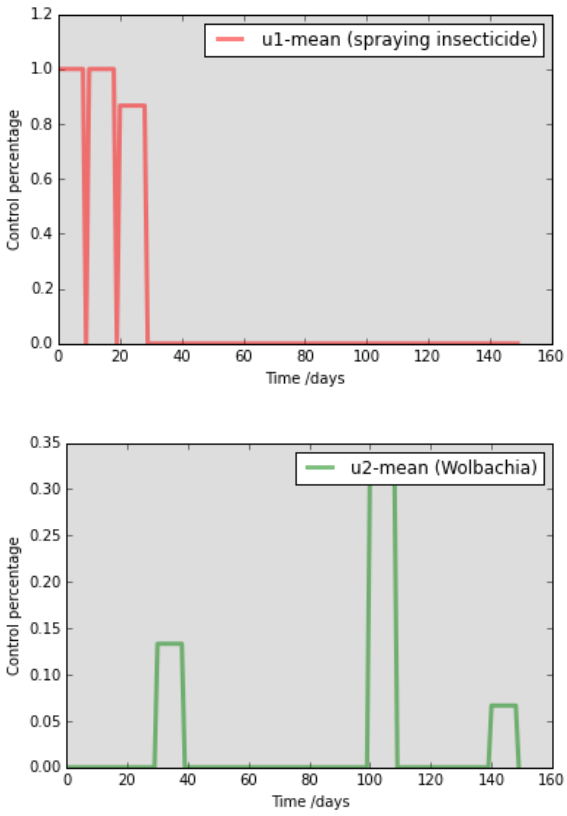


Figure 5. 7 Solved discrete control strategies using MA with 1000 iterations ($m = 15, n = 4$)

Figure 5.7 shows the values of discrete control strategies which are solved by using MA (the detail of the simulation setting is described in section 5.4.6). The control strategies can be considered as discrete since each single control variable was only adjusted every 10 days. The simulation results showed similar tendencies with the MA-based continuous control strategies (as

shown in Figure 5.5) for u_1 and u_2, while, the control u_3 fluctuates during the entire epidemic episode.

Moreover, under special circumstances, specific constraints may be imposed on the control strategies during the implementation. For example, instead of using an unlimited number of adjustments, each control strategy only allowed to be adjusted several times, due to limited resource constraints. In such cases, the constraints should be taken into considerations in our optimal control problem, and therefore, the PMP described in Section 5.4.3 cannot be applied to solve the resulting optimal control problem. Subsequently, the Memetic Algorithm is the only method to be able to handle problems with such constrains using limited adjustments.



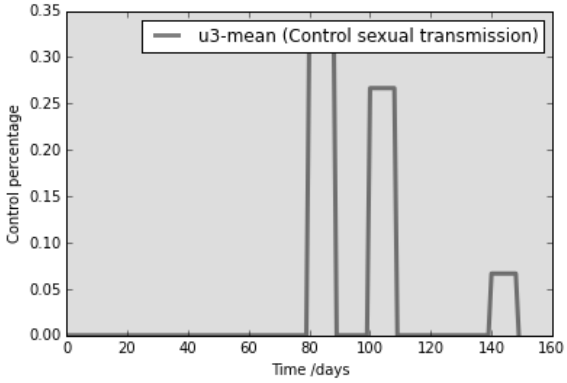


Figure 5. 8 Discrete control strategies with limited adjustments using MA with 1000 iterations ($\boldsymbol{m = 15, n = 4}$)

Figure 5.8 shows the values of discrete control strategies, which are solved by using MA under the constraints (the detail of the simulation settings is described in Section 5.4.6). The constraints require that each control strategy can only be implemented no more than three times. The simulation results show that it is crucial to implement u_1 with high level in the first 30 days of the epidemic, while the significances of u_2 and u_3 are less than u_1.

5.4.6 Simulation Comparisons

To verify the effectiveness of algorithms, in this section, we compared the performance of different control strategies by using an example of the northeastern region of Brazil as a case study. Brazil was the most severely afflicted by the Zika virus in the 2015–2016 outbreak, specifically the northeastern region of Brazil, including Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, and Bahia [237]. The population in the northeastern region of Brazil is about 53 million [211]. Assume the human population in this region was at disease-free conditions: $S_h(0) = 5.3 * 10^7$, $E_h(0) = I_{h1}(0) = I_{h2}(0) = R_h(0) = A_{h(0)} = 0$. Initially, we assume $S_v(0) = 10^8$, $E_v(0) = 10^5$, $I_v(0) = 10^4$. We use the suggested values from Table 1 for the system parameters. The total time of the simulation runs was set to be 150 days, use 0.05 day

the numerical discrete step length (the numerical simulation was implemented with standard Runge-Kutta fourth-order method).

The estimated costs for Zika treatment can be found in a 2017 literature; researchers projected that treatment of chronic Hepatitis C Virus (HCV) infection costs approximately $84,000 per person [238]. Therefore, this research assumed that $c_4 = \$1{,}000$ and $c_5 = \$8{,}000$ per day, and the coefficients of macro-control are $c_1 = \$5 * 10^6, c_2 = \$10^7$, and $c_3 = \$10^6$. Utilize the objective function defined by Eq. (5.36) for the entire epidemic period, simulation results are shown in Table 5.4.

Table 5. 4 Control Strategies comparison (Original objective function)

	Without Control	***PMP Control***	***MA Continuous***	***MA Discrete***	***MA Constraint***
Total infected	40,724	19,454	16,742	16,241	17,072
Obj. Func. Value	$\$3.820 * 10^9$	$\$1.901 * 10^9$	$\$2.065 * 10^9$	$\$1.982 * 10^9$	$\$1.929 * 10^9$
Solution Quality	200.94%	100%	108.62%	104.26%	101.47%

In Table 5.4, the PMP Control column summarizes the results from using Pontryagin's maximum principle algorithm; Results in MA Continuous column use 0.05 days as control step length in MA with a run-length of 20000 iterations; Results in MA Discrete column used 10 days as control step length in MA with run-length of 5000 iterations; MA Constraint column is the MA Discrete with an additional constraint, which requires the control mitigation strategies can only be implemented no more than 3 times. By comparing the performances, results from all 3 MA and PMP controls reduce about 50% of the total cost and about 50-60% of the total infectious than

without control. PMP control has the best performance in reducing the total cost during the epidemic. MA-based control strategies did better in limiting the number of total infections.

Table 5. 5 Control Strategies comparison (Revised objective function)

	Without Control	***PMP Control***	***MA Continuous***	***MA Discrete***	***MA Constraint***
Total infected	40,724	19,454	16,742	16,241	17,072
Obj. Func. Value	$\$3.820 * 10^9$	$\$1.901 * 10^9$	$\$2.065 * 10^9$	$\$1.982 * 10^9$	$\$1.929 * 10^9$
Implementation Cost	0	$\$1.424 * 10^9$	$\$1.92 * 10^9$	$\$9.6 * 10^6$	$\$3.2 * 10^6$
Revised Obj. Func. Value	$\$3.820 * 10^9$	$\$3.325 * 10^9$	$\$3.985 * 10^9$	$\$1.992 * 10^9$	$\$1.931 * 10^9$
Solution Quality	114.89%	100%	119.85%	59.91%	58.08%

To consider the cost of control implement during the epidemic, we compare the simulation results based on the original objective function with revised objective function (Eq. (5.43)) in Table 5.5. We assumed the values of implement cost parameters per time c_6, c_7 and c_8 as 2% of the control cost coefficients c_1, c_2 and c_3. Compare the results from two tables. Even though PMP control has a slight advantage over other threes by using the original objective function, MA discrete and MA constraint algorithms noticeably reduced the implementation costs. Therefore, we suggest MA discrete and MA constraint algorithms when the implementation cost is not negligible.

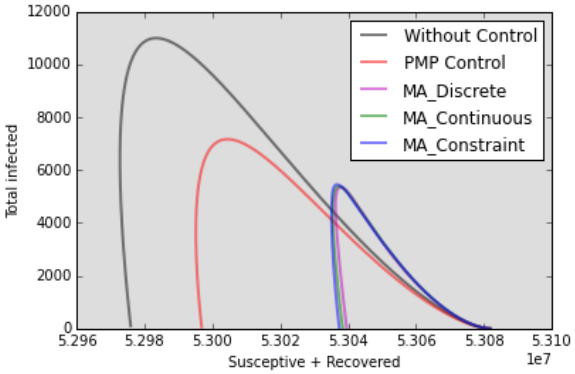


Figure 5. 9 Phases portraits comparisons using standard parameter setting (R_0^*=0.7189)

Although from our simulated results, the PMP and MA algorithms can effectively reduce the overall costs, including the intervention costs and the costs to deal with the infected population during the epidemic, the system tendencies using different control algorithms are similar under the given parameters setting. Figure 5.9 presents the phases portraits based on different algorithms; we find the total infected population went back to zero even without using the control strategies. Therefore, the system can reach the disease-free status no matter using which control strategies. However, the use of different control strategies could significantly reduce the infected population.

The above case studies and simulation comparisons demonstrate that with interventions via the optimal control mechanism, the number of the infected population can be reduced, and the overall cost can be minimized by balancing the intervention costs and costs associated with the infection. Next, we study how the optimal control mechanism would affect the disease dynamics where the system is close to its bifurcation point and a disease-free condition would not be achieved without controls. The main purpose is to see if the control strategies can bring the given Zika virus system back to a disease-free status at the end of the epidemic. We adopt the system (mentioned in Eqs. (5.1)-(5.9)). As we know, the change of stability can be observed easily around the bifurcation point. Therefore, we adjust system parameters a from 0.5 to 1, b from 0.4 to 0.8, c

from 0.5 to 0.8 and γ_{h1} from 0.2 to 0.05. Maintain the rest of parameters with the suggested value from Table 5.1. Let θ to be the free variable from 0.42 to 0.5, without using controls, the stability change can be observed in Fig 5.10.

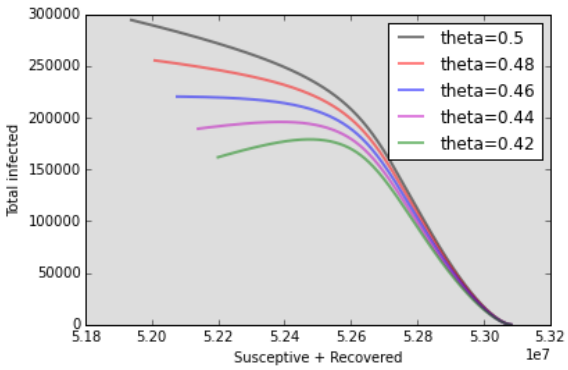


Figure 5. 10 Stability changes around the bifurcation point (without control)

From Fig 5.10, we know the given system cannot automatically go back to disease-free status when θ is around 0.5. In Fig 5.11, we utilized the PMP and MA control strategies into the system under the current parameters setting. Even though there is a slight difference between PMP and MA control strategies in the infected population results, all four control strategies can successfully limit the infected population and control the system back to disease-free status.

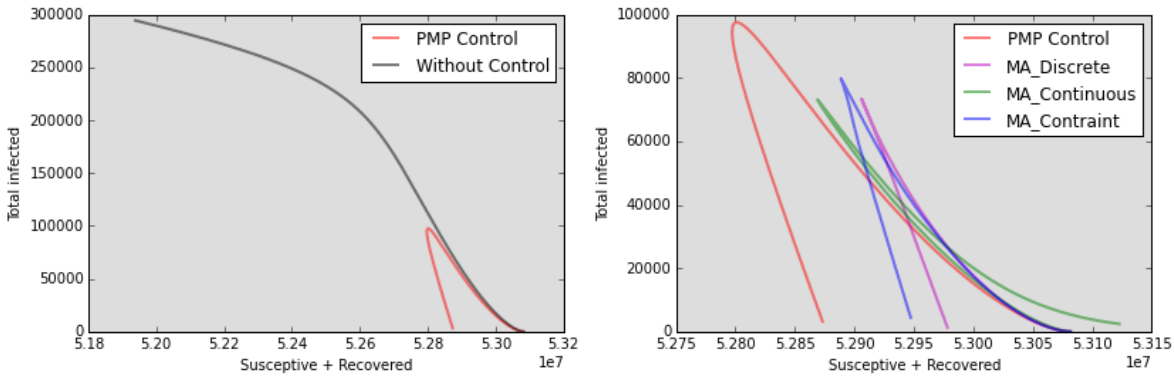


Figure 5. 11 Phases portraits comparisons using higher infecting ability setting (R_0^*=8.0761)

Table 5.6 summarizes and compares the simulation results using the current parameter settings. Since the current R_0 vlaue is higher than the previously assumed value; one can assume that the current Zika Virus epidemic is more contagious than standard Zika Virus. Under this assumption, simulation results show that the epidemic can infect more people and cause more financial losses than the standard Zika Virus epidemics simulated. In the meantime, PMP and MA control strategies can still limit the number of the total infectious and the objective function value.

Table 5. 6 Control Strategies comparison (Revised objective function)

	Without Control	PMP Control	MA Continuous	MA Discrete	MA Constraint
Total infected	750,159	154,796	99,555	101,565	125,247
Obj. Func. Value	$\$1.379 * 10^{11}$	$\$2.278 * 10^{10}$	$\$2.640 * 10^{10}$	$\$2.332 * 10^{10}$	$\$2.416 * 10^{10}$
Implementation Cost	0	$\$1.92 * 10^{9}$	$\$1.92 * 10^{9}$	$\$9.6 * 10^{6}$	$\$3.2 * 10^{6}$
Revised Obj. Func. Value	$\$1.379 * 10^{11}$	$\$2.471 * 10^{10}$	$\$2.832 * 10^{10}$	$\$2.333 * 10^{10}$	$\$2.416 * 10^{10}$
Solution Quality	558.07%	100%	114.60%	94.41%	97.77%

In short, the simulation results verified that PMP and MA control strategies effectively limited both the disease epidemic and financial losses. Under particular cases, these strategies are able to bring an unstable epidemic back to disease-free status. When we consider the implementation costs for the control strategies, both MA Discrete and MA Constraint algorithms are more efficient than the PMP and MA Continuous algorithms.

5.5. Discussion and Conclusion

This paper changes the existing dynamic system of the Zika virus [204] from an enclosed population system to an open population system and then study the system equilibrium, backward

bifurcation analysis, and the optimal control analysis. The results from these studies suggest some potential intervention strategies for mitigating Zika virus epidemics and offer a general analytical framework that can be applied to other infectious diseases. Moreover, the optimal control algorithms presented in Section 4.6 is an innovate approach to solve both continuous and discrete disease control problem. These algorithms have been verified to effectively control the stability of the epidemic system and reduce the total costs due to the disease.

New births and deaths processes for both human and mosquito populations are considered to make the ODE system as an open system, and then, the corresponding Disease-Free Equilibriums (DFE) and disease reproduction rate (R_0) being calculated for the open ODE system. Moreover, the endemic equilibrium is calculated at a dynamic balance condition of the epidemic. Finally, Bifurcation analysis of endemic equilibrium is carried out to study different kinds of equilibriums (Disease-free equilibrium and endemic equilibrium) at the various boundary conditions.

Three commonly adopted Zika virus control strategies are incorporated into the given ODE model to explore the effectiveness of these disease control and mitigation strategies. Using the disease reproduction number R_0 as the underlining metric, control efficacies are compared without considering corresponding costs.

Moreover, an optimal control problem is built to determine the trade-off between the spread of Zika and corresponding costs for these control strategies. In this paper, both conventional Pontryagin's maximum priniple and a new Memetic Algorithm are used to solve the given optimal control problem and their computational results are compared using relative merits. Pontryagin's maximum priniple method can be used to obtain a theoretical optimal solution for the continuous control problems with convex objective functions. Conversely, the specialized MA is proven to be

well-suited for a wide range of optimal control applications. The MA can solve both continuous and discrete control problems with convex or non-convex objective functions plus additional implementation or operation constraints.

In this paper, three popular disease control or intervention strategies for the Zika epidemics are explored, i.e., the control (in percentage) that can improve the death rate of the mosquito population (e.g., spraying and RIDL [226]) (control u_1); the control (in percentage) that can reduce the infectious rate of a susceptible mosquito (e.g., Wolbachia [227]) (control u_2); and the control (in percentage) that can reduce the infection rate among the human population via sexual contacts (control u_3). Our simulated results based on the optimal control problem suggest that the timings and the intensities of implementing for these interventions are very different.

First of all, the disease epidemic mitigation and many social policy intervention systems unlike many real-time control systems in the mechanical and electrical engineering domain, the real-time feedback-control protocol not only is impractical but also cost-prohibitive or impossible. Instead, stage-wise step controls or impulsive control strategies are more practical and cost-effective for disease mitigation systems. Optimal control problems to determine how often and how long the interventions should be involved are the main goals.

Secondly, from our simulated computational studies, the vector population control (u_1) would be most effective when applied in the initial stages of the epidemic, and as intensive as the available resources allowed. The vector infection rate controls (u_2) may not be effective until in the later stages of the Zika epidemic. The controls of human-to-human contacts (u_3), via education or media coverages, are less effective overall and would make more sense to implement them after the epidemic reached its peak.

System parameters such as $\theta, \beta, \kappa, \tau, v_h, \gamma_{h1}, \text{and } \gamma_{h2}$ are key parameters for the disease reproduction number R_0. Among which, $\theta, \beta, \kappa, \text{and } \tau$ are positively correlated to R_0; while $v_h, \gamma_{h1}, \text{and } \gamma_{h2}$ are negatively correlated to R_0. To mitigate the transmission of Zika virus, the system parameters $\theta, \beta, \kappa, \text{and } \tau$ should be decreased and $v_h, \gamma_{h1}, \text{and } \gamma_{h2}$ should be increased as much as possible.

Finally, as our computational experiments suggested, less than a handful of interventions for each control strategy are sufficient to mitigate the Zika epidemics, and the implementation periods for each intervention could be as short as one week or so.

Chapter 6 - A New Evidence Based Impulse Control with Event-triggered Conditions for the Epidemic Dynamic System

Chapter 6 is based on the manuscript “A New Evidence Based Impulse Control with Event-triggered Conditions for the Epidemic Dynamic System” Submitted to Journal of Franklin Institute.

Abstract

This paper is concerned with the epidemic control problem of the infectious disease epidemic system. By using the neural networks to analyze the historical epidemic data, evidence-based impulse control (EBIC) and associated event-triggered controllers are designed to impulsively control the future epidemic. The event-triggered controller is designed to determine the best timing to implement the control. Then the EBIC mechanism can find the optimal level for the control. In the design of EBIC, different types of neural networks like CNN, RNN, and fully connected neural networks are trained. These neural networks are used to learn the relation between the prevalence data and the historical control strategies. This paper also shows the epidemic dynamic system is stable under event-triggered control with and without periodicity. Finally, simulation comparisons are presented to reflect the new method is more intelligent and effective than the existed impulse control methods. Other simulations prove the validity, optimality, and robustness of the EBIC method.

Keywords: Epidemic system; Impulse control; Evidence data; Event-triggered

6.1 Introduction

In the system theory, Impulse control is defined as the intervention by instantaneously moving the process to some new point in the state space at particular discrete time moments [239]. Since Simeonov and Bainov first introduced the system with impulsive inputs [240], the

concept of the impulse control system has been widely applied to study application areas like healthcare, communication, robot design and finance [241-244]. In the epidemic dynamic system, impulse control was used to simulate the disease intervention strategies like vaccination, hospitalization, and isolation. The impulse control epidemic systems were used to study many infectious diseases like HIV, tuberculosis and Ebola virus [245, 246].

However, it is hard to decide the appropriate time to trigger impulse control for the epidemic system. Since the impulse control of infectious disease should be triggered based on the severity of the current epidemic. Therefore, the epidemic system requires a specific event-trigger condition (ETC) to trigger impulse control. Several existed literature set the predefined period to trigger impulse control [247-249]. This literature set the phase-shifting period as the ETC. Other literature adopted the feedback control corresponded into a single variable as the ETC [245, 246, 250, 251]. The ETC will be satisfied when the value of the specific variable reaches the threshold.

Although the previous literatures accomplished to trigger the impulse control by using ETC, these methods still existed obvious deficiencies. In the predefined period ETC, people have to estimate the intervention executed time before the epidemic outbreak. To ensure the effectiveness of the intervention, the estimated time has to reflect the severity of the epidemic. Meaning, the predefined period ETC required excessive precision on the prediction for the epidemic in the future. In the feedback control, the ETC has to judge the tendency of the epidemic only based on one variable. This method may be effective for the system with only a few variables. However, for the system with a large number of variables, this method cannot assess the epidemic severity from all angles.

To overcome the mentioned deficiencies of existed methods, in this paper, we introduce an innovative evidence-based impulse control (EBIC) method. This method is inspired by the recent work by Chen et.al [44]. The mentioned work proposed the concept of evidence-based control, which used the evidence database to generalize the relation between system variables and control outputs. The main features and contributions of this paper are summarized as follows:

(1) Unlike the above-mentioned work [44], this paper not only focuses on the prediction of the magnitude of the control, but also determines the best timing to implement the control by designing two types of ETCs. These ETCs can automatically and intelligently trigger the impulse control based on the current epidemic data.

(2) Compare to the previous impulse control methods [247-249, 251, 252], the EBIC method is less empirical and subjective. EBIC mechanism takes advantage of the historical data, but not directly applies these data and control into the future epidemic. The historical evidence data are used to train the neural networks. With future epidemic data, the trained neural network can predict the specific control strategy to prevent disease transmission.

(3) This paper studied the characteristics of different types of neural networks for the EBIC method. Among which, the fully connected neural network has a fast running time, the convolutional neural network can better analyze the multi-dimension data, the recurrent neural network is more effective in data with time series.

(4) The stability of the epidemic system is discussed under the impulse control triggered by ETCs. We prove the epidemic system is stable under the periodic impulse control. If the impulse control is not stable, the system is stable when the control is triggered by the given ETC.

Therefore, the rest of this paper is organized as follows: Section II defines the problem formulation and preliminaries; Section III details the designs and settings of the ETC and EBIC; Section IV demonstrates the stability and periodicity of the system controlled by EBIC; Section V illustrates the effectiveness of the EBIC method with simulations; Section VI concludes the paper and suggests future work.

6.2 Problem Formulation and Preliminaries

Consider a class of nonlinear continuous-time epidemic system given by:

$$\dot{x}_i = f_i(x) + g_i(x)u \tag{6.1}$$

where $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$, $i = 1,2, \dots, n$ denotes the system state vector, $u \in \mathbb{R}_{[0,\overline{u}]}$ denotes the disease impulse control input, $\overline{u}$ represents the adaptive upper bound of u. $f_i(x)$ and $g_i(x)$, $i = 1,2, \dots, n$ represent the smooth nonlinear epidemic system functions.

$$\Omega = \{x \in \mathbb{R}_+^n | x_1, x_2, \dots, x_n \geq 0\} \tag{6.2}$$

Consider u as an impulsive control strategy to mitigate the spreading of disease. Suppose these impulses occur at time nodes $\tau_1, \dots, \tau_m \in \Gamma$, where m is the number of impulse control activations, Γ is the time nodes set of impulse controls. On the time interval (τ_j, τ_{j+1}), system (6.1) with $u = 0$ describes the process of the epidemic transmission, where $j = 1,2, \dots, m-1$. At the sequences of a time node τ_j, the states of the impulse control system are given by:

$$\dot{x}_i(\tau_j^-) = f_i\big(x(\tau_j)\big) \tag{6.3}$$

$$\dot{x}_i(\tau_j^+) = f_i\big(x(\tau_j)\big) + g_i\left(x\big(\tau_j\big)\right)u(\tau_j)$$

The objective function is required to design the optimal impulse control strategy. The objective function of the epidemic system (1) is combined with the aggregated costs due to the infectious with the costs due to impulse controls on the time interval $[0, T]$. Let x_I represents the

infected population, where $x_I \in x$. Specially, we have $g_I\left(x(\tau_j)\right)u(\tau_j) \leq 0$ represents the reduction of the infected population due to the control. The aggregated costs due to the infectious is defined as: at any given time $t \in [0, T]$, the overall cost with the rate $C_1(x_I(t))$. The disease control cost is defined as the cumulated cost of $C_2(u(\tau_j)), j = 1,2, \dots, m$, this cost represents the consumption of vaccination, antivirus applications, and human resources.

$$J = \int_0^T C_1(x_I(t))dt + \sum_{j=1}^{m} C_2(u(\tau_j)) \tag{6.4}$$

To mitigate the spreading of the epidemic with a limited budget, the optimal impulse control strategy should be expected to minimize J under the constraints defined by equations in the system (6.1)

$$\text{Minimize } J(u(\tau_j)) \, , j = 1,2, \dots, m \tag{6.5}$$

Many literatures already designed the impulse control strategies for the epidemic system (1) [245, 246, 251-256]. However, there are two major weaknesses existed in these methods for solving the given optimal impulse control problem: 1) Some methods adopted the feedback control, which can only limit the disease transmission without considering the cost of control consumption [245, 246, 251, 252, 254]. Meaning, they only considered aggregated infected costs in the objective function. 2) The rest of the methods required pre-defined time $\tau_1, \dots, \tau_m$ or period for implementing the control law [253, 255, 256]. Meaning, the accurate prejudgment for the control implement time is necessary to calculate the value of impulse control u.

To improve the existed impulse control methods, our objective is to design an evidence-based impulse control (EBIC) with two different event-triggered conditions (ETC) to activate the control. The ETC can determine the time epoch to trigger the impulse control. If the ETC triggers the control, the system will handover to EBIC. Then the EBIC mechanism determines

the optimal level of the control. To achieve the above-mentioned goal, we make the following assumptions for the given problem.

Assumption 1: Impulses occur at the time $\tau_1, \dots, \tau_m \in \Gamma, \Gamma \neq \emptyset$. For this problem, the event-triggered condition only depends on the epidemic situation, which are the values of system state variables x_i.

Assumption 2: $\exists\, x^* \in \Omega,\ x^* = [x_1^*, \dots, x_I^* \dots x_n^*]^T, x_I^* \neq 0$. That point satisfies $f_i(x^*) = 0,\ i = 1,2, \dots, n$. Specifically, the real symmetric matrix $\hat{A}(x^*)$ is strictly nonnegative defined, where $\hat{A}(x) = A(x) + A^T(x)$, $A(x)$ is the Jacobi matrix of the system (1) without control

$$A(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{6.6}$$

6.3 Event-triggered control condition and Evidence based impulse control

In this chapter, an event-triggered control scheme using only the system state vector is provided for the nonlinear epidemic system (6.1). Considering the high cost of the control implementation, the ETC is expected to trigger the impulse control only at the necessary time epochs. Therefore, two system variable based ETCs are defined to reduce the negative influence due to the infections with limited control triggers. These ETCs are designed to trigger the EBIC by considering the automaticity, holistic and intellectuality. Meanwhile, this chapter discusses how to calculate the value of EBIC after the ETC is triggered. When calculating the control level of the EBIC, we don't directly use the historical control data for the current epidemic. Because of the empirical control strategy may not be optimal, even it was optimal to control the historical prevalence. In this section, the EBIC calculation will utilize the neural networks to learn from the evidence prevalence and control data. The observed historical database is used to train neural

network, which is expected to determine the accurate value of EBIC based on the current epidemic information from state variables.

6.3.1 Event-triggered and Control Mechanism

In this section, the event-triggered mechanism is described to trigger the EBIC, which is used as the disease control input to mitigate the epidemic described on the system (1). Since the control is triggered with limited times as we mentioned, the event-triggered condition is determined based on the epidemic severity. For the epidemic system, the epidemic severity is reflected by the values of system state variables. Therefore, the ETCs in this paper are defined by using the current system state variables. The details of two ETCs used in this paper are mentioned in section 6.3.2 and 6.3.3.

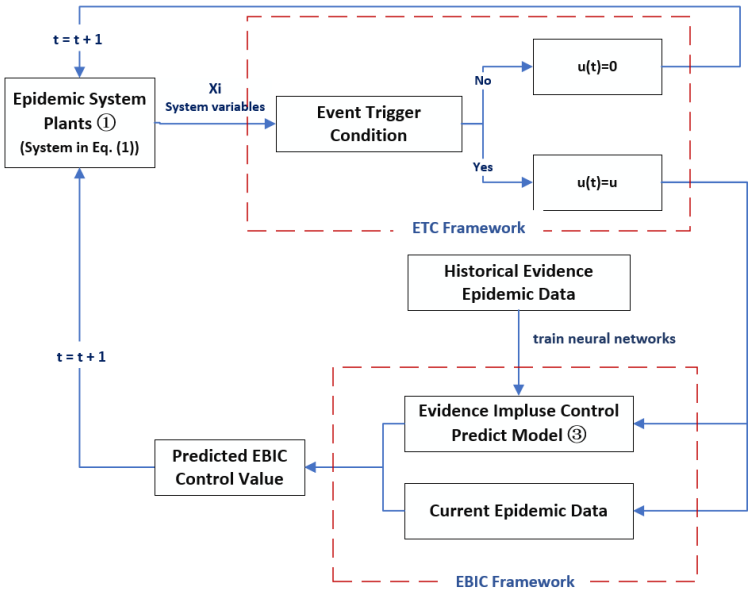


Figure 6. 1 Flowchart of the Event-triggered and Control Mechanism

From the previous literature for an epidemic impulse system, most of the existed feedback triggers ETCs were defined by the responses from one criterion variable [250, 253]. For example, considering the infected population x_I as the criteria variable, the ETC can be defined as:

$$\begin{cases} Trigger\ (\overline{u(t)} = \bar{u}), & if\ x_I(\tau) - x_I(\tau - \eta) \geq \kappa\ and\ x_I(\tau) \geq x_I(\tau - 1) \\ Not\ Trigger\ (\overline{u(t)} = 0), & else \end{cases} \tag{6.7}$$

where τ is current time node, η is epidemic previewing time length and the potential starting point for ETC, κ is the adaptive threshold coefficient.

If the ETC in Eq. (6.7) is satisfied, the EBIC will be triggered to control the epidemic system. Unlike most of the existed methods in event-triggered control [245, 246, 254], this section triggers the impulse control by adjusting the current upper bound of the control variable $\overline{u(t)}$. If the ETC is not satisfied, $\overline{u(t)}$ will be set as zero, meaning the control cannot be triggered at the current time. Otherwise, if the ETC is satisfied, the $\overline{u(t)}$ will be set as the predefined $\bar{u}$. The value of EBIC is received from the learning model, which is trained by the historical evidence database. The details of the model training process are mentioned in section 6.3.4.

6.3.2 Convolutional Time Series Event Trigger Condition (CTSETC)

Although ETC mentioned in Eq. (6.7) is effective for some simple epidemic systems, the same ETC may not be able to determine the triggers for the system with multiple variables. To synthetically reflect the severity of the epidemic, the ETC should take advantage of the information from all variables instead of just a single variable. However, processing the data from multiple variables into a single 0-1 output trigger signal is challenging. In addition, to determine the trend of the epidemic, we should analyze the time series of the system variables. Therefore, in the following, the CTSETC is introduced in two steps. In the first step, the system variables matrix is transferred into a time series vector by using the convolutional kernel. In the second step, the received vector is processed by the time series analysis methods to evaluate the current epidemic severity and then determine if or not to trigger the control.

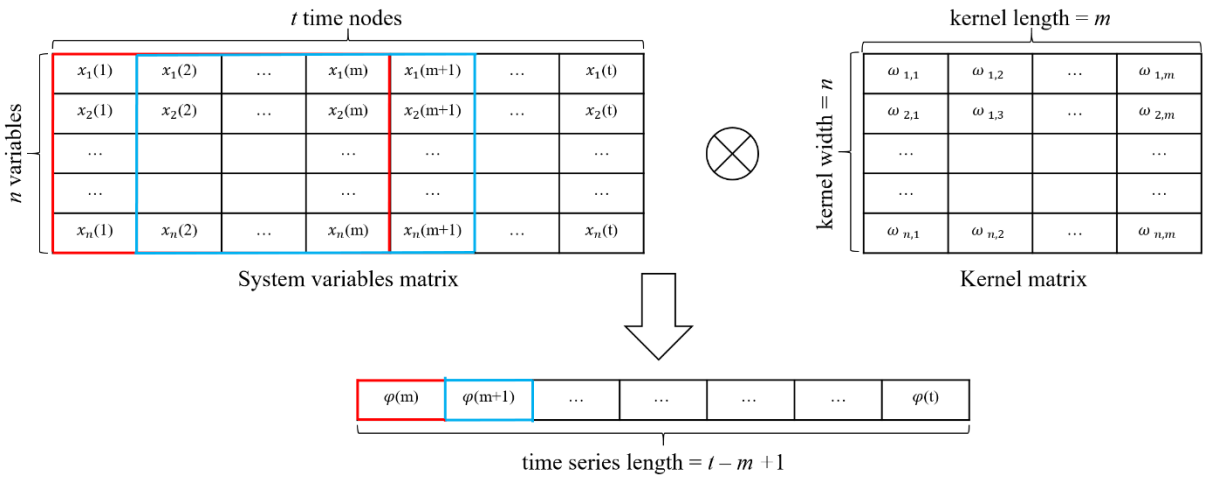


Figure 6. 2 Using Kernel matrix to obtain the time series

At the time epoch t, the historical system variables received from the impulse control system can be compacted into a $n * t$ matrix, where n is the number of the variables. To avoid the difficulty of comparing multiple time series, CTSETC simplifies the system variables matrix into one dimension time series by using a $n * m$ convolutional kernel matrix, where m is the length of the kernel. In the kernel matrix, each parameter $\omega_{i,k}$ represents the weight of the corresponding system variable, where $i = 1,2, \dots, n; k = 1,2, \dots, m$.

$$\varphi(t) = \sum_{i=1}^{n} \sum_{k=1}^{m} \omega_{i,k}\, x_i(t - m + k) \tag{6.8}$$

The time series φ is calculated by the Eq. (8). The higher value of $\varphi(t)$ represents an epidemic with a severe outbreak. To determine the trend of the current epidemic, time series methods are used to analyze φ. In statistics, moving average, control chart, and regression analysis are typical trend analysis methods [257]. In this section, we illustrate the weighted moving average (WMA) as an example.

$$WMA_p = \frac{\sum_{j=1}^{p} j\varphi(t-p+j)}{p*(p+1)/2} \tag{6.9}$$

Where WMA_p represents the moving average for last p nodes. For given $s, l \in \mathbb{Z}^+$, where $l > s$. Set the specific threshold percentage $\delta \in \mathbb{R}_{[0,1]}$. If there exists:

$$WMA_s - WMA_l > \delta WMA_l \tag{6.10}$$

Then the short term WMA is significant higher than the long term WMA, meaning the time series φ has the growth trend. Since φ can reflect the severity of the epidemic, the ETC has to trigger the control when Eq. (6.10) is satisfied.

The CTSETC is able to automatically trigger the control without using predefined control to implement time. In addition, the CTSETC can take advantage of the information from all system variables not only in the current states but also in the past states. However, human intervention is still existed in defining the weights of the kernel matrix and the threshold parameter. Therefore, the CTSETC cannot be considered as intelligent ETC.

6.3.3 Convolutional Evidence-based Event Trigger Condition (CEBETC)

In this section, another alternative ETC is presented by using the historical database to intelligently trigger the impulse control. In the CEBETC, the historical database is required to store the information for both system variables and the corresponding trigger condition. To avoid contingency, this information may be obtained from the historical epidemic in different years (scenarios). Since the peaks of epidemic outbreaks and the corresponding control triggers existed in different months, it is hard to only use the time characters to design the control trigger condition.

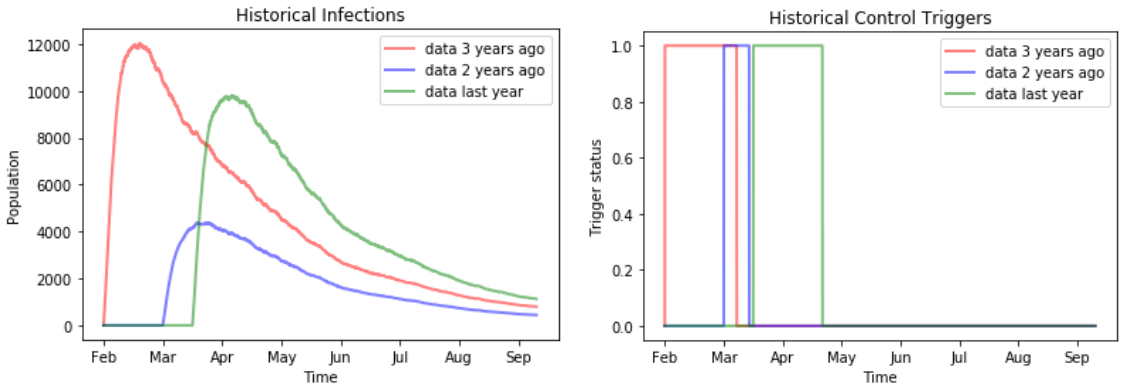


Figure 6. 3 Simulated historical infections and control triggers data example illustration

In addition, the control trigger status may be different even within the same number of infected populations. In Table.6.1, we compared historical control trigger status in different years. Sometimes, the control is triggered by the lower infected population. However, for another moment, the control is not triggered even with the higher population. Therefore, the ETC cannot be designed only based on the value of a single system variable.

Table 6. 1 The comparison of historical trigger situation illustration

Year	Data	Infected Population	Control Trigger Situation
Last year	March 21st	5998	Yes
	May 13th	6103	No
2 years ago	March 12th	3850	Yes
	April 9th	3990	No
3 years ago	February 8th	8017	Yes
	March 15th	8146	No

To effectively trigger the control by properly using historical data, we design the CEBETC as a potential ETC in this section. The evidence database included historical system vectors as input data and historical control trigger records as output data, are utilized to train the convolutional neural network (CNN). We consider CNN since it performances well in the feature learning from multi-dimensions data. For the CEBETC, CNN can better learn the valid features in the data with different infection magnitudes and time epochs. Therefore, the trained CNN can be considered as the ETC to determine the control trigger.

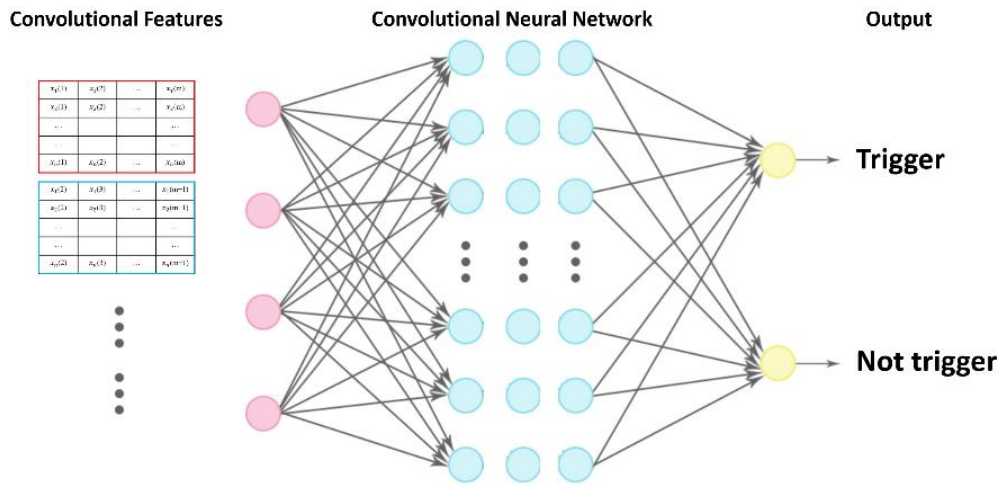


Figure 6. 4 The training process of the Convolutional Evidence-based Event Trigger Condition (CEBETC)

To illustrate the process of the CEBETC, we define the evidence database by

$$S_{CEBETC} = \{\mathbb{x}_{CEBETC}{}^{(i)}, \mathbb{y}_{CEBETC}{}^{(i)}\}_{i=1}^{N} \tag{6.11}$$

where $\mathbb{x}_{CEBETC}$ represents the convolutional features set, $\mathbb{y}_{CEBETC}$ represents the historical control trigger records, N is the size of the database. For each data, $\mathbb{x}_{CEBETC}{}^{(i)}$ is a $n * m$ matrix, which includes the features of n state variables in last m time nodes. $\mathbb{y}_{CEBETC}{}^{(i)} \in \{0,1\}$, where $\mathbb{y}_{CEBETC}{}^{(i)} = 1$ represents the control was triggered, $\mathbb{y}_{CEBETC}{}^{(i)} = 0$ represents the control was not triggered.

S_{CEBETC} can be used to train CNN model w, which can predict the output $\hat{y}(t)$ based on the current system state vector $x(t)$, the prediction of the output can be described as:

$$\hat{y}(x(t)) = w(x(t)) = w_1(w_2(w_3(x(t)))), t \in [0, T] \tag{6.12}$$

where w_1, w_2, w_3 denote the embedding, inference, and reconstruction subnet functions. These functions can be designed as convolutional, pooling and fully connected layers based on the different types of features in S_{CEBETC} [258]. In this paper, we designed the CEBETC neural network by starting with one convolutional layer followed by one flatten layer and two dense layers. The first three layers used a rectified linear activation function, the last one layer

considered the ‘softmax’ activation function. We adopt ‘adam’ optimizer and ‘cross-entropy’ loss function when training the neural network. During the simulation, the default training process takes 50 epochs.

6.3.4. Evidence Based Impulse Control (EBIC) design

Based on the ETCs from section 6.3.2 and 6.3.3, the system already determines the time epochs to trigger the EBIC. In this section, evidence learning models are designed to determine the optimal level of the EBIC value. By analyzing the historic evidence database, the evidence learning model can be used to learn the rules to optimize the EBIC control level. In the real world, epidemic control strategies are designed by studying the epidemic history. Meaning, the public health scientist will apply the historical validated control strategy into the current epidemic. This method is empirical since the current disease may be different from the historical disease in the transmission routes, transmission capacity, and the prevention method. Therefore, the historical control strategy may not work for the current epidemic.

In this section, the evidence learning model will be expected to predict the stronger control level when the ETC is satisfied. To guarantee the prediction accuracy, the historical evidence database should include enough proven effective control data. The historical evidence database for EBIC prediction is defined by:

$$S_{EBIC} = \{\mathbb{x}_{EBIC}{}^{(i)}, \mathbb{y}_{EBIC}{}^{(i)}\}_{i=1}^{N} \tag{6.13}$$

where $\mathbb{x}_{EBIC}$ represents the system variables set, $\mathbb{y}_{EBIC}$ represents the set of EBIC control levels, N is the size of the database. To analyze the database S_{EBIC}, the CNN model mentioned in section III.B still can be used. The major difference is the output set. In the CEBETC model, $\mathbb{y}_{CEBETC}$ only has two possible output values, which represent trigger or not trigger. In the EBIC

prediction model, $\mathbb{y}_{EBIC}$ may have multiple sets for output values, which represent different EBIC control levels.

Recurrent neural network (RNN) is another type of evidence learning model can be used to analyze the S_{EBIC}. RNN is a class of neural networks that allow previous outputs to be used as inputs while having hidden states [259]. RNN can not only learn from the current input variables but also "remember" things during the training. We consider the RNN in this section since the RNN has been proved to effectively analysis the time series. Since the outputs of the database S_{EBIC} represent the EBIC control level, which has continuity in the time series. Meaning, RNN may be able to fast and accurately classify the EBIC control levels by learning the continuity of predicted outputs. Similarly, modified RNNs like Echo State Networks (ESNs), Gated Recurrent Unit (GRU) and Long short-term memory (LSTM) can also be applied to analyze database S_{EBIC}.

However, due to the more complex architecture and connections, RNN has a slower training process than the general artificial neural network, even with the same input dimension, output dimension, and batch size. In addition, because of the inherently sequential behavior of the long sequences, some problems like slow convergence, degradation, and memory leakage may exist in RNN training. To improve the training performance, we consider a revised bidirectional unrolled RNN (BURNN) below:

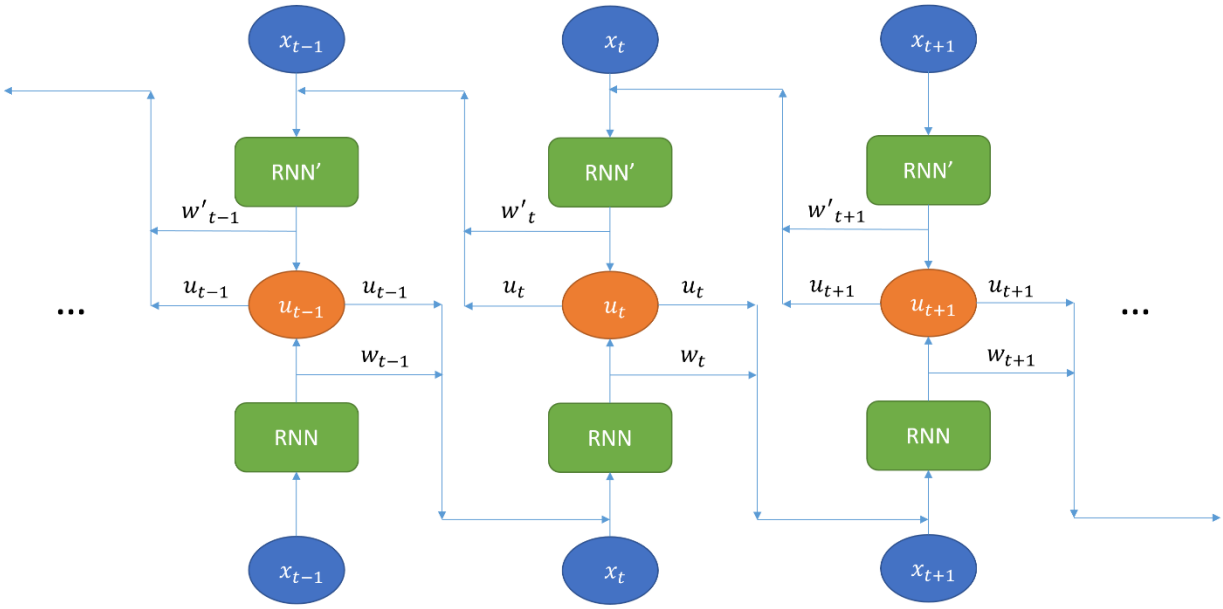


Figure 6. 5 Diagram of the bidirectional unrolled RNN

To overcome the limitation of traditional RNN in the slow convergence, BURNN adopts the bidirectional method to design the RNN architecture. Bidirectional RNN is trained forward and backward by two separate recurrent nets, both are connected into the same output layer. Therefore, the bidirectional RNN takes advantage of all available input information in the past and future [260]. Because of the special architecture, bidirectional RNN is more effective in prediction accuracy with less training iterations than unidirectional RNN [261]. Moreover, training a bidirectional RNN required more parameters, layers, and connections, which typically takes longer training time than unidirectional RNN. Therefore, BURNN adopted the unrolling process, which computes the output using the current RNN cell by chaining with the last trained RNN cell. The unrolling design can process a large amount of data with only a few parameters. Meaning, the BURNN can be trained with a significantly shorter time than bidirectional RNN.

Same to the design of the CEBETC, the training process of the neural networks in EBIC considered the 'adam' optimizer and 'cross-entropy' loss function. During the simulation, the training process takes 50 epochs.

6.4 Stability and Periodic Analysis

In this chapter, the stability of both non-control and impulse control epidemic systems are analyzed. Specifically, we analyzed the periodicity for the given system under the event-triggered impulse control. By comparing the steady states of the non-control and impulse control systems, shows that impulse control is effective in reducing the number of infected populations when reaching the steady-state of the epidemic system.

6.4.1 Stability Analysis for non-control Epidemic System

In this section, we consider $u = 0$ for the system (1), represents the non-control epidemic system. And discuss the equilibrium and stability of non-control epidemic system given by

$$\dot{x}_i = f_i(x) \tag{6.14}$$

Definition 1 [262]: $x^e \in \Omega,\ x^* = [x_1^e, \dots, x_I^e \dots x_n^e]^T$ is considered as an equilibrium of non-control system, if and only if $f_i(x^e) = 0$, for all $i = 1,2, \dots, n$. In addition, if $x_I^e = 0,\ x^e$ is called disease free equilibrium; otherwise, x^e is called endemic equilibrium.

Theorem 1 (Krasovskii theorem) [263]: For any $x' \in \Omega$, suppose $f_i(x') = 0$, $i = 1,2, \dots, n$. Let $A(x)$ be the Jacobi matrix. If $\hat{A}(x) = A(x) + A^T(x)$ is negative definite in a neighborhood of x', then x' is asymptotically stable and $V(x) = f^T(x)f(x)$ is a Lyapunov function, where $f(x) = [f_1(x), f_2(x) \dots f_n(x)]^T$.

Lemma 1: x^* is an endemic equilibrium. There exists a closed set $\Omega^* \subseteq \Omega$, the non-control system is asymptotically stable in a neighborhood Ω^*.

Proof: Based on assumption 2, $f_i(x^*) = 0$ and $x_I^* \neq 0$. x^* satisfies the definition of endemic equilibrium in Definition 1.

In addition, $f_i(x), i = 1,2, \dots, n$ are smooth nonlinear functions. Therefore $\frac{\partial f_i}{\partial x_j}$ are continuous, $i, j = 1,2, \dots, n$.

According to $\hat{A}(x^*)$ is strictly nonnegative defined. There exists $\delta \in \mathbb{R}_+^n$ that $\hat{A}(x)$ is strictly nonnegative defined, for every $x \in [x^* - \delta, x^* + \delta]$. Therefore, $[x^* - \delta, x^* + \delta]$ can be defined as a neighborhood Ω^*. Use the Theorem 1, the non-control system is asymptotically stable in a neighborhood Ω^*.

Based on the definition of Lyapunov stability [263]. For every $\varepsilon > 0$, there exists $t^* > 0$, $\varepsilon, t^* \in \mathbb{R}$. Satisfies that

$$\|\Phi(t; x_0, t_0) - x^*\| \leq \varepsilon, \forall t \geq t_0 + t^* \tag{6.15}$$

where $x_0 \in \Omega^*$ represents any given initial state of non-control system in Eq. (6.14), t_0 is the initial time, $\Phi(t; x_0, t_0) \in \Omega$ represents the system trajectory start from x_0 at t_0.

6.4.2 Stability Analysis for Event-triggered Impulse Control System with periodicity

This section discusses the stability of the event triggered impulse control system, which has the periodicity. To illustrate the period phenomenon existed in the system, we consider the single feedback criteria mentioned in Eq. (6.7) as ETC to trigger the control system.

Definition 2: Let $\Omega_M = \{x(t) | t \in \Gamma\}$ represents the impulse set, $\Omega_N = \{x(t) | t \in [0, T]/\Gamma\}$ represents the phase set. φ_1 is called impulse mapping, $\varphi_1 \in \Omega_M \to \Omega_N$; φ_2 is called system mapping, $\varphi_2 \in \Omega_N \to \Omega_N \cup \Omega_M$. $\varphi_2(x, t)$ represents the terminal point of trajectory described by system in Eq. (1), starting from point x running with time period t, $x \in \Omega_N$, $t \in [0, T]/\Gamma$.

Definition 3 [264]: If there is a point x^a in phase set Ω_N and a time period t_p such that $\varphi_2(x^a, t_p) = x^b \in \Omega_M$, it also has $\varphi_1(x^b) = x^a \in \Omega_N$, then $\varphi_2(x^a, t_p)$ is said to be the order-1 periodic solution with period t_p.

Lemma 2: The event-triggered control system in Eq. (6.3) has the order-1 periodic solution, if there exists $u(\tau_j) \geq \frac{x_I(\tau_j) - x_I(\tau_j - \eta)}{|g_I(x(\tau_j))|}$, $\forall \tau_j \in \Gamma$.

Proof: Based on assumption 1, $\Gamma \neq \emptyset$. Therefore, at least exists one time $\tau_j \in \Gamma$ can trigger the event-triggering condition. Using the single feedback ETC criteria such that:

$$x_I(\tau_j) - x_I(\tau_j - \eta) \geq \kappa \tag{6.16}$$

Since $u(\tau_j) \geq \frac{x_I(\tau_j) - x_I(\tau_j - \eta)}{|g_I(x(\tau_j))|}$, based on Eq. (6.3) there exists

$$x_I(\tau_j^+) \leq x_I(\tau_j - \eta) \leq x_I(\tau_j) - \kappa \tag{6.17}$$

In addition, $x_I(\tau_j) \geq x_I(\tau_j - \eta)$ meaning the trajectory x_I has the tendency to increase at time range $[\tau_j - \eta, \tau_j)$, therefore $x_I(\tau_j)$ did not reach the equilibrium under the non-control condition, which means

$$x_I(\tau_j^+) \leq x_I(\tau_j - \eta) < x_I(\tau_j) \leq x_I^* \tag{6.18}$$

Since x_I^* is asymptotically stable, $x_I(\tau_j^+)$ also has the tendency to increase. Accord to $x_I(\tau_j^+) \leq x_I(\tau_j - \eta)$, there exist a time $t^p \geq \tau_j^+$ that trajectory $\varphi_2(x_I(\tau_j^+), t^p)$ will reach $x_I(\tau_j - \eta)$ again. Meaning, there must exist another time $\tau_{j+1} \geq \tau_j^+ + \eta$ can trigger the event-triggering condition. At τ_{j+1}, it satisfies

$$x_I(\tau_{j+1}) \leq x_I(\tau_j) \tag{6.19}$$

If $x_I(\tau_{j+1}) = x_I(\tau_j)$, then it is obvious that the event-triggered control system (6.1) has the order-1 periodic solution. Otherwise, $x_I(\tau_{j+1}) < x_I(\tau_j)$ exists, we can define a non-increasing sequence $S = \{x_I(\tau_j), x_I(\tau_{j+1}) \ldots\}$.

Based feasible region of system variables defined in Eq. (6.2), we have that

$$x_I(t) \geq 0, \forall t \in [0, T] \tag{6.20}$$

Therefore, the sequence S has a lower bound c, where $c \geq 0$.

$$\lim_{n\to\infty} x_I(\tau_{j+n}) = c \tag{6.21}$$

For every $\varepsilon > 0$, there exists a positive integer K . For all $n > K$, it satisfies

$$\|x_I(\tau_{j+n}) - c\| \leq \varepsilon \tag{6.22}$$

When we consider a greater enough K with a small enough ε, it satisfies

$$x_I(\tau_{j+n}) = c = x_I(\tau_{j+n+1}) \tag{6.23}$$

Meaning, event-triggered control system in Eq.(6.3) has the order-1 periodic solution.

Lemma 3: If the event-triggered control system in Eq.(6.3) has the order-1 periodic solution, the period starts from $\tau_j{}^* \in \Gamma$. Then sequence $x_I(t)$ is bounded by a positive number B, when $t \in [\tau_j{}^*, T]$. Specifically, it satisfies $B \leq x_I^*$.

Proof: When the periodic exists in the system (6.3), if satisfies

$$x_I(\tau_j^-) = x_I(\tau_{j+1}^-), \forall \tau_j > \tau_j{}^* \tag{6.24}$$

During the impulse control time interval, it satisfies

$$x_I(t) \leq x_I(\tau_j^-), \forall t \in [\tau_j, \tau_{j+1}] \tag{6.25}$$

Consider B equal to the value of $x_I(\tau_j{}^*)$, then sequence $x_I(t)$ is bounded

$$x_I(t) \leq B, \forall t \in [\tau_j{}^*, T] \tag{6.26}$$

According to Eq (6.29),

$$B = x_I(\tau_j{}^*) \leq x_I^* \tag{6.27}$$

Therefore, when system trajectory reach the order-1 periodic, the infected population of event-triggered system is always less than or equal to the value of endemic equilibrium. Meaning, event-triggered system can be effective than non-control system in limiting the spread of disease when the system reach the stability or periodic.

6.4.3 Stability Analysis for Event-triggered Impulse Control System without periodicity

Lemma 4: If the event-triggered control system in Eq.(6.3) doesn't have the order-1 periodic solution, the system will converge into x^*.

Proof: Since the control system doesn't have the order-1 periodic solution, meaning that: $\forall x^a \in \Omega_N, t_p > 0$, such that $\varphi_2(x^a, t_p) = x^b \in \Omega_M$. It always has that $\varphi_1(x^b) = x^c \in \Omega_N$, where $x^a \neq x^c$.

Therefore, in any two contiguous impulse control trigger phases $[\tau_{j-1}, \tau_j)$ and $[\tau_j, \tau_{j+1})$, it satisfies that:

$$\Omega_{N(\tau_{j-1},\tau_j)} \cap \Omega_{N(\tau_j,\tau_{j+1})} = \emptyset, \forall \tau_j \in \Gamma \tag{6.28}$$

In addition, based on the characteristic of the epidemic system in Eq. (1), there exists a $\tau_{j'} \in \Gamma$. $\forall \tau_j > \tau_{j'}, \tau_j \in \Gamma$, satisfies that $x(t)$ is monotonous at $\Omega_{N(\tau_j,\tau_{j+1})}$.

Let $x'(\tau_j)$ to be any random selected point from $\Omega_{N(\tau_j,\tau_{j+1})}$, define the sequence $S' = \{ x'(\tau_{j'}), x'(\tau_{j'+1}) \ldots \}$, where S' is monotonous at Ω_N.

Based on the Lemma 1, x^* is an endemic equilibrium for the non-control system. Meaning, sequence S' has a limitation at Ω_N:

$$\lim_{n\to\infty} x'(\tau_{j'+n}) = x^* \tag{6.29}$$

In conclusion, the given impulse control system can at least reach the same equilibrium of the non-control system. In particular, if the impulse control system has the periodic with suitable control value, this system can reach a better steady state which has the lower infected population. Therefore, no matter the periodicity exists or not, the impulse control system in Eq. (6.3) is stable in Ω.

6.5 Simulation Studies

In this chapter, we illustrate the proposed EBIC method using numerical simulation. The simulation results prove that the EBIC method has the ability to predict the control levels based on the evidence database. By comparing with existed methods, the ETC-triggered EBIC method has better performance in both reducing the infected population and saving the total cost during the epidemic. In addition, we set up another numerical simulation using a multi-dimension system for Zika Virus epidemic, to test the best ETC setting and neural network design for EBIC. Finally, the robustness simulation is used to prove the EBIC can obtain valid information from the epidemic data with shifts.

6.5.1 Simulations for the effectiveness analysis

In this section, we compare the proposed EBIC control method with the other existed event-trigger control methods. The goal of this comparison is to verify the effectiveness of the EBIC method. This comparison is set up by a second-order standard SIS compartmental model, including a single vaccination control variable:

$$\frac{dS}{dt} = -0.01SI + 0.008I - uI \tag{6.30}$$

$$\frac{dI}{dt} = 0.01SI - 0.008I + uI$$

Where S represents the susceptive population, I represents the infected population, u represents the percentage of vaccination control usage. The initial values of the susceptive

population and infected population are 0.93 and 0.07, meaning there are 7% of the total population are infected at the initial stage. The objective function is defined by the overall cost due to the infectious with the cost spent on the vaccination control. For this simulation, we set the total population size as 100,000.

$$Minimize\ J(u) = \left(\int_0^T 10 * I(t)\, dt + \sum_{j=1}^{m} 200 * I(\tau_j) * u(\tau_j)^2\right) * population \quad (6.31)$$

Since the EBIC method required the pre-training process using historical data, we considered the simulation data with different initial values in 5 scenarios as the evidence database. Among which, 70% of the evidence database was used as training data; 30% of the evidence database was used as testing data. The simulation was set up by using the system in Eq. (6.30) with exactly the same objective function in Eq. (6.31). During the data collection, the control values are calculated by using the Pontryagin's Maximum Principe (PMP) based optimal control algorithm, which is introduced by our previous publication[41, 50]. As we mentioned in section 6.3.4, the evidence database consists of both evidence input and output data. Where the input data is collected from observing the system variables, the output data is collected from the records of the optimal control values obtained via the PMP method.

From the previous literature, other methods were introduced by using the impulse controls to limit the spread of the epidemic. Some of them considered the pre-defined value for the vaccination control variables [245, 251, 254, 255]. Meaning, they defined the values of vaccination usage percentages before the epidemic. When the impulse control is triggered by ETC, the system will use the predefined value as the control level. Other existed literature considered the feedback control, meaning the control level is depends on the value of the system variables [246, 252]. Typically, in the epidemic system, the more infected population will cause a higher value of the control variable.

In the following, we compare the EBIC with the existed impulse control methods, to verify the effectiveness of the EBIC method. The simulation comparisons adopt the unanimous trigger condition, which is the feedback trigger ETC mentioned in Eq. (6.7). The phase portrait comparisons are shown in Figure 6.6, and the simulations results are summarized in Table 6.2.

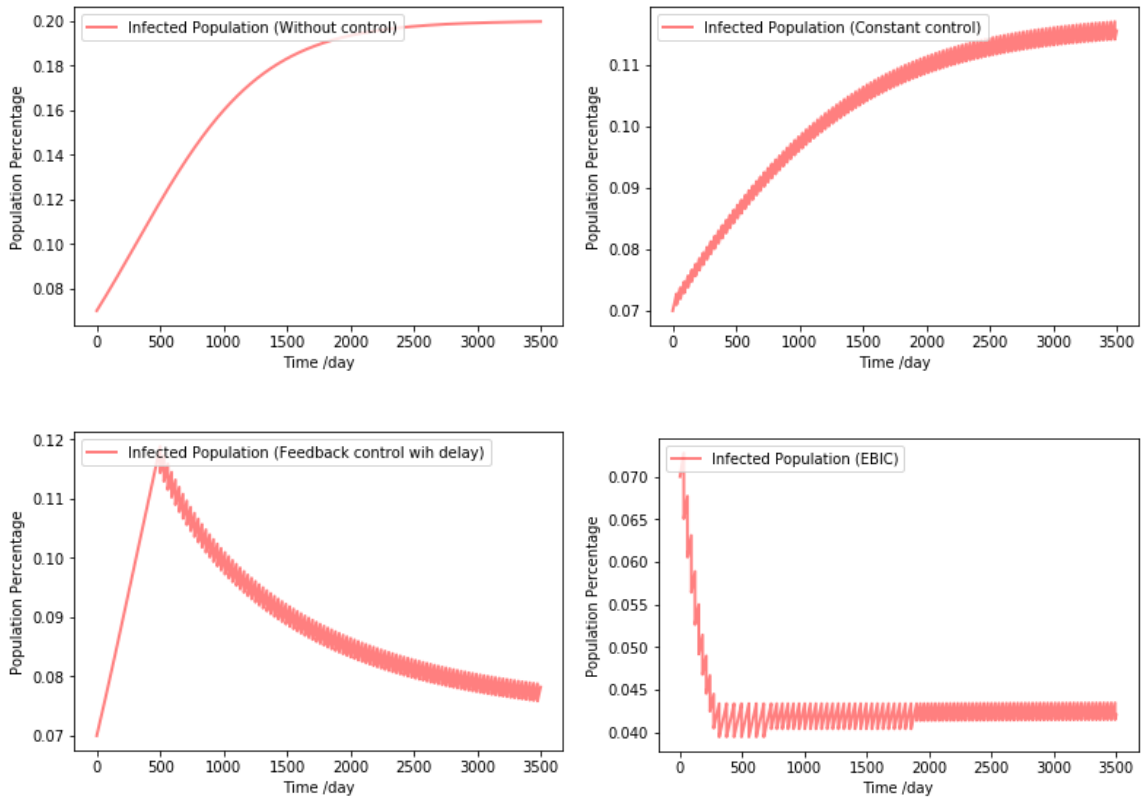


Figure 6. 6 Phase portrait comparisons for different impulse control methods in time series

Through the comparisons in Table 6.2, the EBIC method shows the effectiveness in reducing objective function value during the entire simulation period. In the meanwhile, EBIC spent less in the control cost than predefined control and feedback control as well. In particular, the predefined control method has a similar average value of control than the EBIC method, however, EBIC uses less control trigger times. In the meantime, the feedback control method has a similar control trigger times than the EBIC method, however, EBIC spends less on the value of each impulse control.

Table 6. 2 Efficiencies and steady states comparisons for different impulse control methods

	Objective Function	Cost due to Control	Number of Controls	Average Value of Control	Stable Status	Equilibrium Point (S,I)	Periodic Lower Bound (S,I)	Periodic Upper Bound (S,I)
Without Control	$5.95*10^8$	0	0	0	Endemic Equilibrium	0.8, 0.2	None	None
Predefined Control	$3.62*10^8$	$2.99*10^6$	116	0.5	Periodic	None	0.8794, 0.1206	0.8832, 0.1168
Feedback Control(delay)	$3.16*10^8$	$5.02*10^6$	100	0.754	Periodic	None	0.9268, 0.0732	0.9241, 0.0759
EBIC	$1.52*10^8$	$1.21*10^6$	94	0.502	Periodic	None	0.9587, 0.0413	0.9568, 0.0432

Through the simulation comparison in Figure 6.6, all of the predefined control, feedback control and EBIC methods can reach the periodic steady-state, which has a lower infected population percentage than the equilibrium point of the non-control system. Among these three control methods, EBIC receives better values in both lower bound and upper bound when the system reaches the periodic.

6.5.2. Simulations of ETC comparisons

In this and the following simulations, we considered a multi-dimensional dynamic system for the Zika Virus epidemic, including a single control variable [35]. In this section, we compare the EBIC control method with the well-known PMP optimal control method in controlling the dynamic system of Zika Virus. Since the evidence database of the EBIC control is simulated by the PMP optimal control method, the EBIC control is expected to obtain the features of the PMP optimal control. Therefore, the object of this section is to verify the similarity in the trajectories of two types of control methods. To ensure the comparison equity, this simulation doesn't consider the event-trigger conditions.

$$\frac{dS_h}{dt} = -ab\frac{I_v}{N_h}S_h - \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h \tag{6.32}$$

$$\frac{dE_h}{dt} = \theta\left(ab\frac{I_v}{N_h}S_h + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h\right) - v_h E_h \quad (6.33)$$

$$\frac{dI_{h1}}{dt} = v_h E_h - \gamma_{h1} I_{h1} \quad (6.34)$$

$$\frac{dI_{h2}}{dt} = \gamma_{h1} I_{h1} - \gamma_{h2} I_{h2} \quad (6.35)$$

$$\frac{dA_h}{dt} = (1-\theta)\left(ab\frac{I_v}{N_h}S_h + \beta\frac{\kappa E_h + I_{h1} + \tau I_{h2}}{N_h}S_h\right) - \gamma_h A_h \quad (6.36)$$

$$\frac{dR_h}{dt} = \gamma_{h2} I_{h2} + \gamma_h A_h \quad (6.37)$$

$$\frac{dS_v}{dt} = \mu_v N_v - ac\frac{\eta E_h + I_{h1}}{N_h}S_v - \mu_v(1+u)S_v \quad (6.38)$$

$$\frac{dE_v}{dt} = ac\frac{\eta E_h + I_{h1}}{N_h}S_v - (v_v + \mu_v(1+u))E_v \quad (6.39)$$

$$\frac{dI_v}{dt} = v_v E_v - \mu_v(1+u)I_v \quad (6.40)$$

Where $S_h(t), E_h(t), I_{h1}(t), I_{h2}(t), A_h(t), R_h(t)$ respectively represent the population of susceptible, exposed, symptomatically infected, convalescent, asymptomatically infected, recovered human, $S_v(t), E_v(t), I_v(t)$ respectively represent the population of susceptible, exposed, infectious vectors. u is the control (in percentage) that can improve the death rate of the mosquito population (e.g., spraying the insecticide). The parameter descriptions and the initial values of the variables are summarized in the appendix.

$$Minimize\ J(u) = \int_0^T 8000 * I_{h1}(t) + 1000 * I_{h2}(t)dt + \sum_{j=1}^{m} 5 * 10^6 * u(\tau_j)^2 \quad (6.41)$$

The objective function is defined by the combining the overall costs due to infectious treatment and spreading the insecticide. Since this problem will be solved by using EBIC framework, evidence database is necessary for the pre-training process. To simplify the data collection process, we will use the PMP-based optimal control algorithm mentioned in section

6.3.1 to simulate the evidence database. Same like the section 6.5.1, the 70% of the evidence database is used as training set, the rest 30% is used as testing set.

In the Figure 6.7, the EBIC control is compared with the PMP optimal control. These system variables are calculated by using the PMP optimal control method. Under the assumption of the EBIC control is always triggered, the EBIC is calculated based on the same system variable values. This comparison assumed that 250% of the normal death rate of mosquito is the maximum value by spraying insecticide.

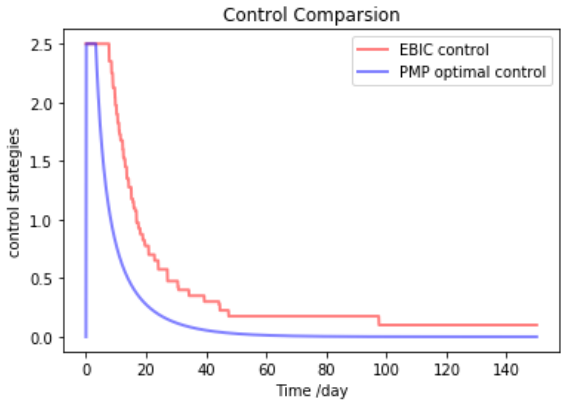


Figure 6. 7 EBIC and PMP optimal control values comparison

The comparison proves that EBIC control is feasible since it has a similar trajectory in time series than PMP optimal control. However, in the entire time series, the values of EBIC control are significantly higher than PMP optimal control. Meaning, if we assumed the EBIC control is always triggered, the value gaps between EBIC control and PMP optimal control will be wasteful. Therefore, we will compare different trigger methods with EBIC control in the following section, the goal is using the limited trigger times to minimize the total cost and the number of infections.

6.5.3 Trigger methods comparison for EBIC

In this section, the feedback trigger, CTSETC, and CEBETC methods are compared by using the given Zika Virus epidemic system. As we mentioned in section 6.3.2 and 6.3.3, CTSETC can synthetically evaluate the current epidemic; CEBETC can intelligently trigger the impulse control. The object of this comparison is to verify the effectiveness of the proposed ETCs in limiting the transmission of Zika Virus.

Table 6. 3 Efficiencies comparisons for different ETCs

METHOD	TRIGGER METHOD	TOTAL INFECTIONS	NUMBER OF TRIGGERS	COST OF INFECTIONS	COST OF CONTROL	OBJECTIVE FUNCTION	IMPROVE %
Without control	N/A	22944	0	4.13*10^9	0	4.13*10^9	0
PMP-based Optimal control	Continuous	6141	Continuous	1.10*10^9	5.16*10^8	1.62*10^9	60.78%
EBIC (Single variable)	Feedback Trigger	10930	192	1.58*10^9	4.26*10^8	2.01*10^9	51.33%
	CEBT	12401	106	1.78*10^9	3.31*10^8	2.11*10^9	48.91%
EBIC (Kernel system variable)	CTST	9561	412	1.42*10^9	5.05*10^8	1.93*10^9	53.27%
	CEBT	7013	875	1.29*10^9	7.95*10^8	2.09*10^9	49.39%
EBIC (Single variable with moving average)	Feedback Trigger	11458	157	1.73*10^9	4.15*10^8	2.15*10^9	47.94%
	CEBT	14940	91	2.69*10^9	2.84*10^8	2.98*10^9	27.85%
EBIC (Kernel system variable with moving average)	CTST	10511	337	1.59*10^9	5.13*10^8	2.12*10^9	48.67%
	CEBT	10803	693	1.65*10^9	7.08*10^8	2.35*10^9	43.10%

In Table 6.3, we compare the feedback trigger, CTSETC, and CEBETC trigger methods by using the EBIC control methods with different variable types. The comparison results are synthetically reflected by the infection numbers, control trigger numbers, and the total cost during the epidemics. It is apparent that PMP-based optimal control is the best to minimize the value of the total cost during the epidemic (objective function). However, the PMP method can only provide continuous control, which is hard to be executed in the real world. For the EBIC method with single variable criteria, the feedback trigger method is compared with the CEBETC

method; for the kernel processed system variable, the CTSETC method is compared with the CEBETC method. From the comparison, CEBETC is less effective than feedback triggers and CTSETC methods in both the infected population and cost function.

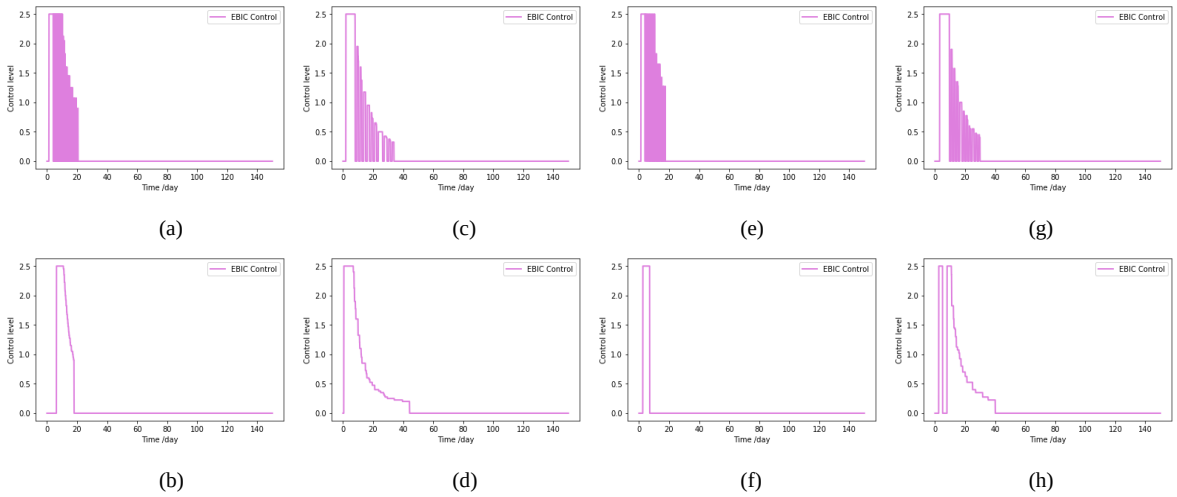


Figure 6. 8 Control levels comparisons for: (a),(b) EBIC (single variable) with Feedback trigger/CEBT; (c),(d) EBIC (kernel system variable) with CTST/CEBT; (e),(f) EBIC (single variable with moving average) with Feedback trigger/CEBT; (g),(h) EBIC (kernel system variable with moving average) with CTST/CEBT

Even though CEBETC method receives a higher value in the cost and infected population, it can intelligently trigger the impulse control without using human intervention at ETC criteria. In Figure 6.8, we compared the EBIC control triggered by feedback trigger/CTSETC with the EBIC control triggered by CEBETC. From the comparison, the CEBETC can trigger the control at almost same time period than the feedback trigger/CTSETC method. Therefore, the CEBETC method has the potential to be improved through the training by the dataset with more relevant system information.

6.5.4 Simulations of neural network architecture design for EBIC

In this section, we still utilize the dynamic system for Zika Virus to research the neural network architecture of EBIC method. As we mentioned in the section 6.3.4, many types of neural networks can be used as classification models to classify the EBIC control level. In the

following, we use the same training set in section 6.5.3 to compare the classification accuracies and speeds of different neural networks like fully connected network (FCN), CNN, RNN, GRU and LSTM.

The historical evidence dataset S_{EBIC} is generated by the 5 years simulation of dynamic model for Zika Virus. The evidence control set is calculated by using PMP based optimal control method. In addition, we classified the control variables at each time epoch into 100 control level sets based on the value. In the validation process, we randomly selected 70% data from the S_{EBIC} as the training set, 30% data as the validation set. To test the potential overfitting, we also compare the accuracy rates by using the cross validation.

Table 6. 4 EBIC architecture comparisons for different neural networks

	50 iterations running time	**Accuracy in training**	**Accuracy in cross validation**	**200 iterations running time**	**Accuracy in training**	**Accuracy in cross validation**
FCN	19.92s	68.71%	67.37%	77.48s	75.98%	75.31%
FCN(dropout)	24.48s	71.64%	70.01%	93.89s	77.40%	75.71%
CNN	706.70s	77.54%	76.70%	3032.89s	77.88%	76.63%
CNN(dropout)	740.79s	77.29%	75.55%	3225.02s	77.18%	75.89%
RNN	75.14s	77.08%	74.26%	331.72s	78.24%	76.31%
RNN(dropout)	85.65s	77.30%	77.21%	339.03s	77.74%	77.10%
GRU	163.84s	77.41%	74.50%	595.51S	78.14%	77.86%
GRU(dropout)	173.82s	76.58%	76.04%	717.20s	78.80%	76.58%
LSTM	177.89s	77.03%	73.88%	779.24s	78.55%	77.58%
LSTM(dropout)	180.91s	76.01%	75.13%	791.45s	78.20%	77.49%

In the Table 6.4, we compared the running times and classification accuracy rates in both 50 iterations and 200 iterations for different neural networks. The running time is defined as the time to train and validate the neural network. The classification accuracy rate is defined as how many percentages of the control variables are exactly on the same level as the PMP optimal control algorithm.

Through the comparison, FCN has the fastest training speed, in the meantime, it also requires more iterations to convergence than other neural networks. Most of the other neural networks can convergence into around 77% accuracy rate in 50 iterations; around 78% accuracy rate in 200 iterations. However, the training speeds are significantly different. Standard RNN takes around 80 seconds to convergence into a 77% accuracy rate, the rest of them require 160 seconds or higher to convergence into a similar rate. Meaning, RNN is the most efficient neural network by taking both training speed and accuracy rate into account.

In addition, researchers already reached a consensus in using cross-validation to test the potential overfitting existed in the classification model [243, 265]. Typically, if there is a significant difference between the predicted accuracies with and without cross-validation, meaning the trained model may exist overfitting. From Table 6.4, although there is a slight difference between the accuracy rates with and without cross-validation, most of the gaps are less than 2 percentages. Meaning, this model doesn't include obvious overfitting through the training process.

Table 6. 5 RNN architecture improvement comparisons with 50 iterations

Experiment	Unroll	Bias vector	Return last output	One direction running time	Accuracy with one direction	Bidirection running time	Accuracy with bidirection
#1	N	N	N	75.14s	77.08%	63.24s	76.06%
#2	N	N	Y	50.64s	77.73%	47.54s	76.36%
#3	N	Y	N	64.69s	76.47%	64.20s	75.57%
#4	N	Y	Y	58.21s	77.09%	52.83s	76.53%
#5	Y	N	N	56.02s	76.28%	50.68s	76.70%
#6	Y	Y	N	63.45s	75.55%	51.91s	76.34%
#7	Y	N	Y	48.21s	77.78%	39.85s	77.32%
#8	Y	Y	Y	48.73s	76.40%	39.42s	76.53%

As we mentioned in the section 6.3.4, standard RNN can be further improved by adding components in architecture design. From Table 6.4, we already knew the best-predicted accuracy

rate is around 77%-78% using the given database. To further improve the efficiency of the RNN architecture, we design the experiments to speed up convergence. Methods like unrolling, adding bias vector, returning last output of the sequence and bidirectional training are tested. Through the experiments, the best performance was received by BURNN, which includes the unrolling, only returning last output of the sequence and bidirectional training. BURNN can finish the 50 iterations of training in 39.85s, which is significantly faster than standard RNN.

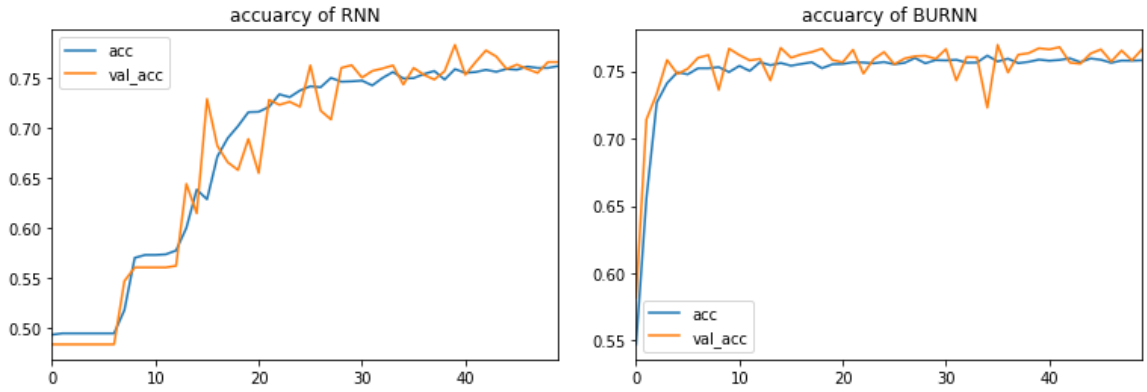


Figure 6. 9 Convergence plots comparison for standard RNN and BURNN

In addition, since bidirectional training can fully take advantage of all available input information in the past and future, BURNN can train the neural network with fewer iterations. From Figure 6.9, we know that BURNN can convergence into a 77% accuracy rate using less than 10 iterations. However, the standard RNN requires round 40 iterations to reach the same value of accuracy.

6.5.5 Simulations of Robustness and Validity Analysis

In this section, we simulate the real-world evidence data with the time shifts and magnitude differences. These data are used to train the neural networks for calculating the control variables. The goal of this simulation is to verify the robustness of the EBIC method under the unperfect data. In addition, the final simulation compares the validities of historical

controls and EBIC control. The goal of this comparison is to verify that EBIC control has a better performance than the empirical control strategy for the future epidemic.

Similar to the data collection for Section 6.5.4, the historical evidence dataset S_{EBIC} is generated by the 5 years simulation of a dynamic model for the Zika Virus. However, in this section, the historical data are simulated with time shifts and magnitude differences in the epidemic. The infections related variables and control strategies from the historical simulations are shown in Figure 6.10.

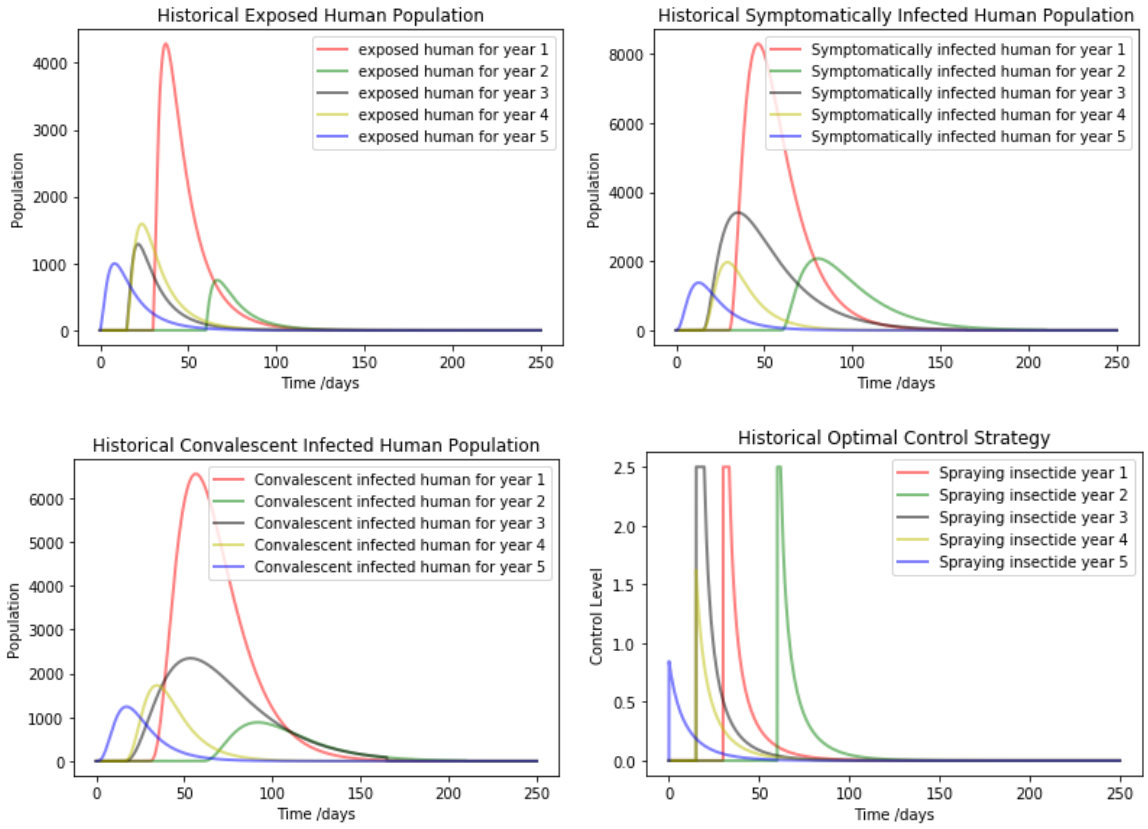


Figure 6. 10 Illustration of the evidence data with time shifts and magnitude differences

Randomly selected 70% of the total evidence data as the training set; 30% of the total evidence data as the validation set. Then, we train the BURNN by the training set and compare the accuracy and loss function for both the training set and validation set. The comparison is trained with 200 iterations, and the details are shown in Figure 6.11.

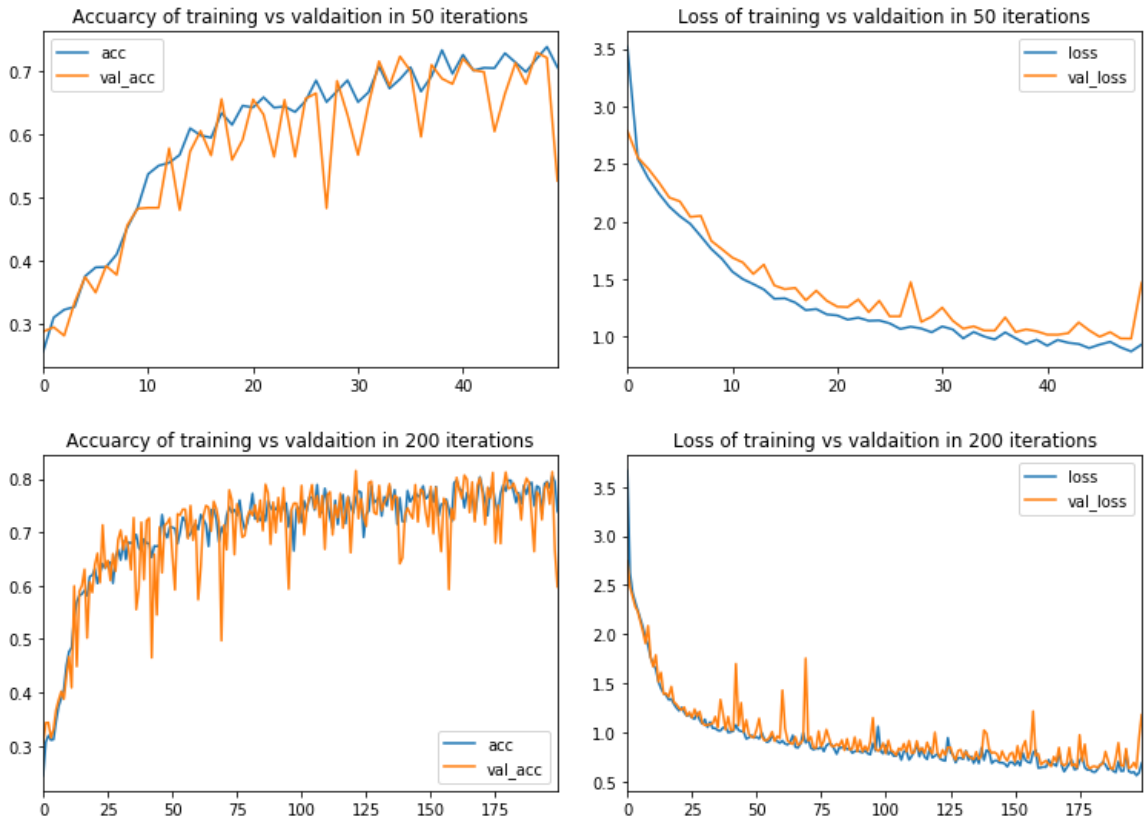


Figure 6. 11 Accuracy and loss comparisons for the BURNN training with time shifts and magnitude differences data

Through the comparison, the difference of accuracy and loss between the training set and the validation set are tiny. Meaning, even though the overfitting may exist in the training process, the influence still can be negligible. That comparison verifies that the EBIC method is robust if the training and validation sets are randomly selected from the same resource.

However, in the real world, the training and validation sets typically come from different resources. Therefore, in the following comparison, we consider the Zika Virus data from 5 years of simulation as the training set. Then, we consider another 1-year Zika Virus data with different time shifts and magnitude differences as the validation data. The comparison results are shown in Figure 6.12.

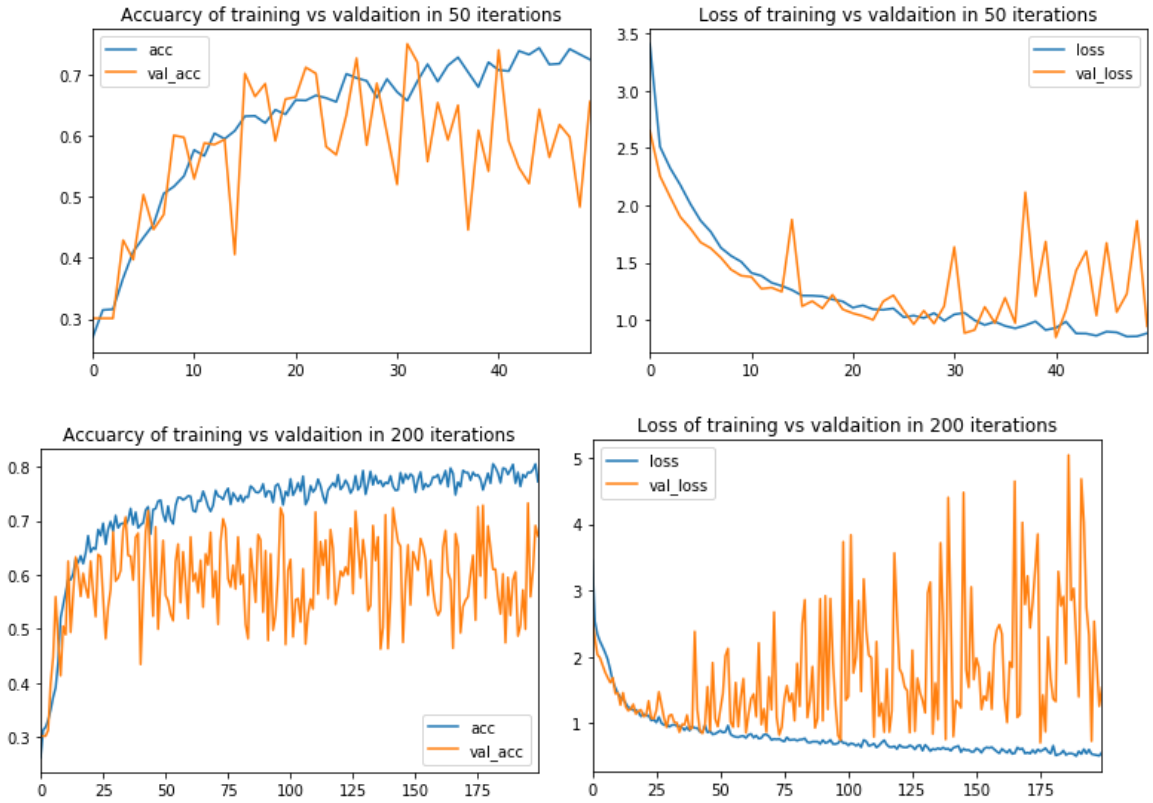


Figure 6. 12 Accuracy and loss comparisons for the BURNN training validated by the different resource

When the validation set comes from a different resource as the training set, there exists a significant difference between the accuracies from two sets. The validation accuracy is obviously lower than the accuracy of the training set. Moreover, the loss value and the variation level of the validation set increases with more training iterations. Meaning, the overfitting exists when the training process takes more than 50 iterations. For the training process after the 50^{th} iteration, even though the accuracy of the training set is still increasing, the performance of the neural network becomes worse. Therefore, to ensure the robustness, we shouldn't consider too many iterations of the training process.

The final simulation is to compare the efficiencies of EBIC control with 5 historical control strategies for the upcoming epidemic. The 5 historical control strategies are generated

from the above evidence data in 5 historical years. The EBIC control is set up by the BURNN, which is trained by using the above evidence data in 5 historical years. The upcoming epidemic is simulated by using the different disease system parameters and initials for the variables. The comparison results are shown in Figure 6.13.

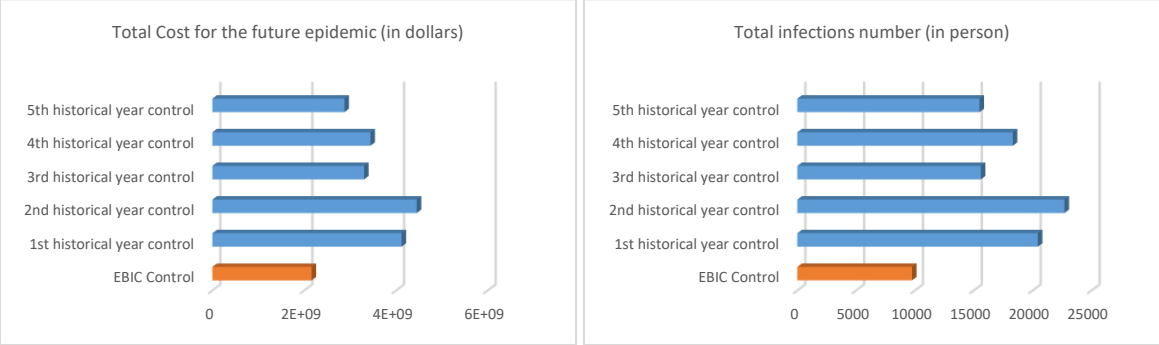


Figure 6.13. Total cost and infections number comparisons

Figure 6.13 shows that the EBIC has a lower total cost with the lower total numbers of the infections. Meaning, the EBIC has been proved to learn the control strategies from the historical evidence, then utilizes the learned rules to limit the transmission of the future epidemic. This comparison also proves that EBIC is obviously better than empirical strategies in controlling the upcoming epidemics.

6.5 Conclusion

Aiming to reduce the cost of unnecessary epidemic controls, the new disease control method has been proposed with the EBIC and ETCs for the epidemic dynamic systems. By taking advantage of the historical epidemic and control data, EBIC trains several neural networks to control the spreading of infectious diseases. Therefore, EBIC can intelligently analyze the future epidemic status and provide the appropriate control strategies. In addition, ETCs have been designed to avoid executing unnecessary control strategies. These ETCs have been triggered only if the current prevalence has the potential to deteriorate. The stability of the epidemic system under the event triggered conditions that have been proved by this paper, no

matter the system has the periodicity or not. Five groups of simulation comparisons have been given to prove the validity, accuracy, optimality, and robustness of the proposed method.

Designing two types of ETC is one major contribution of this work. Both CTSETC and CEBETC are able to determine the best time epoch of impulse control. CTSETC utilizes the kernel matrix method to synthetically integrate information from multi-variables. The integrated epidemic information will be analyzed by the time series method to automatically trigger the disease control. Moreover, CEBETC achieves one more step than CTSETC. CEBETC trains the convolutional neural networks to obtain the relation between historical epidemic data with the control trigger data. Then, the trained neural networks are used to intelligently trigger the epidemic impulse control.

Another major contribution of this work is the design of the EBIC mechanism. The EBIC method proposes a new method to determine the optimal level of impulse control after the ETC is triggered. The EBIC method takes advantage of the evidence epidemic data and control strategies. Similar to the CEBETC method, the EBIC method trains the neural networks to learn the relation between evidence epidemic data and control strategies. Many neural networks like fully-connected networks, CNN, RNN, LSTM, and GRU are tested for the EBIC method. Through the simulation comparisons, the revised BURNN receives good performance in both predicted accuracy and running time.

This paper also proves the epidemic system with impulse control is stable under the ETCs. The simulation part of this paper verifies that the EBIC predicted control has a similar trajectory with theoretical optimal control. In addition, the simulation reflects the robustness of the EBIC method, and provides the suggestions to avoid the overfitting during the training process.

This paper only provides the theoretical methods in the design of the EBIC and ETCs. To make this contribution valuable in real-world epidemic control, the future work of this paper will focus on real-world data collection and implementation. Instead of using simulation data, future work will collect the historical epidemic data from health organizations to train the neural networks. The expected result is to predict on-time disease control strategies for the policymakers.

Chapter 7 - Modeling Learning and Forgetting Processes with the corresponding impacts on Human Behaviors in Infectious Disease Epidemics

Chapter 7 is based on the manuscript “Modeling learning and forgetting process with the corresponding impacts on human behaviors in infectious disease epidemic” Published in Computers & Industrial Engineering [36].

Abstract

This chapter presents two new mathematical models, an information forgetting curve (IFC) model and a memory reception fading and cumulating (MRFC) model, to examine forgetting and learning behaviors of individuals during an infectious disease epidemic. Both models consider how epidemic prevalence and community behavior-change information may affect agent emotions and subsequently influence an individual's behavior changes during an epidemic. The IFC model utilizes a forgetting curve to process epidemic information, and the MRFC model formulates disease information variations using the Itô diffusion process. Sensitivity analysis and simulation comparisons showed that the MRFC model more accurately describes the epidemic with high lethal rate gets high attention. The author also demonstrated that MRFC model has higher sensitivity parameters and is more flexible on wide ranges of infection rates than the IFC model. However, the IFC model is a better suited for widespread, low-risk mortality epidemics, such as seasonal influenza, the infection information and protective behavior have close relationships among the susceptible population. An agent-based simulation model also developed to mimic the epidemic prevalence of the 2009 Chicago H1N1 using public available historical data sets by IFC model.

Keywords: Forgetting and learning, epidemics, behavior changes, agent-based simulation.

7.1 Introduction

Human disease awareness and related behavior changes during disease epidemics have recently attracted considerable research attention [266]. In order to be more accurately predict a disease epidemic and estimate its potential impacts, however, a comprehensive understanding of information dissemination within human contact networks and the effects of this information on human emotions, awareness, and behavior must increase. Extensive literature and studies have investigated how information affects human behaviors, but minimal research has focused on human memory and forgetting/learning processes related to disease information, or the process in which information may be forgotten and relearned during an epidemic episode. This paper proposes two new mathematical models to investigate the effects of information in disease transmission, including the forgetting and learning phenomenon.

The human brain cannot store an infinite amount of retrieved information. In 1968, Atkinson and Shiffrin first classified memory as long term and short term [267]. Engle *et.al* defined short-term memory as retainable for a short period of time (usually from 6 to 600 seconds) but unable to be manipulated; however, they did not detail how long-term or short-term memory relates to people forgetting information [268]. Ebbinghaus *et.al* experimentally investigated how the process of forgetting proceeds with influences of time or daily events, hypothesizing that, although a memory series is gradually forgotten, memories that have been learned twice fade more slowly compared to memories that have been learned once [269]. Wingfield *et.al* proposed a "forgetting curve" to show the process of memory loss over a period from 20 minutes to 31 days [270]. A recent paper has indicated that the people learn language also following the forgetting curve [271], and experiments have been conducted to increase understanding of the learning and

forgetting phenomenon [272, 273]. In 1985, Brainerd *et. al* concluded that forgetting is governed by various laws and therefore requires unique theoretical assumptions [274].

Notable learning and forgetting mathematical models have been proposed. In 1976, Carlson *et. al* introduced the variable regression variable forgetting (VRVF) model [275], and, in 1990, Elmaghraby proposed the variable regression invariant forgetting (VRIF) model, which corrected errors in previous forgetting models and accommodated a finite horizon [276]. In 1996, Jaber *et. al* proposed the learn-forget curve model (LFCM), which showed that forgetting is dependent on some factors such as the learning slope, the quantity produced and the minimum production breaks [277]. In 1997, Jaber *et. al* compared these three models, and in recent years, Jaber *et. al* reviewed factors that influence forgetting and incorporated the job similarity factor into the LFCM [278]. In 2002, Sikström & Jaber provided an elaborate review of forgetting curves in psychology and industrial engineering literature [279]. Because the previous papers did not consider the forgetting phenomenon in disease, however, this paper proposes an information forgetting curve model (IFC) to describe how disease information fades over time during an epidemic, following a forgetting curve through time and thereby influencing final disease memory. In 2012, Sikström & Jaber updated their research in the modeling of learning and forgetting area [280]. They proposed a Depletion-Power-Integration-Latency (DPIL) Model. This model considered the depletion of the encoding resource as forgetting and learning behavior when the system repetitively performs a task. More than fitting the historical dataset and calculate the settings of optimal performance, this model discussed how learning can interact with the forgetting by modeling the repetitively encoding can increase the memory strength.

Stochastic factors can influence the information perception process and new information can diversely affect human memory when agents receive new information and forget previous

information. Researchers have shown that biased media coverage, misleading personal experiences, and anxieties can cause people to process information with unwarranted confidence or uncertain judgment [281]. Zhao *et.al* also discussed the stochastic change rate of perception to infectious disease risk [282]. This paper proposes a memory reception fading and cumulating (MRFC) model to describe the stochastic phenomenon of human memory as it pertains to disease information based on learning and forgetting.

Agents have unique understandings based on identical information, and they react with distinct switch behaviors. In 2009, Chen concluded that agents learn prevalence through the spread of information and can adjust human behavior during a disease epidemic [283]. Funk *et al.* found that the spread of disease awareness significantly decreases the infection rate [284], and Kiss *et al.* proved that the diffusion of disease information increases risk awareness and causes the host population to take infection prevention measures [285]. In 2015, Zhao *et al.* proposed a disease model using a spatial evolutionary game to illustrate the impact of information dissemination on human behavior in an epidemic, proving that how an agent feels depends on information content and context [286]. In other words, agents demonstrate unique perspectives for the same disease information, resulting in diverse emotional responses. When disease information is positive, agents may have minimal concerns about the disease; negative information, however, may increase agent's awareness. Hence, certain types of information could alter agents' moods or emotions [287]. Chen *et al.* also modeled how disease information, such as the number of infected individuals and the number of susceptible individuals who choose the switching behaviors, impact agents' fears about the epidemic [41]. This paper applies an agent-based model to determine how disease information can cause diverse human behaviors, as well as use of the one-factor-at-a-time method (OFAT) to conduct sensitivity analysis for various parameter settings in the IFC and

MRFC models, highlighting the influence of parameter settings for the model and comparing agents' epidemic behaviors using the IFC, MRFC, and no-memory models.

This paper also discusses the 2009 H1N1 influenza epidemic. By combining historical infection data in affected cities with corresponding population characteristics, the authors restored the 2009 H1N1 prevalence in Chicago, a typical H1N1-affected city. This paper also investigates how the phenomenon of memory fading and behavior-switched protection influence epidemic spreading.

7.2 Information Forgetting Curve Model

7.2.1 Contact Network and Disease Information

Contact networks have been widely applied to many implementations of disease transmission [288-290]. Zhao *et al.* introduced the concept of disease information dissemination and its effect on epidemic disease transmission by proposing that agents gain disease information from two layers: local and global contact networks [40]. Local information is gained from neighboring agents, while global information is acquired from all locations. The researchers used a spatial evolutionary game to figure out if agents switch behavior based on payoff [282]. In general, local contact networks contain many social cliques and more readily transmit pathogens, while global contact networks are usually dominated by non-face-to-face contacts and random long-distance connections [291].

Rapidly increasing advanced technology and popularity of the internet, social media, and new-media broadcasting channels have revolutionized the ways about information transmission and human communication. Sahneh *et al.* changed traditional global and local contact network divisions to disease transmission and information transmission divisions [292]. A disease transmission contact network (DTCN) is comprised of daily face-to-face contacts, such as family

members, neighbors, and colleagues, while an information transmission contact network (ITCN) includes all contacts in an agent's social media. In general, contacts in a DTCN are a subset of contacts in an ITCN (Figure 7.1).

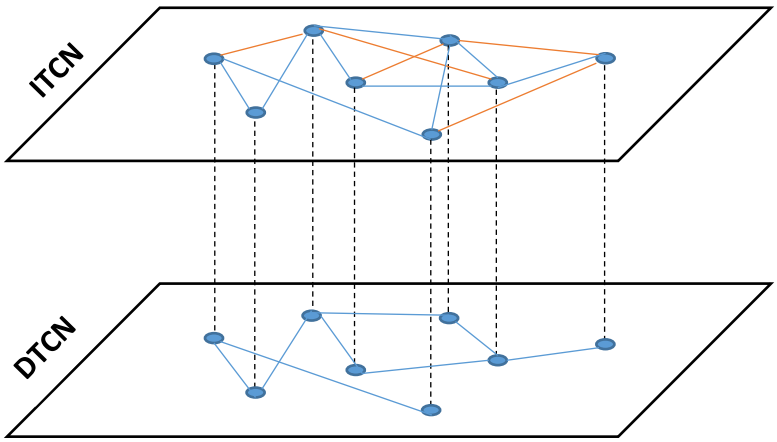


Figure 7. 1 Schematics of a two contact networks: ITCN and DTCN

Recent advancements in the information technology industry and the Internet of Things are spurring a rapid transition into an information era. Subsequently, disease information and its dissemination via modern information systems, such as social media, virtual communities, alternative media, and traditional media broadcasting, have begun to significantly influence disease transmission [293]. However, current research on the effects of disease information are primarily limited to the infected information (i.e., disease prevalence) [54, 294, 295]. In 2009, Chen found that self-protection qualities in a disease can positively influence disease transmission [283]. Agents frequently choose to use protective measures in an epidemic to reduce infection risks; these behaviors are known as switching behaviors. When agents choose not to take any protective measures in an epidemic, their behavior is referred to as normal behavior. In order to most accurately describe an individual's perception of information related to an ongoing epidemic, this research refers to a measure as perceived disease information (PDI), as defined in equation

(7.1), which can be divided into infected information (denoted as $I_i(t)$) and switch behavior information (denoted as $sw_i(t)$):

$$PDI_i(t) = \alpha I_i(t) + (1-\alpha) sw_i(t), \text{ where } 0 \le \alpha \le 1 \quad (7.1)$$

In equation (7.1), $PDI_i(t)$ represents new disease information of agent i at time t; $I_i(t)$ denotes the estimation of infected population percentages of agent i's contact network at time t, which represents the infected information; $sw_i(t)$ is the estimation of switch behavior population percentages of agent i's contact network at time t, which represents behavior-switched information; weighted parameter α denotes how an individual agent weights the proportion of infected information.

Since ITCN and DTCN can be utilized to collect information, $I_i(t)$ and $sw_i(t)$ can be divided based on the information source. $I_i^{ITCN}(t)$ and $I_i^{DTCN}(t)$ represent infected information collected by ITCN and DTCN, $sw_i^{ITCN}(t)$ and $sw_i^{DTCN}(t)$ represent behavior-switch information collected by ITCN and DTCN, and β is the weight parameter to determine the proportion of information sources. Thus, $I_i(t)$ and $sw_i(t)$ can be described as

$$I_i(t) = \beta I_i^{ITCN}(t) + (1-\beta) I_i^{DTCN}(t) \quad (7.2)$$

$$sw_i(t) = \beta sw_i^{ITCN}(t) + (1-\beta) sw_i^{DTCN}(t) \quad (7.3)$$

Then,

$$PDI_i(t) = \alpha\beta I_i^{ITCN}(t) + \alpha(1-\beta) I_i^{DTCN}(t) + (1-\alpha)\beta sw_i^{ITCN}(t) + (1-\alpha)(1-\beta) sw_i^{DTCN}(t) \quad (2.4)$$

Since I_i^{ITCN}, $I_i^{DTCN}(t)$, $sw_i^{ITCN}(t)$, and $sw_i^{DTCN}(t)$ are four percentages; also we have $\alpha\beta + \alpha(1-\beta) + (1-\alpha)\beta + (1-\alpha)(1-\beta) = 1$, the range of perceived disease information $PDI_i(t)$ is [0,1]. The high $PDI_i(t)$ indicates increased seriousness of the disease.

7.2.2 Information Forgetting in a Disease Epidemic

Human memory is a system that can store and retrieve information [296]. In 1913, Ebbinghaus introduced the forgetting curve to describe the information process of forgetting over time [269]. In 1991, Wixted *et. al* presented mathematical functions to represent the process of forgetting, including experiments to demonstrate and analyze the forgetting curve based on those forgetting functions [297]. The authors considered recall, recognition, and saving as measures of memory. The article also studied several materials to be remembered, including present words, faces, nonsense syllables, and graphic images. In addition, serval different subjects were used in the experiments, and the experiments were carried out at various time intervals. The study found that the process of forgetting can be represented mathematically using the following simple power function of time:

$$y = at^{-b} \tag{7.5}$$

In equation (7.5), y is a memory performance measure for the strength of the memory trace, or the proportion recalled by memory, and t represents time (one day as a unit period) [37]. Parameter a is the degree of learning, which represents the estimated level of performance after one unit of time, and parameter b is the rate of forgetting, where a and b range from 0 to 1 [38]. Therefore, this power function, y, ranges from 0 to infinity; that is, as t tends to zero, y tends to infinity, and as t tends to infinity, y tends to zero. Figure 7.2 (a) and (b) show the trend of function y when the degree of learning ($0 \leqslant a \leqslant 1$) and rate of forgetting ($0 \leq b \leq 1$) differ.

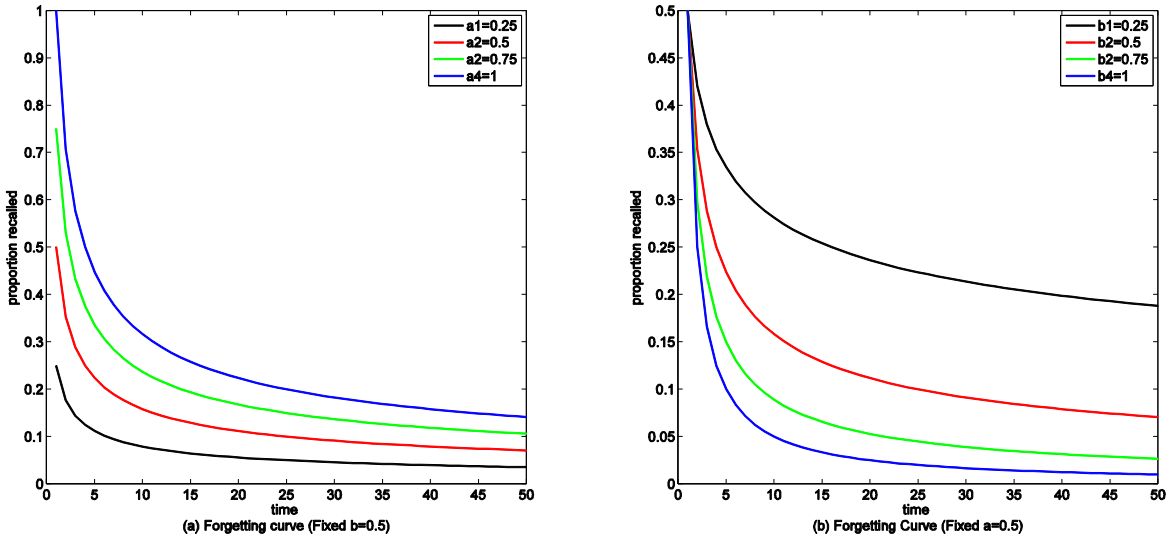


Figure 7. 2 Forgetting curves

Disease parameters such as infectious periods, numbers of infections, and disease susceptibility are considered disease knowledge or information and can be remembered or recalled by an individual agent [298]. This paper assumes that, similar to normal daily information, disease information acquired via disease parameters incorporates the learning and forgetting phenomenon. When agents receive similar information that they received before, corresponding memory is reinforced, causing slow memory fading. However, when signal of prevalence aggravation has not stimulated human memory, memory about the information follows the forgetting curve.

$$w_t = \frac{\int_t^{t+1} at^{-b} dt}{\int_0^{t_f} at^{-b} dt} \tag{7.6}$$

$$w_0 = \frac{\int_0^1 at^{-b} dt}{\int_0^{t_f} at^{-b} dt} = \frac{1}{{t_f}^{b+1}-1} \tag{7.7}$$

Equation (7.6) defines memory residual weight of disease information. In the equation, w_t represents the memory performance proportion for integration of y_t in one day at total memory performance on t days before current day, where the integration of at^{-b} from 0 to t_f represents

the sum of memory portions; integration of at^{-b} from t to $t+1$ represents the memory performance proportion in the $t-$th day. Equation (7.7) shows formulation of the special case (performance proportion of memory in the day before current day) w_0, meaning feasible longest memory epoch t_f and forgetting rate b. Therefore, the mathematical equation of final disease information (FDI) in the IFC model can be described as

$$FDI_i(t) = w_0 PDI_i(t) + w_1 PDI_i(t-1) + \cdots + w_t PDI_i(0) \tag{7.8}$$

where $FDI_i(t)$ represents FDI of agent i at time t. Since $I(t)$ and $sw(t)$ are percentages from 0 to 1, $PDI_i(t)$ has the range [0,1]. In addition, $\sum_{j=0}^{t} w_j = 1$, meaning the range of $FDI_i(t)$ is also [0,1]. For agent i, $FDI_i(t)$ reflects disease severity cognition at time t:

$$FDI_i(t) = w_0 PDI_i(t) + (1-w_0) FDI_i(t-1) \tag{7.9}$$

FDI also could be exponential smoothing, as shown in equation (7.9). Specifically, $w_0 PDI_i(t)$ could be new information learning, and $(1-w_0)FDI_i(t-1)$ could be past information forgetting. w_0 is a crucial adjective parameter in the IFC model. In the case of $w_0 = 1$, the agent is assumed to be memoryless, so FDI is equal to PDI and the agent chooses behavior based only on current information.

7.2.3 Fear Factor and Human Behavior in Disease

Epidemic information is often disseminated in correlation with the spread of disease. In addition to acquiring disease information through social networks, newspaper, or TV news, agents also gain disease information via communication with colleagues and families in their contact networks. This method of attaining disease-related information often includes an emotional response that influences agents' immediate behavior changes. For example, if agents know many individuals have become infected or expired at the outbreak of an infectious disease, they may fear the disease and take protective measures, such as decreasing travel, wearing masks, or becoming

vaccinated, to prevent infection. The process of transforming information to emotion and then to action is illustrated in Figure 7.3.

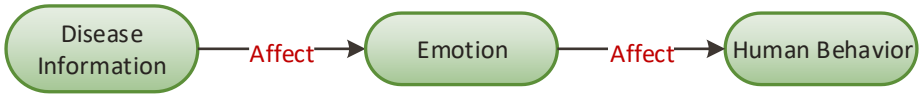


Figure 7. 3 Process of disease information affects human behavior

Chen *et al.* proposed a model to describe how disease information can affect individuals' emotions by introducing an individual fear factor ($IFF_i(t)$) [41]. Where t represents the current time, i represents the sequence number of the agent. Their model showed that individuals demonstrate diverse disease perceptions when they know infected individuals or switch individuals, thereby altering their emotions. For example, an increasing number of infected individuals in a neighborhood and increasingly negative media about the disease enhance, an individual's concern about the disease, consequently increasing the individual's fear factor. If fewer individuals become infected in neighboring areas, an individual tends to feel safer rather than fearful, resulting in a minimal individual fear factor. This paper assumes that disease information and an individual's emotions have significant correlation, as shown in equation (7.10). When the number of final disease information $FDI_i(t)$ is large (more close to 1), an agent has strong fear emotion about the disease; when the number of $FDI_i(t)$ is small (more close to 0), an agent has weak fear emotion about the disease.

$$FDI_i(t) \propto IFF_i(t) \tag{7.10}$$

Steimer stated that emotions such as fear can result in defensive behaviors [299]. That is, when agents feel fear, they tend to demonstrate self-protective behavior. Chen *et al.* described the relationship between emotion and human behavior using a logistic function [41]. An individual with a large fear factor tends to choose switch behavior with large probability, and an individual with a small fear factor tends to choose switch behavior with a small probability. Because emotion

and switch behavior have a positive correlation, this paper assumes that when $FDI_i(t)$ of agent *i* at time *t* is large (closer to 1 than 0), agent *i* possesses a strong fear emotion and he/she will likely choose switch behavior. However, when $FDI_i(t)$ of agent *i* at time *t* is low (closer to 0 than 1), agent *i* has a weak fear emotion, so agent *i* will likely choose to do nothing (i.e., normal behavior).

7.3 Memory Reception, Fading, and Cumulating Model

7.3.1 Memory Reception and Fading

Based on the study by Atkinson and Shiffrin, the memory fading process occurs at the same time as the information reception process [267]. The main objective of this section is to establish the memory fading model for received disease information. Similar to equation (7.4), gained information $\gamma_i(t)$ of agent i over time period t can be divided into infected information and surrounding switching behavior information collected on the ITCN and DTCN, $\gamma_i(t)$ represents all information transmitted from the information sources to an agent. However, the collected information varies with time, so $\gamma_i(t)$ is defined as changes in disease information of agent i from time $t-1$ to t. For example, $\Delta I_i^{ITCN}(t)$ is equal to $I_i^{ITCN}(t) - I_i^{ITCN}(t-1)$, which represents the change of infected information in ITCN from time *t-1* to time *t*.

$$\gamma_i(t) = \alpha\beta\Delta I_i^{ITCN}(t) + \alpha(1-\beta)\Delta I_i^{DTCN}(t) + (1-\alpha)\beta\Delta sw_i^{ITCN}(t) + (1-\alpha)(1-\beta)\Delta sw_i^{DTCN}(t) \quad (7.11)$$

$$\gamma_i'(t) = (\gamma_i(t) + 1)/2 \quad (7.12)$$

In real-world scenarios, individuals constantly perceive new information, and they tend to respond to and receive new cognition that occur infrequently [300]. During a disease epidemic, agents tend to receive strong stimulations of fresh information about the disease, thereby creating distinct contrasts in their memories. For example, when an agent receives new that breast cancer can be contagious (even it is not true), the perception regarding this new information is strong, so

the agent will diligently consider and remember this information. In addition, models presented in this paper incorporate a discount factor between received disease information $\gamma_i(t)$ and an individual's perception of disease information, defined as $PDI_i(t)$, since people typically overlook minor trifles and overemphasize crucial issues in their minds. Therefore, this research used the Hill equation as the discount between $\gamma_i(t)$ and $PDI_i(t)$. Since free ligand concentration must be positive, a transformation was applied in equation (7.12) to maintain positive received disease information $\gamma_i(t)$.

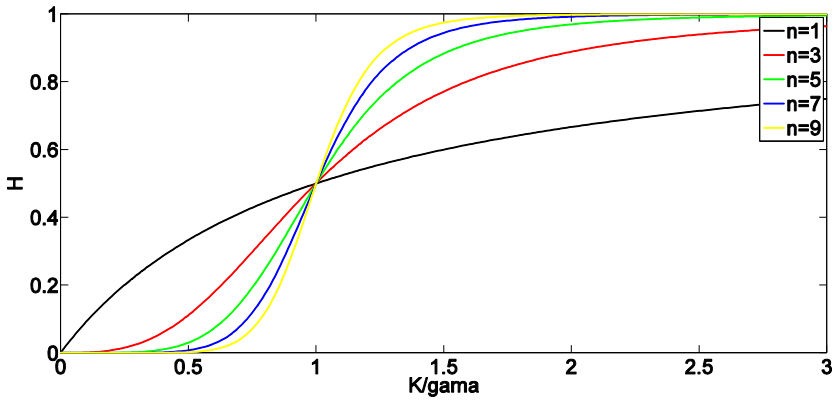


Figure 7. 4 Hill equation

The Hill equation (as shown in Figure 7.4), widely used in biochemistry and pharmacology, describes the fraction variation of a macromolecule in the molecular binding process [301]. Zhao *et.al* first introduced the Hill equation to describe memory perception rate in the disease transmission process [40]. In their assumption, the combination of new disease information and memory of past information is similar to the macromolecule binding process. The reasonable boundary of the Hill equation is between 0 and 1, this number from the hill equation represents the rate of information learning. The memory fading process is shown in equation (7.14), where ε is the forgetting rate, meaning that an individual could lose his/her memory of the disease over time. Disease information processing of agent i $\mu_i(t)$ represents adjusted disease information after

memory fading. c is an adjustive constant ($c = 0.5$) with a range of $-0.5 \leq H(\gamma_i(t), n) \leq 0.5$ since $\mu_i(t)$ is the parameter to represent new information about an epidemic and can be positive or negative, indicating whether the disease prevalence could be aggravated or mitigated, respectively.

$$H(\gamma_i(t), n) = \frac{(\gamma_i(t))^n}{(K)^n + (\gamma_i(t))^n} - c \tag{7.13}$$

$$\mu_i(t) = H(\gamma_i(t), n) - \varepsilon \tag{7.14}$$

$H(\gamma_i(t), n)$ in equation (7.13) and (7.14) represents newly gained disease information via a revised Hill equation; $\mu_i(t)$ is the disease information processing of agent i, which is a process variable between newly gained information $\gamma_i(t)$ and perceived disease information $PDI_i(t)$. K is the equilibrium constant and $n > 1$ is the Hill coefficient assumed to have a positive cooperative influence (binding) property in this application.

7.3.2 Information Cumulation

Because disease-related information may vary throughout an epidemic, an individual's perceived disease information, $PDI_i(t)$ also changes during the epidemic. As mentioned, an individual's perception of disease information contains both learning and forgetting processes; that is, the perception process is affected by an agent's previous memory and newly acquired disease information. Melanie *et.al* found that information reporting by mass media leads to the behavior change because media can disseminate information to many agents, subsequently influencing social networks and agents' decisions [302]. In addition, agents typically pay minimal attention to switch-protective behaviors in an epidemic when people have access to media coverage of the disease [47]. Zhao *et.al* proposed that memory accumulation and fading for disease information can be found using a stochastic differential equation of Itô drift-diffusion process, including a

drifting factor and random walk, to predict when an individual switches to a protective behavior [40]. This research identifies perceived disease information $PDI_i(t)$ as disease prevalence information remembered by agent *i* at time t. As described in Section 3.1, the processing disease information $\mu_i(t)$ represents the difference between gained information and perceived disease information. The larger the difference of the current disease information, the larger the value of the processing disease information $\mu_i(t)$, meaning that $\mu_i(t)$ is the change magnitude of disease information used to represent a drifting factor in the Itô drift-diffusion process. An uncertain factor, also present during an epidemic due to population diversity and uncertainties, can be modeled as a random walk in the Itô drift-diffusion process. Similarly, in this paper, we assumed that the stochastic process of perceived disease information $PDI_i(t)$ in the MRFC model is an Itô drift-diffusion process as

$$dPDI_i(t) = \mu_i(t)dt + \sigma dZ_t \tag{7.15}$$

where $Z = \{Z_t : t \in [0, \infty)\}$ is standard Brownian motion with a mean of 0 and standard deviation of 1. $\mu_i(t)$ is a drafting factor that represents the processing disease information, and $\sigma(t)$ represents the variance of randomness in the population. The initial value $PDI_i(0)$ is calculated based on the definition in equation (7.4); the definition of $PDI_i(t)$ is shown in equation (7.16).

$$PDI_i(t) = PDI_i(t-1) + \int_{t-1}^{t}(\mu_i(t)dt + \sigma dZ_t) \tag{7.16}$$

Because an individual's memory and disease information vary over time, PDI can change due to memory fading of prior information and continuous new information updates. Therefore, final disease information $FDI_i(t)$ is not only the sum of all new information acquired by an individual *i* at time *t*, but it is also affected by faded memory of prior information. Considering the

exponential smoothing memory fading method in equation (7.9), final disease information $FDI_i(t)$ in the MRFC model can be defined as

$$FDI_i(t) = w_0 PDI_i(t) + (1 - w_0) FDI_i(t-1) \tag{7.17}$$

The entire process of disease information is shown in Figure 7.5.

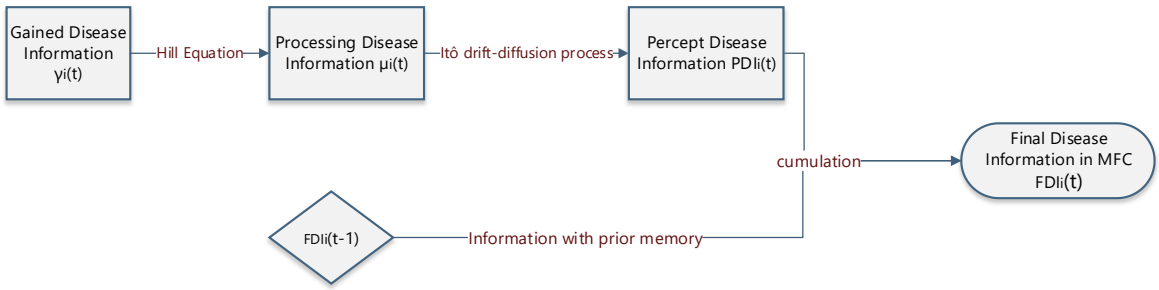


Figure 7. 5 Flowchart of disease information in the MRFC model

7.4 Agent-Based Modeling

7.4.1 Agent-Based Modeling

This section discusses agent-based modeling simulation of epidemic transmission to determine if IFC and MRFC models are effective in the real world. Because individuals have unique memories, moods, and behaviors, the unit of simulation must be individual, hence the use of the agent-based model. Agents are typically categorized into four types: switch susceptible, normal susceptible, infected, and recovery. An infected agent can contaminate a nearby switch/normal susceptible agent with a varying infection rate based on the switching behavior of susceptible agents. Switch susceptible agents have lower infection rates than normal susceptible agents because they choose behaviors to protect themselves from disease. An infected agent has a probability of recovering after reaching the recovery period, and an infected agent becomes a recovery agent after completing the recovery process (as shown in Figure 7.6).

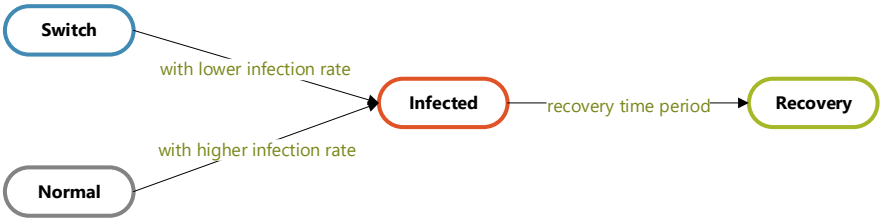


Figure 7. 6Flowchart of disease transmission

The agent-based epidemic model in this research references the basic susceptible, Infectious and Recovered (SIR) model framework built by Kermack-McKendrick [47]. Kermack-McKendrick theory assumes no births, deaths, or travel into or out of the population and that every agent has an equal chance of contacting with any other agent. In terms of an epidemic virus, the model assumes no dormant and latent periods in the disease. Viral mutation is not considered.

The agent-based simulation model primarily defines information reception, memory fading, and behavior switching through the IFC and MRFC models. Upon model initialization all agents are randomly arranged spatially in two dimensional (2-D) simulation space. When the simulation begins agents randomly move to nearby areas or remain at their current location. Disease information is calculated based on the prevalence and switch-behavior information at each time epoch following agent movement, and then susceptible agents reconsider whether or not to choose switching behavior based on the updated $PDI(t)$ calculated by the IFC or MRFC models. Infected agents can infect nearby susceptible individuals at a certain infection rate. The simulation terminates when all infected agents reach recovery, signaling the end of the epidemic. Otherwise simulation continues until the given maximum simulation time.

Agent-based models have gradually become mainstream due to rapid advancements in hardware and software computing power. Several agent-based simulation and modeling environments, such as Swarm, Mason, and NetLogo, were developed to help researchers study detail behaviors of their models. Netlogo is the most researcher-recommended software because

of the friendly programming interface and ease to code [67]. Netlogo software also includes a library with a large amount of example models. Based on these advantages, this research utilized Netlogo as the simulation platform.

The simulation framework in this paper is based on the epiDEM framework, an existing example model in the Netlogo library [303]. The simulation is innovative because it defines switch-susceptible agents, which were not considered in the original Netlogo library Moreover, the IFC and MRFC models are embedded into the model to determine if susceptible agents will switch to a protective behavior. This model also highlights susceptible population average fear factors that can be used to describe psychological emotion variations during epidemics.

Simulation results reported total switch populations and accumulated infections results. The initial testing simulation model was set in a 51×51 2-D grid. Each agent could randomly move to the grid nearby or stay at the current position, and each susceptible agent could be infected if and only if exist infections existed in his/her surrounding grids, referred to as the DTCN. In addition, the agents were assumed to receive disease information (infection rate and switch rate) from the entire simulation map, referred to as the ITCN. The recovery time for each agent followed a normal distribution (assume $\mu = 30$, $\sigma = 7.5$), and following the recovery time, each infected individual became a recovered individual and was not infected nor reinfected [304]. This simulation also assumed that the infected rates for normal and switched-susceptible agents are 15% and 5% to clearly embody effects of the protective measures.

The simulation contained an initial population of 1000 agents randomly placed on the simulation map. Among those agents, 5% was randomly selected to be infected individuals based on the binomial distribution and 10% was randomly selected to be the switched population based on binomial distribution. No overlapping occurred between switched and infected agents. The

white agent in the graphical user interface (GUI) represented non-switched susceptible agents, purple agents represented switched susceptible agents, red agents (also orange and pink agents in simulation of Section 7.5) represented infections, and green agents represented recovered agents.

Simulations were run on a workstation equipped with an Intel-based central processing unit i7-6700K and 32GB memory (RAM). System parameters such as infection rates or specific model parameters could be set manually prior to the simulation runs. The primary objective of the simulation was to monitor the trends or changes of crucial system variables (e.g., switched population, infected population, average fear factor).

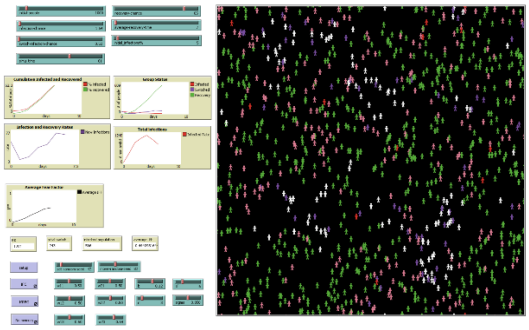


Figure 7. 7Netlogo model and simulation GUI

7.4.2 Sensitivity Analysis

Although Sections 7.2 and 7.3 of this paper detail IFC and MRFC models and Section 4 presents Netlogo models and simulation setups to study model behaviors, the significant influences of parameters such as infected information weight α, were not yet discussed. This section utilizes the OFAT method to analyze the sensitivity by changing the setting of each parameter in the IFC and MRFC models. The main goal is to analysis model behaviors caused by altered parameter settings. In addition, parameter ranges also reflect the limit application range of crucial variables. The application range can be used to determine model flexibility.

Three fixed random seed values were used in each simulation to eliminate effects from the random process and ensure the presence of only one variable in each sensitive analysis simulation. Figures 7.8 and 7.9 show the average number of three example simulation results (with common random seeds). The common random seeds setting ensures the analyzed parameter is the only variable.

The IFC model contains three crucial parameters: infected information weight α, behavior switch information weight β (equation (7.4)), and memory performance proportion w_t (equation (7.5)), which is determined by the rate of forgetting b and longest memory epoch t_f (equation (7.10)). Therefore, this section considers parameters α, β, b, and t_f (set as variables of information reception and forgetting process in the Netlogo simulation environment) by changing the value of parameters in their corresponding reasonable range and then analyzing the varying tendencies of infected rate and a total population of switch.

Sensitivity analysis results of the IFC model are shown in Figure 7.8. The red and blue lines illustrate variations of cumulated infections and total switch population, respectively, corresponding to parameter setting variations. As shown, cumulated infections decreased and switch population increased with the growth of α, β, and b. In addition, the variation range resulting from changing β was more pronounced than the variation ranged when α was changed, proving that β is more sensitive to total switch population than α. For the longest memory epoch t_f, however, the t_f cause the lower total switched population and higher infections. In general, $\alpha, \beta,$ and t_f are sensitive, but b is not sensitive.

The MRFC model contains six parameters: infected information weight α, behavior switch information weight β (equation (7.4)), the power of Hill equation n, equilibrium constant K (equation (7.12)), forgetting process constant ε, and maximum random range σ (equation (7.14)).

In addition to α and β, n and σ are also advantageous for sensitivity analysis because n determines the information reception level and σ determines the randomness level. However, K and ε identify the learning and forgetting process, which means K and ε should be fixed at the simulation initiation.

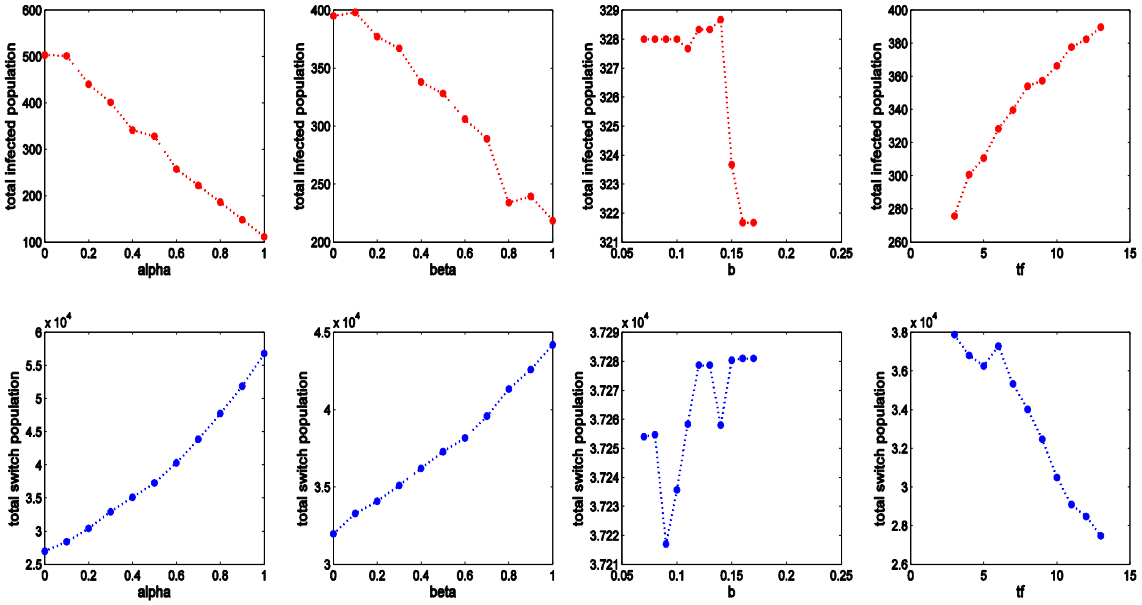


Figure 7. 8 IFC model sensitivity analysis

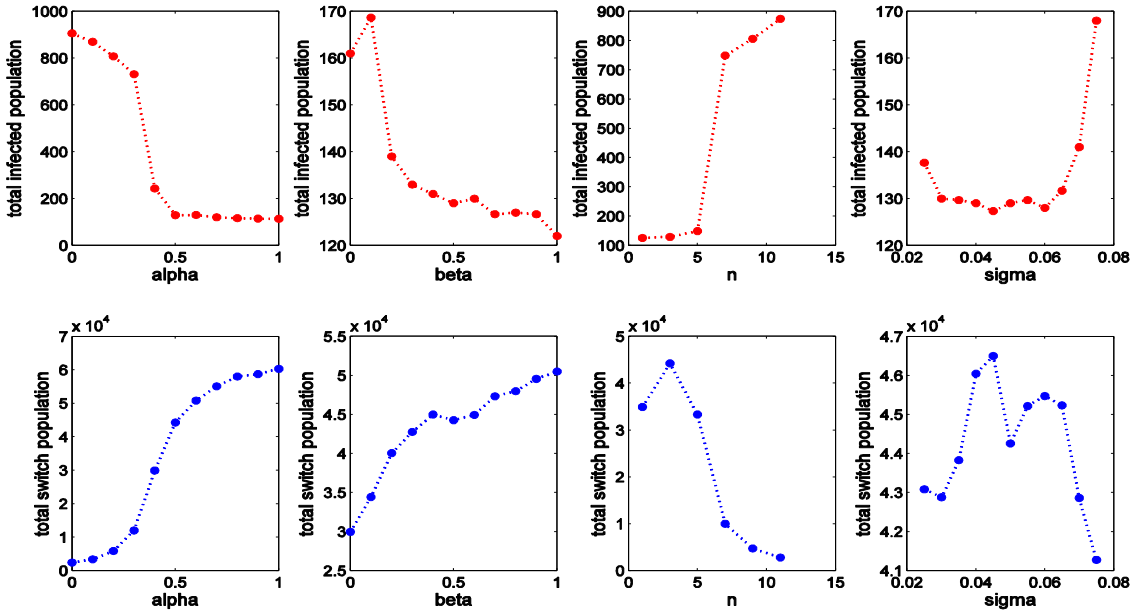


Figure 7. 9 MRFC model sensitivity analysis

MRFC model sensitivity analysis results for the total infected population and the total switched population are shown in Figure 7.9. Similar to the IFC model, cumulated infections tended to decrease and switch population increased with the growth of α and β. However, the MRFC model was more sensitive to α than β, contrary to results from the IFC model. For the parameter n, cumulated infections and switch population were very sensitive when $n \in [3, 7]$. For the variation parameter (σ), the cumulated infections and switch populations became more random as σ increased, resulting in a population diversity parameter σ below 0.6, which was not a significant parameter for system sensitivity.

In order to determine whether the IFC or MRFC model is more sensitive to parameter changes, this study compared sensitivity analysis results of the models shown in Figure 7.8 and 7.9. Comparison showed that increasing α and β resulted in a decrease of total infections and total switching populations in both models. In addition, parameter α demonstrated greater sensitivity than β in the models. Table 4.1 presents the covering range of total infected and total switched populations when common parameters α and β were changed in both models. Sensitivity levels of the MRFC model were generally higher than the IFC model, the only exception is the total switch population changing by β. The infected population in the MRFC model varied from 11.3% to 90.6% (95% confidence interval [6.5%, 93.9%]), and the switched population ranged from 2.4 to 60.3 thousands (95% confidence interval [1.2, 79.7] thousands). Infected population variation was much smaller for the IFC model, ranging between 11.1% and 52.2% (95% confidence interval [8.2%, 67.7%]). Similarly, the switched population varied only between 26.9 and 56.8 thousands (95% confidence interval [16.2, 68.3] thousands). The MRFC model demonstrated less robust to the parameter changes than the IFC model when modeling general epidemics because the model

is more sensitive to parameter settings. Therefore, the MRFC model could be used to model epidemics with high variations of infected and switched population ranges.

Table 7. 1 Comparison of model flexibilities based on sensitivity analysis data

	Range of infections by changing α	Range of infections by changing β	Total switch population range by changing α	Total switch population range by changing β
IFC model	11.1%-52.2%	21.8%-41.4%	26.9-56.8 thousands	31.9-44.2 thousands
95% confidence interval	8.2%-67.7%	4.2%-63.4%	16.2-68.3 thousands	20.2-53.4 thousands
MRFC model	11.3%-90.6%	12.2%-16.8%	2.4-60.3 thousands	29.9-50.5 thousands
95% confidence interval	6.5%-93.9%	9.4%-20.8%	1.2-79.7 thousands	17.1-61.6 thousands

The second part of sensitivity analysis focuses on infection track variations when crucial parameters are changed. Using OFAT sensitivity analysis, the number of infections was inversely correlated to the number of switch populations, eliminating the need to track percentages of both the infected population and the switched population. Therefore, this portion of analysis focused on the percentage of infection population, which more accurately reflects the severity level of epidemics. Simulation comparison is shown in Figure 7.10. For the IFC model, all shapes of tracks with different parameter settings were similar. The total infected population throughout the epidemic decreased when α and β increased, thereby corresponding to OFAT sensitivity analysis results. In comparison, the tracks of infection population percentage with various parameter settings demonstrated multiple shapes in the MRFC model. Parameter α was the key to controlling the midterm (approximately Day 25) epidemic performance of the model. Therefore, if agents focus more on infection information rather than switch behavior information (with higher α), total infections will decrease. However, parameter β significantly determines epidemic performance of late periods (approximately Days 25–50)., meaning that increased attention to local disease information rather than global information will decrease total infections in the second half of the simulation. The second part of the sensitivity analysis proved that the

MRFC model has greater flexibility than the IFC model in the range of infections and switched populations and in the shape variations in infection population percentage tracks.

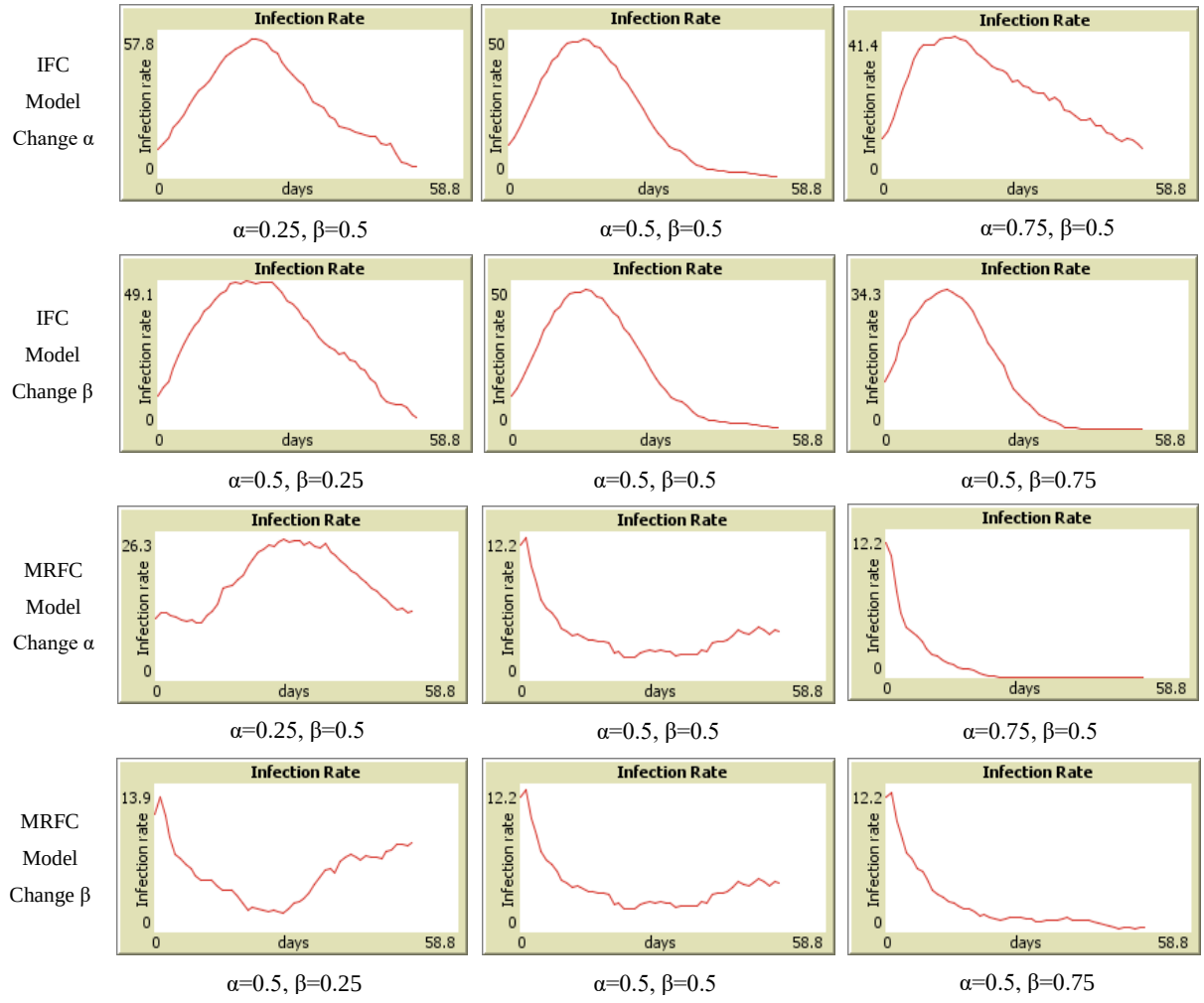


Figure 7. 10 Sensitivity analysis of infection population percentage tracks for IFC and MRFC models

7.4.3 Simulation Comparisons

This section describes simulation runs to compare IFC and MRFC model performances, including use of a no memory model as the baseline in each comparison. Populations of each type of agent and percentages of switched population were considered for each simulation run. Common random seed and other parameters ($\alpha=\beta=0.5$) were used to ensure that variations

occurred due to different modeling methods only. Simulation run times were set to 70 days, although in most cases in our simulations, the epidemic ended prior to that limit. All populations, percentages of infected/switched populations, and average fear factors were compared, as illustrated in Figure 7.11. Populations contain the number of infected, susceptible, switched, and recovery agents, and infected/switched rates are defined as the infections/switch-susceptible populations divided by the total non-cured population. The average fear factor represents the average number of fear factors for the entire susceptible population.

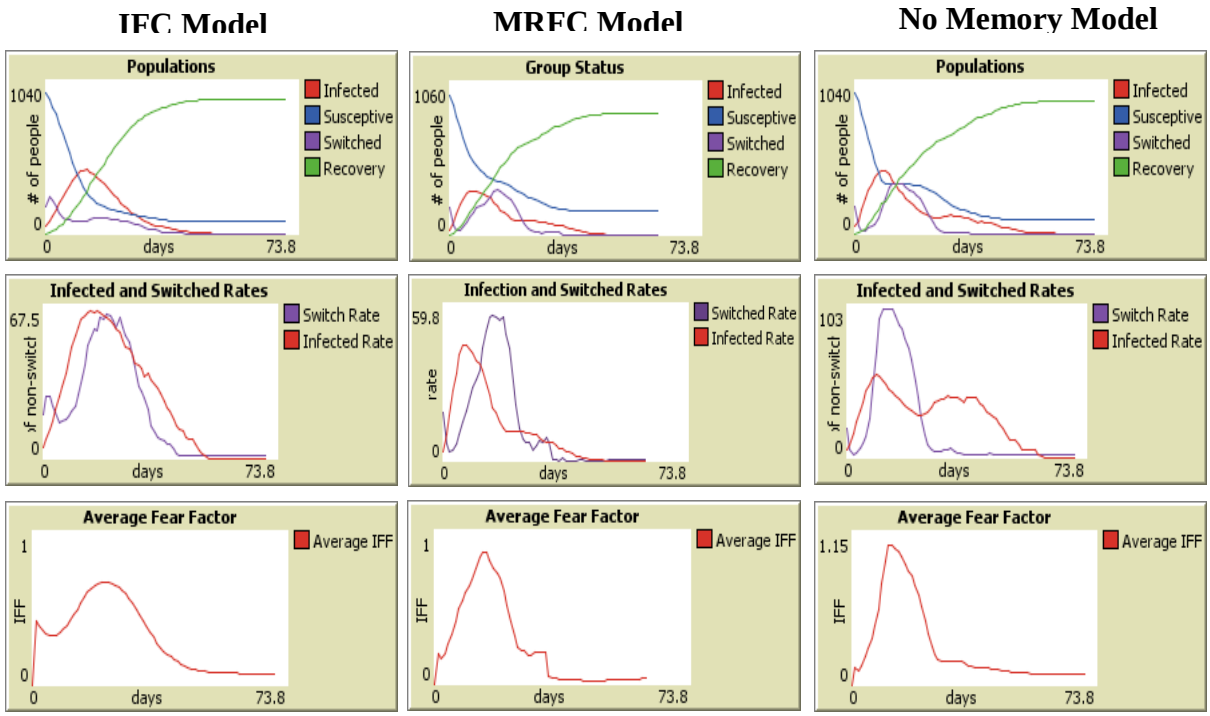


Figure 7. 11 Comparison of IFC, MRFC, and no-memory models

As shown in Figure 7.11, infected individuals in the IFC model totaled 914, 847 in the MRFC model, and 902 in the no-memory model. The total switched population in the IFC model was 3299, 3978 for the MRFC model, and 4642 for the no-memory model. In general, the IFC and MRFC models demonstrated similar performances, although the MRFC model showed relatively more switched-susceptible agents in the mid stage of the epidemic (10–20 days).

Because in the MRFC model, the process to gain information can immediately detect epidemic variation information. Therefore, the MRFC model can be applied to highly lethal or infectious epidemics, such as the SARS epidemic in 2003. Agents typically pay close attention to new epidemic developments, thereby causing rapid, elevated fear emotions in the population and motivating a majority of individuals to take protective measures even if only a minimal number of infections has been reported.

Both the IFC model and the no-memory model showed relatively lower switched behavior than the MRFC model, making them better-suited to model epidemics with higher numbers of infectious and minimal lethal situations, such as an influenza epidemic, in which people frequently pay less attention to new disease developments because the risk of death or other adverse consequences are relatively low. Although overall performances of the IFC model and no-memory model were similar, the tendencies of infected population percentages differed significantly. The switched population in the IFC model followed the decline of the infected population approximately 30 days after the onset of the epidemic. However, in the no-memory model, the infected population demonstrated a second peak around 40 days due to the quick decline of the switched population. Because agents in the IFC model retain memory of past epidemic information, they maintained their switching behavior for an extended period even as the epidemic neared completion. Nevertheless, the susceptible population in the no-memory model does not retain past disease information, so a majority of individuals switched behavior back to normal immediately, allowing the epidemic to spread.

7.5.1 H1N1 Case Study

7.5.1 H1N1 Pandemic in 2009

H1N1, also known as swine flu, is a form of influenza in pigs that can be transmitted to people by exposure to infected droplets [305]. H1N1 is an orthomyxovirus, a subtype of influenza A that is the most common cause of seasonal human flu. H1N1 viruses attack the human immune system, attaching and replicating within infected cells. A person infected by the H1N1 virus will develop a progressive lower respiratory tract disease that could result in respiratory failure [306]. The first human case of H1N1 virus was reported in Mexico in 2009, quickly spreading to the United States and the world and resulting in a pandemic outbreak [307]. Determination of H1N1 infection is difficult based on symptoms because influenza symptoms are nonspecific, typically lasting four to six days with an effective infection period continuing for approximately seven days.

Prior to the outbreak of the H1N1 virus pandemic in 2009, minimal information was available about the disease and people had limited or no disease awareness or past memory. Immediately following the initial disease outbreak, people began to understand and focus on the new strain of influenza A. Therefore, in order to more accurately describe how people learned about and forgot H1N1 virus pandemic information, this research utilized data from the pandemic from several days (5 days) after the outbreak since people began to have memories about the disease after that time.

Many reports documented the H1N1 virus pandemic in 2009 [308]. The first graph in Figure 7.12 illustrates pandemic H1N1 virus infection among New York City residents hospitalized from May 29 to July 1, 2009. At the preliminary stage of the pandemic, people had minimal memories of the disease and paid limited attention to it, resulting in an increase in infection cases. After several days of infection, people developed increased awareness and

memory about the disease and began to take preventive measures; at the end of June 2009 the number of infected cases gradually declined. Pandemic prevalence showed a similar trend in Chicago from June 7 to July 10, 2009, and in Shanghai, China, from June 29 to July 29, 2009, the number of infection cases initially increased and then decreased during July. Mexico City showed a similar trend from April 17 to May 17, 2009. This research used the graphs in Figure 7.12 to determine whether the learning and forgetting phenomenon affected the pandemic trend in the four cities.

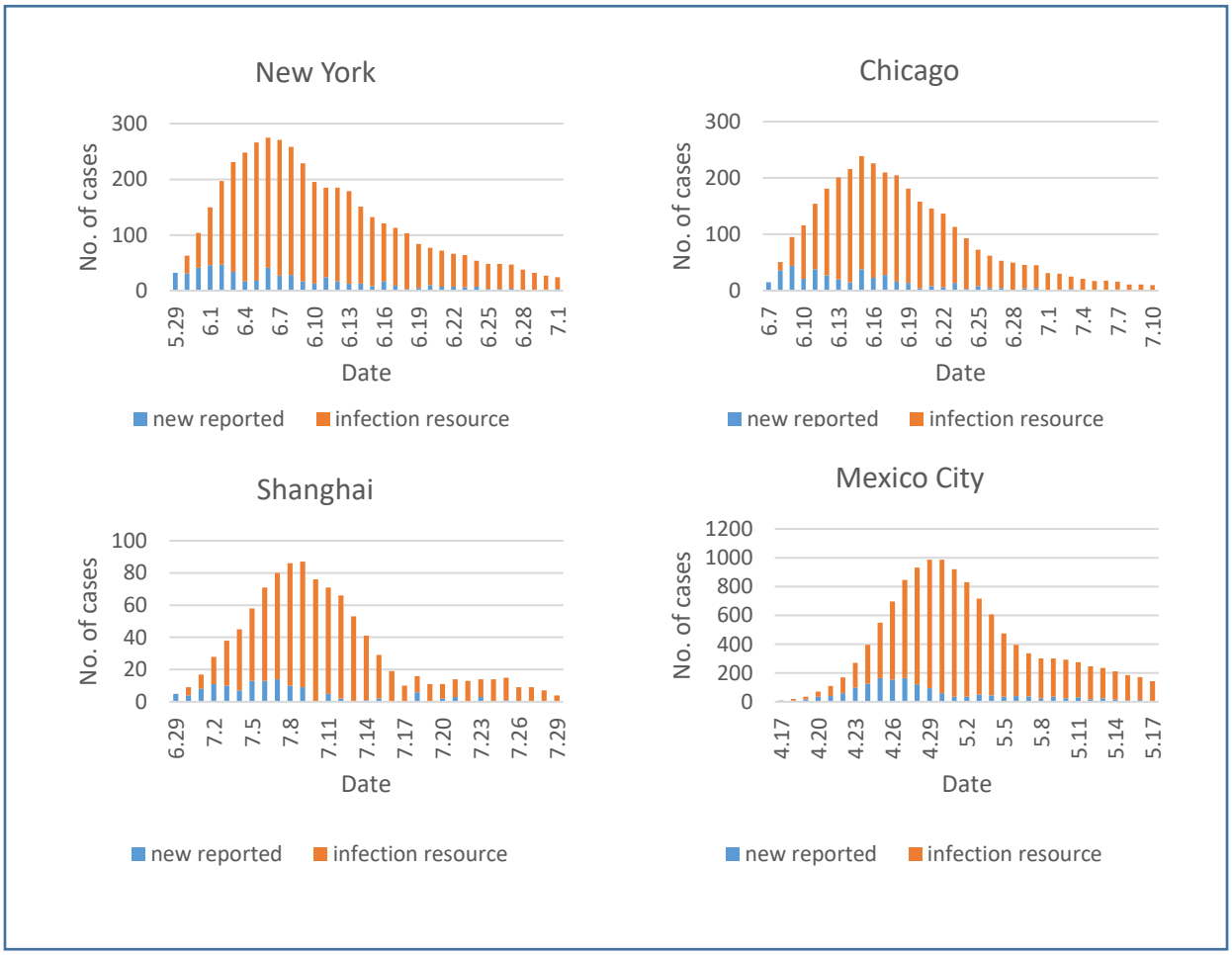


Figure 7. 12 H1N1 infection data in 2009

Table 7. 2 Infection characteristic percentages by cities

Regions	Age Group of Infection Percentage				Urban/Non-Urban Percentage	
	Infant aged 0-4	Student aged 5-19	Adult aged 20-64	Elder aged >64	Urban	Rural
New York	6.7%[54]	20%[54]	58.8%[54]	14.5%[54]	87.9[55]	12.1[55]
Shanghai	16%[56]	51%[56]	30%[56]	3%[56]	59[57]	41[57]
Chicago	3.8%[58]	34.2%[58]	60%[58]	2%[58]	89.3[59]	10.7[59]
Mexico City	16%[60]	20%[60]	56%[60]	8%[60]	78.8[61]	21.2[61]

Table 7.2 shows the proportional distribution of infection cases based on age group and total population in four cities. From May 1 to June 30, 2009, the highest infection rate occurred in students 5–19 years old and adults 20–64 years old. The infection rate for students was approximately three times higher than the infection rate for infants under 4 years old, with the exception of Mexico City. The numbers of infected cases in urban areas were significantly higher than non-urban areas.

In order to determine if receiving and forgetting information during an epidemic can influence the spread of disease, this paper assumes that people can both learn new information and forget prior information over time. The IFC and MRFC models were applied to simulate an outbreak of the H1N1 virus pandemic in 2009 for the Chicago using report data to determine if the IFC model or MRFC model can better reproduce infection population trends throughout the pandemic.

7.5.2 Infection Rate Calculation with Historical Epidemic Data

Infection rate is the probability of a susceptible agent to be infected at time *t*. It can be defined mathematically as the new infected population divided by the susceptible population [36]:

$$\text{Infection rate} = \frac{new\ infections\ (t)}{susceptible\ agents\ in\ risky\ (t)} * K \tag{7.18}$$

In equation (7.18), *new infections(t)* represents infections caused by the original infection after last time epoch, and K is a constant used to adjust the infection rate within different time periods. For example, if the infection rate is calculated based on a one-day period, the infection rate in 10 days can be determined by assigning a value of 10 to K. *Susceptible agents in risky(t)* have face-to-face contact with the infected population in their DTCN. Even if the data of the original infected population are known, however, accurate recording of the number of potential contacts for each infected agent is nearly impossible.

This section uses three methods to estimate the average number of contacts for an infected individual. The first method, the average-contact-based method, utilizes the estimated average number of contacts from the report. Mossong *et.al* recorded physical contact behavior for 7290 participants from eight countries with various age ranges [309]. The average number of contacts for these participants was 13.4 per day (standard deviation was 10.6).

The second method, the age-ranges-based method, is based on the average number of contacts from various age ranges. Valle *et.al* utilized U.S. Census Bureau data from 2000 to analyze contacts per person [310]. They found that the adult group (between 20 and 60 years old) had the highest number of contacts (approximately 20) per day. Children and elderly groups had the lowest number of contacts (approximately 10). Therefore, the susceptible population was defined as $\sum_i I_i(t)c_i$, where i is the corresponding age group *I*, $I_i(t)$ represents how many infected individuals are present in age group i in time t, and c_i represents the average number of contacts in age group i.

The third method, the urban-population-percentage-based method, considers the average number of contacts in areas with various population densities. Read *et.al* researched the relationship between daily contacts and population densities [311]. They found that urban citizens

have a higher number of contacts than rural citizens. This method most accurately reflects the epidemic performance difference between rural and urban populations.

Infection rates were calculated at each time epoch in each city using these three methods. Table 7.3 presents the average numbers with 95% confidence intervals of the infection rates. Among the four cities, Shanghai had the highest infection rate and Chicago had the lowest infection rate, a difference that could be attributed to the areas' different population densities and epidemic outbreak locations. However, the infection rate of H1N1 in 2009 remained approximately 1%-2%. This study used the average-contact-based method to calculate the infection rate in the case simulation.

Table 7. 3 Calculated infected rates using three methods in four cities

Regions	Average-contact-based method	Age-ranges-based method	Urban-population percentage-based method
New York	1.30% [1.05%-1.54%]	1.08% [0.87%-1.28%]	1.22% [0.99%-1.45%]
Shanghai	2.39% [1.71%-3.05%]	1.88% [1.36%-2.41%]	2.22% [1.60%-2.85%]
Chicago	1.08% [0.91%-1.25%]	1.01% [0.85%-1.16%]	1.16% [0.98%-1.34%]
Mexico City	1.90% [1.12%-2.69%]	1.59% [0.93%-2.24%]	1.86% [1.09%-2.62%]

This research also investigated whether behavior switching reduces the chances for infection. Unfortunately, no data reported how many agents protected themselves from H1N1 infection using switch behavior such as protective masks, vaccinations, or staying at home. To most accurately determine the H1N1 infection rate for behavior switched population, therefore, this research used existing reference data showing switch behavior effectiveness to be 50%–86% [312-314]. This case study used the median of the range as the effectiveness of the switched behavior, meaning the infection rate of behavior-switched agents was approximately 33% of the normal infection rate.

7.5.3 Simulation of 2009 Chicago H1N1 Case

As mentioned, Chicago was one of the cities which severely afflicted by H1N1 in the 2009 virus pandemic. This mega city has an approximate population of 2.705 million people (third largest in the United States) and a population density of 4582.3 people per km^2 (fourth larges in the United States) [315]. Chicago has 588.26 km^2 of the land area, which has a strip-shaped, but the northeast portion of Chicago has the highest population density. The fast-paced urban lifestyle and highly concentrated population in this area contributed to the extensive H1N1 outbreak. This section uses Chicago as a special H1N1 case to conduct epidemic simulation using historical data and analyzing the fear emotion with learning/forgetting behavior.

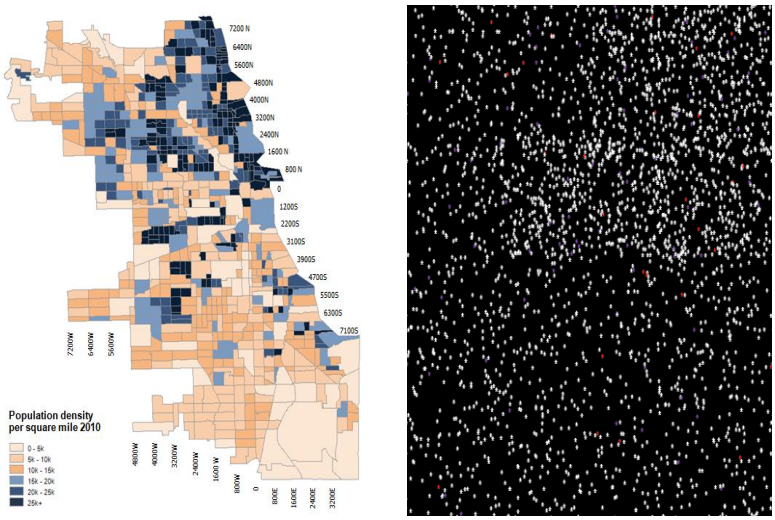


Figure 7. 13 Comparison of population density map with simulation initial setting in Chicago

The Chicago H1N1 virus epidemic simulation was set in a 60 * 120 grid 2-D space with each grid representing 0.2858 * 0.2858 = 0.0817km^2. A total of 2,705 agents were located in the simulation space, and each agent represented a group of 1000 individuals. Since June 7, 2009, there were 25 H1N1 reported infections, or the initial infected agents (red) were 1% of the total

agent groups; each infected group had one infected person. Initial switched-behavior agents were assumed to comprise approximately 3% of the total population. All infected, switched, and susceptible agents were randomly placed according to population densities in Chicago, as shown in Figure 7.13.

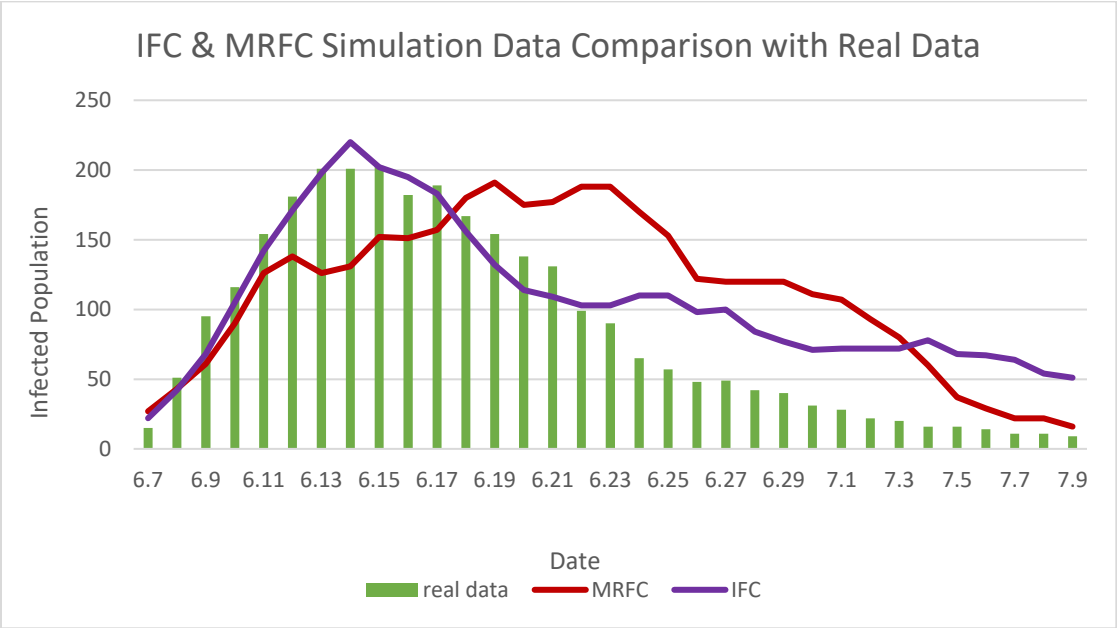


Figure 7. 14 Simulation results comparison of the IFC and MRFC models

Simulations runs were set up with identical initial settings and random seeds, but with two models (the IFC and MRFC models). Simulation results related to infection source, average fear factor, and switch groups in the two models are shown in Figure 7.14. The simulated time period was June 7–10, 2009 (33 days). In the IFC model simulation, although the total number of infections (492) was slightly higher than the real situation (410), the tendency of the epidemic was very similar to the historical data. In addition, the average fear factor (not shown in Figure 7.14 due to space limitations) in the susceptible population and the number of individuals in the switched-behavior group (Figure 7.15) were similar to the prevalence data reported in 2009. However, simulation results of the MRFC model did not correlate well with the historical data.

The peak of the H1N1 outbreak occurred around June 25 in the MRFC simulation, when the real-world epidemic was nearly finished, and the switched-group number was nearly half of the total groups (Figure 7.15) at the beginning of the prevalence, which is not realistic. In conclusion, results of the IFC model simulation runs fits better to the 2009 Chicago H1N1 historical data.

Figure 7.15 uses GUI to show epidemic tendency in the IFC and MRFC models. In the interface, red, orange, and pink agents represent the infected group with 1, 2, and 3 infected people, respectively, white agents represent normal-behavior susceptible agents, violet agents are switch-behavior susceptible agents, and green agent groups contain recovered people.

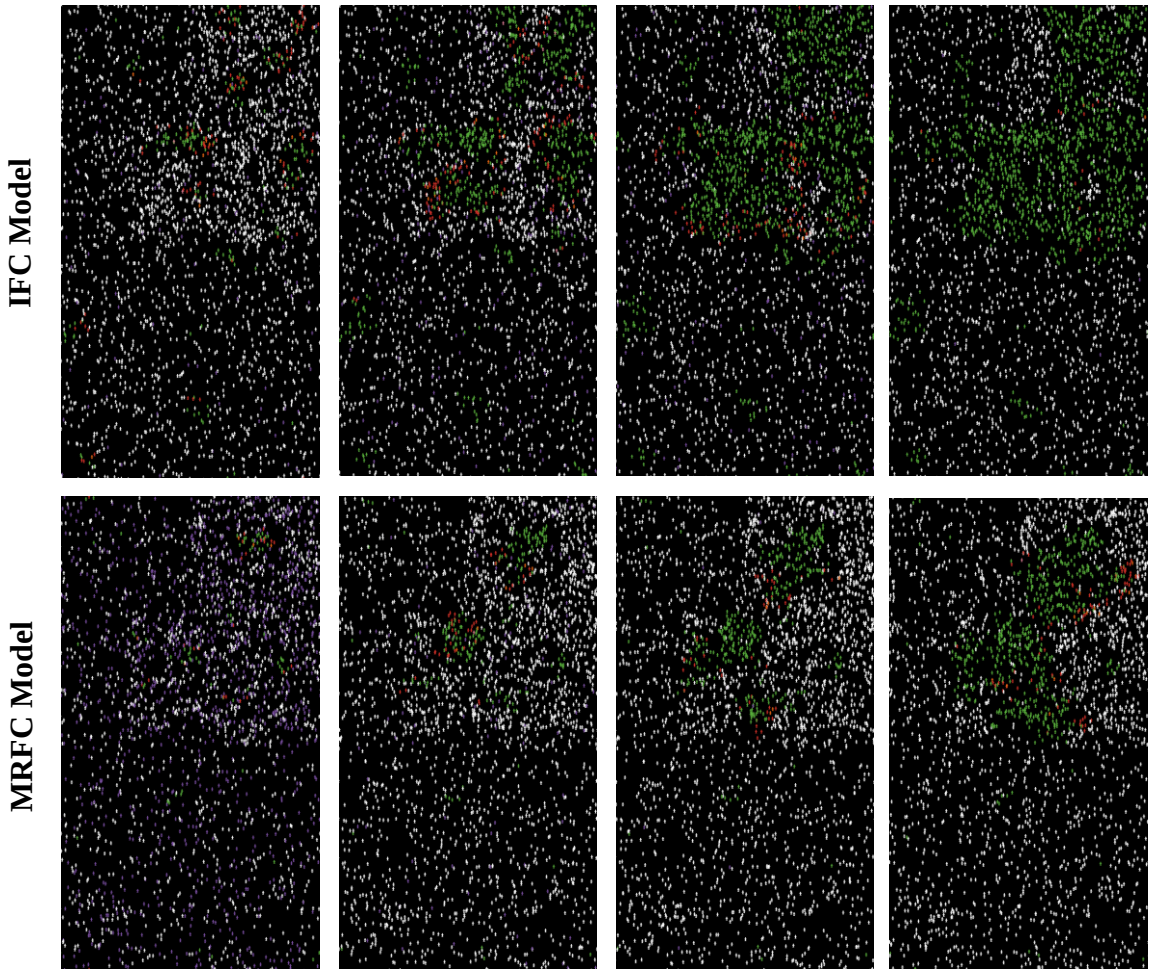


Figure 7. 15 IFC and MRFC model topographic chart for 2009 H1N1 in Chicago

On Day 6 of IFC model simulation runs only a few agent groups in the north part of the city were infected, and almost no agents switched their behavior to prevent infection. On Day 14 the epidemic area expanded into the entire downtown and northern areas of the city. With the increasing number of newly infected individuals, more and more agents gained awareness of the H1N1 virus epidemic from the information learning process. The number of switched-agent groups also grew significantly due to increasing fear of the H1N1 virus. On Day 22 the infected population was primarily located at the edge of highly populated areas, and the number of infected individuals began to decrease from the overall peak. Since memory fading requires several days (usually around 5 days in our simulation), the number of switched groups did not show significant decline. The prevalence nearly ended on Day 30 when the number of infected agents approached zero. Very few agent groups (around 10 percent) decided to switch their behavior, meaning that most agents had already forgotten that the H1N1 virus is a highly pathogenic influenza.

Results from the MRFC model simulation showed that most of the susceptible population decided to switch their behavior at the beginning of the epidemic (around Day 6) due to fears caused by sudden disease prevalence, thereby preventing the early epidemic diffusion. In the metaphase of the epidemic fewer people protected themselves by switching their behavior even though there was no strong signal of prevalence deterioration. The result of returning to normal behavior caused the epidemic outbreak to become more serious near the end of the simulation. However, simulation results from the MRFC model may not actually reflect characteristics of 2009's H1N1 virus scenario in Chicago in regards to epidemic infection tracks and behavior patterns of susceptible agents.

In general, the IFC model accurately recreated the complete process of Chicago's H1N1 epidemic in 2009. The time zone and infection tendency from this simulation resembled real historical data. Prevalence began in the northeast portion of the city and spread through downtown. Suburb areas were not influenced significantly from the H1N1 virus. The average fear factor and number of switched agents also followed the tendency of epidemics with a short delay (2–3 days), potentially indicating the time needed to process new disease information and allow for memory fading.

7.6 Summary

This paper investigated the assumption that disease information can influence an individual's fear emotion and that agents' emotions potentially affect behavior during an epidemic. This study used two mathematical models (IFC and MRFC models) to discuss disease information fading and learning processes. Both models synthesized disease information on local and global levels with infection information and switched-behavior information, thereby providing comprehensive disease information to agents. However, the performances of the two models are different in modeling techniques, sensitivity levels and simulation performances (show in Table 7.4).

Table 7. 4 Comparison of model characteristics

	IFC Model	MRFC Model
Information source	ITCN (local) and DTCN (global)	ITCN (local) and DTCN (global)
Information types	Infection and behavior	Infection and behavior
Information cumulative process	Learning and forgetting	Learning and forgetting
Randomness process	No randomness	Randomness in Itô stochastic process
Information completeness	Based on complete disease information for each agent	Only based on Δinformation at each time epoch
Information transformation	No information transformation	Information transformation by hill equation

Sensitivity and model flexibility	Relative normal sensitivity and flexibility	Relative higher sensitivity and flexibility
Relationship between infections and switched behavior	Relative higher relationship	Relative normal relationship
Real data simulation performance	Fit the tracks of epidemics	Obviously different from historical data
Potential scope of application	Highly infectious and low lethal epidemics	Highly lethal epidemics

Modeling technique is the primary difference between the IFC and MRFC models. The IFC model assumes agents can obtain a complete picture of the epidemic via information from local daily contacts or global news coverage. This prevalence information affects the protective behavior of agents by changing their fear emotion level. However, the MRFC model assumes agents can detect discrepancies in disease information. The Hill equation transformation showed that agents usually ignore minor discrepancies and pay attention to major inconsistencies. Cumulative information transformation to knowledge was modeled mathematically by an Itô stochastic diffusion process.

The differences of modeling methods reduced the significant imparity in simulation results. The IFC model more accurately describes epidemics with high infectious ability and low lethality. Although the IFC model has less sensitivity and flexibility than the MRFC model, it more precisely restored the tracks of 2009's H1N1 virus epidemic in Chicago. Therefore, the MRFC model should be applied for highly lethal epidemics. Moreover, the simulation results show a weak relationship between infections and switched behavior in the MRFC model, which results that did not correlate to current data sets.

In conclusion, a dynamic agent-based model can be used to mimic real-world epidemic situations and explain disease transmission, behavior changes, and distribution of prevalence panic. Moreover, agent-based simulation with real data restored the historical H1N1 virus

influenza data from Chicago in 2009. Therefore, the future work of this research should explore more in the areas of the potential applications of both agent-based epidemic model and forgetting and learning model in following areas: (1) Health Organizations and Disease Control Centers can utilize the model presented in this research to evaluate and experiment the possible impacts and influences using various control strategies. For example, the agent-based epidemic model can be used to forecast the effects of information dissemination through media and broadcast. The model presented in this research can also reflect the effectiveness of public health education regarding the underlining disease in the researched area. (2) Our model can also be applied to examine the self-protection ability of the general public during a spontaneous or unannounced epidemic. By simulation experiments aimed to the specific area, the disease outbreaks in a localized area can be detected and subsequently mitigated by the public health agencies. (3) The forgetting and learning model presented in this research can be applied to other areas of application to reflect the diversified opinions or spontaneous behaviors within a heterogeneous population. For example, *the* investor's diversified perspectives and investing patterns on the stock market or the popularity or the electability of a public figure in the commercial or political campaigns can be modeled using the forgetting and learning curve presented in this research."

Chapter 8 - Conclusion, Contribution and Future Research

8.1 Conclusions

This dissertation introduces a new mathematical modeling based epidemic control method. This method starts from the mathematical modeling for different types of diseases, which have different transmission routes and characters. The system and parameter analyses provide the potential control variables for the epidemic model. Several numerical control methods are proposed by using nonlinear programming, heuristic algorithms, and machine learning. The resulting analytical and numerical simulations help us to validate the effectiveness of the control strategies and provide the information to the policymakers.

Main conclusions drawn from this dissertation are:

1, From the historically reported data, the ZVL has shown a significate difference in the distribution of human infections for each age group. The children and teenagers have more possibilities to be infected by ZVL. This dissertation summarizes the six historical ZVL outbreaks from different countries. By using kernel density estimation to process the data, we find that the children have 5 times higher possibility to be infected than others. This finding is also modeled by a differential equation and simulated.

2, The future ZVL epidemic of eight most severely affected countries are predicted by the model with the real-world data. The prediction result shows that the prevalence will relieve in Asian countries like India, Bangladesh, and Iraq; the prevalence in Brazil will maintain at the current level in the next couple years; and the prevalence in the East Africa countries like Sudan, South Sudan and Ethiopia are expected to deteriorate.

3, By comparing the changes in R_0 for ZVL, this dissertation finds that average biting rate to dogs (b_{fd}), the death rate of dogs (μ_d), the death rate of sand flies (μ_f) are more sensitive

than other parameters. Meaning, changing the values of these three parameters is the most efficient way to limit the ZVL transmission. Therefore, protecting the dogs by insecticide collars, dog treatments and spraying the insecticide to sand flies are suggested. Culling the dogs may have side effects to control the ZVL transmission.

4, Considering the Zika Virus epidemic when we ignore the cost of the control strategies. Using the Wolbachia to infect the susceptible mosquitoes is the best control method to cut down the transmission route between mosquitos to human. Spraying the insecticide can only reduce about 45% of human infections than Wolbachia. The influence of controlling human sexual transmission is negligible.

5, When considering the cost of the control strategies for Zika Virus, this dissertation suggests to spray the insecticide to minimize the mosquito population at the early stage of the epidemic. During the middle stage, Wolbachia should be used to lower down the infected rate of susceptible mosquitoes. Human sexual transmission control should be considered around the peak of human infections.

6, Disease epidemic mitigation and many social policy intervention systems unlike many real-time control systems in the mechanical and electrical engineering domain, the real-time feedback-control protocol is not only impractical but also cost-prohibitive or impossible. Instead, stage-wise step controls or impulsive control strategies are more practical and cost-effective for disease mitigation systems. Optimal control problems to determine the frequency and action time of the interventions are involved as the main goals.

8.2 Contributions

Major contributions of this dissertation to the area of epidemic system modeling, infectious disease optimal control and evidence epidemic data analysis are listed as follows:

1. This dissertation validates that dynamic systems combined with differential equation modeling can be used to model the transmission of different vector-borne diseases. These models can also be used to explain the underlying infections for other epidemics by changing the parameters and variables.

2. This dissertation designs control variables in time series. These control strategies act on the established epidemic models to limit the negative influence of the transmission of vector-borne diseases. Through the numerical simulation, the control strategies can be validated at a lower cost than the real-world experiments. In addition, the numerical simulations are duplicable by adjusting the control variables with different values.

3. For the first time, this dissertation introduces a partial differential equation model for vector-borne disease to model the age structure for human infections. This model has been used to explain the high proportion of ZVL infections in children and the teenage group.

4. This dissertation first introduces a heuristic-based optimal control method, which can solve the epidemic optimal control problem with the highly nonlinear objective function. Moreover, this method can revise the time length of the control steps. Therefore, the implementation costs of changing the control level can be minimized by this method.

5. This dissertation innovatively proposes an evidence-based control method. This method can take advantage of historical epidemic data. With the help of the analysis results of the evidence data, the new method will be able to design a suitable control strategy based on the current prevalence information.

6. This dissertation also presents several event-triggered methods to automatically and intelligently trigger epidemic control. The event-triggered methods are designed to

comprehensively analyze the current prevalence based on the collected information. This method is able to save the unnecessary waste of control at an inopportune time.

7, This dissertation introduces an agent-based model to describe all of the epidemic, emotion and disease information transmission. This model illustrates the disease information can affect human emotion; the emotion changes will response to the human behavior changes for each individual. This model also discusses the human forgetting and learning process during the epidemic information transmission.

8, This dissertation proposes a method to systematically analyze the R_0, equilibrium and bifurcation of the epidemic dynamic system. This method can be used to study the static state, stability and control efficiency for the given system. The analysis results will be benefited to design the appropriate strategies to control the vector-borne disease.

8.3 Future work

Major future works to the area of epidemic system modeling, infectious disease optimal control and evidence epidemic data analysis are listed as follows:

1. Currently, the most epidemic model utilizes the collected data as initials. The future work should develop data-corrected models, which can take advantage of real-time epidemic data to correct the model during the simulation.

2. In this dissertation, most of the control strategies are theoretical. The future work will fully study the epidemic control methods in the real world, to make controlling model to better reflect the reality.

3. Currently, the nonlinear optimization-based algorithm can only solve the optimal control problem with the convex objective function. The newly proposed heuristic and evidence-

based algorithms have a long-running time. Therefore, a faster and more comprehensive optimal control algorithm is expected in future work.

4, This dissertation only mentions the optimal control for the dynamic disease model. Future work should include agent-based optimization, which can develop optimal control strategies at the individual level.

5, This dissertation mentions how to use simulated data to train the neural networks, which can be used to predict the future epidemic strategy. However, that is not enough. Future work should more focus on the self-supervision and uncertainty quantification for evidence-based epidemic control.

References

1. Taylor, L.H., S.M. Latham, and M.E. Woolhouse, *Risk factors for human disease emergence.* Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 2001. **356**(1411): p. 983-989.
2. Papagrigorakis, M.J., et al., *DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the Plague of Athens.* International Journal of Infectious Diseases, 2006. **10**(3): p. 206-214.
3. Gottfried, R.S., *Black death*. 2010: Simon and Schuster.
4. Morens, D.M., G.K. Folkers, and A.S. Fauci, *Emerging infections: a perpetual challenge.* The Lancet infectious diseases, 2008. **8**(11): p. 710-719.
5. Ryan, J.R., *Pandemic influenza: emergency planning and community preparedness*. 2008: CRC Press.
6. Sankaranantham, M., *HIV–Is a cure possible?* Indian journal of sexually transmitted diseases and AIDS, 2019. **40**(1): p. 1.
7. Goudsblom, J., *Public health and the civilizing process.* The Milbank Quarterly, 1986. **64**(2): p. 161-188.
8. Majra, J. and A. Gur, *India needs a great sanitary awakening.* Indian journal of occupational and environmental medicine, 2008. **12**(3): p. 143-143.
9. Rosenkrantz, B.G., *Public health and the state: changing views in Massachusetts, 1842-1936*. 1972: Harvard University Press.
10. Schoenbaum, M.A. and W.T. Disney, *Modeling alternative mitigation strategies for a hypothetical outbreak of foot-and-mouth disease in the United States.* Preventive veterinary medicine, 2003. **58**(1-2): p. 25-52.
11. WHO. *World Health Organization. Leishmansis*, . 2017 [cited 2017 Oct 6th]; Available from: http://www.who.int/leishmaniasis/burden/en/.
12. Dasbach, E.J., E.H. Elbasha, and R.P. Insinga, *Mathematical models for predicting the epidemiologic and economic impact of vaccination against human papillomavirus infection and disease.* Epidemiologic reviews, 2006. **28**(1): p. 88-100.
13. Nianogo, R.A. and O.A. Arah, *Agent-based modeling of noncommunicable diseases: a systematic review.* American journal of public health, 2015. **105**(3): p. e20-e31.
14. An, L., *Modeling human decisions in coupled human and natural systems: Review of agent-based models.* Ecological Modelling, 2012. **229**: p. 25-36.
15. Real, L.A. and R. Biek, *Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes.* Journal of the Royal Society Interface, 2007. **4**(16): p. 935-948.
16. Peterson, A.T., et al., *Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2005. **99**(9): p. 647-655.
17. Holt, R.D. and J. Pickering, *Infectious disease and species coexistence: a model of Lotka-Volterra form.* The American Naturalist, 1985. **126**(2): p. 196-211.
18. Castillo-Chavez, C., et al., *Epidemiological models with age structure, proportionate mixing, and cross-immunity.* Journal of mathematical biology, 1989. **27**(3): p. 233-258.
19. Frias-Martinez, E., G. Williamson, and V. Frias-Martinez. *An agent-based model of epidemic spread using human mobility and social network information*. in *2011 IEEE*

third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing. 2011. IEEE.
20. Zhou, Y., Z. Ma, and F. Brauer, *A discrete epidemic model for SARS transmission and control in China.* Mathematical and Computer Modelling, 2004. **40**(13): p. 1491-1506.
21. Wang, L., L. Chen, and J.J. Nieto, *The dynamics of an epidemic model for pest control with impulsive effect.* Nonlinear Analysis: Real World Applications, 2010. **11**(3): p. 1374-1386.
22. De la Sen, M. and S. Alonso-Quesada, *Vaccination strategies based on feedback control techniques for a general SEIR-epidemic model.* Applied Mathematics and Computation, 2011. **218**(7): p. 3888-3904.
23. Zaman, G., Y.H. Kang, and I.H. Jung, *Stability analysis and optimal vaccination of an SIR epidemic model.* BioSystems, 2008. **93**(3): p. 240-249.
24. Kar, T. and A. Batabyal, *Stability analysis and optimal control of an SIR epidemic model with vaccination.* Biosystems, 2011. **104**(2-3): p. 127-135.
25. Chen, L. and J. Sun, *Global stability and optimal control of an SIRS epidemic model on heterogeneous networks.* Physica A: Statistical Mechanics and its Applications, 2014. **410**: p. 196-204.
26. Lemos-Paião, A.P., C.J. Silva, and D.F. Torres, *An epidemic model for cholera with optimal control treatment.* Journal of Computational and Applied Mathematics, 2017. **318**: p. 168-180.
27. Lashari, A.A. and G. Zaman, *Optimal control of a vector borne disease with horizontal transmission.* Nonlinear Analysis: Real World Applications, 2012. **13**(1): p. 203-212.
28. Okosun, K.O., R. Ouifki, and N. Marcus, *Optimal control analysis of a malaria disease transmission model that includes treatment and vaccination with waning immunity.* Biosystems, 2011. **106**(2-3): p. 136-145.
29. Meyers, L., *Contact network epidemiology: Bond percolation applied to infectious disease prediction and control.* Bulletin of the American Mathematical Society, 2007. **44**(1): p. 63-86.
30. Bi, K., et al., *A new zoonotic visceral leishmaniasis dynamic transmission model with age-structure.* Chaos, Solitons & Fractals, 2020. **133**: p. 109622.
31. Bachinsky, A.G. and L.P. Nizolenko, *A universal model of epidemic: optimizing interventions.* Universal Journal of Public Health, 2014. **2**(4): p. 111-117.
32. Joshi, H.R., et al., *Optimal control methods applied to disease models.* Contemporary Mathematics, 2006. **410**: p. 187-208.
33. Agusto, F.B. and I.M. ELmojtaba, *Optimal control and cost-effective analysis of malaria/visceral leishmaniasis co-infection.* PloS one, 2017. **12**(2): p. e0171102-e0171102.
34. Bi, K., et al., *Current visceral leishmaniasis research: a research review to inspire future study.* BioMed research international, 2018. **2018**.
35. Bi, K., et al., *A Memetic Algorithm for Solving Optimal Control Problems of Zika Virus Epidemic with Equilibriums and Backward Bifurcation Analysis.* Communications in Nonlinear Science and Numerical Simulation, 2020: p. 105176.
36. Bi, K., et al., *Modeling learning and forgetting processes with the corresponding impacts on human behaviors in infectious disease epidemics.* Computers & Industrial Engineering, 2019. **129**: p. 563-577.

37. Soltanolkottabi, M., *Modeling social response to disease spread using spatial game theory*. 2019.
38. Ochoche, J.M., *A mathematical model for the transmission dynamics of cholera with control strategy.* International Journal of Science and Technology, 2013. **2**(11): p. 212-217.
39. Perez, L. and S. Dragicevic, *An agent-based approach for modeling dynamics of contagious disease spread.* International journal of health geographics, 2009. **8**(1): p. 50.
40. Zhao, S., *Spontaneous changes of human behaviors and intervention strategies: human and animal diseases*. 2017, Kansas State University.
41. Chen, Y., et al., *Modeling individual fear factor with optimal control in a disease-dynamic system.* Chaos, Solitons & Fractals, 2017. **104**: p. 531-545.
42. Leamer, E.E., *Sensitivity analyses would help.* The American Economic Review, 1985. **75**(3): p. 308-313.
43. Pontryagin, L.S., *Mathematical theory of optimal processes*. 2018: Routledge.
44. Chen, Y., et al., *A new evidence-based optimal control in healthcare delivery: A better clinical treatment management for septic patients.* Computers & Industrial Engineering, 2019. **137**: p. 106010.
45. Nowzari, C., V.M. Preciado, and G.J. Pappas, *Analysis and control of epidemics: A survey of spreading processes on complex networks.* IEEE Control Systems Magazine, 2016. **36**(1): p. 26-46.
46. Blackwood, J.C. and L.M. Childs, *An introduction to compartmental modeling for the budding infectious disease modeler.* Letters in Biomathematics, 2018. **5**(1): p. 195-221.
47. Kermack, W.O. and A.G. McKendrick, *A contribution to the mathematical theory of epidemics.* Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character, 1927. **115**(772): p. 700-721.
48. Nåsell, I., *The quasi-stationary distribution of the closed endemic SIS model.* Advances in Applied Probability, 1996. **28**(3): p. 895-932.
49. Li, M.Y. and J.S. Muldowney, *Global stability for the SEIR model in epidemiology.* Mathematical biosciences, 1995. **125**(2): p. 155-164.
50. Zhao, S., et al., *Zoonotic visceral leishmaniasis transmission: modeling, backward bifurcation, and optimal control.* Journal of mathematical biology, 2016. **73**(6-7): p. 1525-1560.
51. Gilbert, N. and P. Terna, *How to build and use agent-based models in social science.* Mind & Society, 2000. **1**(1): p. 57-72.
52. Tracy, M., M. Cerdá, and K.M. Keyes, *Agent-based modeling in public health: current applications and future directions.* Annual review of public health, 2018. **39**: p. 77-94.
53. Hunter, E., B. Mac Namee, and J. Kelleher, *An open-data-driven agent-based model to simulate infectious disease outbreaks.* PloS one, 2018. **13**(12).
54. Van den Driessche, P. and J. Watmough, *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission.* Mathematical biosciences, 2002. **180**(1-2): p. 29-48.
55. Lin, X., H.W. Hethcote, and P. Van den Driessche, *An epidemiological model for HIV/AIDS with proportional recruitment.* Mathematical biosciences, 1993. **118**(2): p. 181-195.

56. Sun, R., *Global stability of the endemic equilibrium of multigroup SIR models with nonlinear incidence.* Computers & Mathematics with Applications, 2010. **60**(8): p. 2286-2291.
57. Wang, W., *Backward bifurcation of an epidemic model with treatment.* Mathematical biosciences, 2006. **201**(1-2): p. 58-71.
58. Ouarda, T. and J. Labadie, *Chance-constrained optimal control for multireservoir system optimization and risk analysis.* Stochastic environmental research and risk assessment, 2001. **15**(3): p. 185-204.
59. Trélat, E., *Optimal control and applications to aerospace: some results and challenges.* Journal of Optimization Theory and Applications, 2012. **154**(3): p. 713-758.
60. Hug-Glanzmann, G. and G. Andersson, *Decentralized optimal power flow control for overlapping areas in power systems.* IEEE Transactions on Power Systems, 2009. **24**(1): p. 327-336.
61. Øksendal, B. and A. Sulem. *A maximum principle for optimal control of stochastic systems with delay, with applications to finance*. in *Optimal control and partial differential equations (Paris, 4 December 2000)*. 2001.
62. Karrakchou, J., M. Rachik, and S. Gourari, *Optimal control and infectiology: application to an HIV/AIDS model.* Applied mathematics and computation, 2006. **177**(2): p. 807-818.
63. Caetano, M.A.L. and T. Yoneyama, *Optimal and sub‐optimal control in Dengue epidemics.* Optimal control applications and methods, 2001. **22**(2): p. 63-73.
64. Makinde, O.D. and K.O. Okosun, *Impact of chemo-therapy on optimal control of malaria disease with infected immigrants.* BioSystems, 2011. **104**(1): p. 32-41.
65. Bock, H.G., et al., *Constrained optimal feedback control of systems governed by large differential algebraic equations,* in *Real-Time PDE-constrained optimization*. 2007, SIAM. p. 3-24.
66. Keller, J.P., L. Gerardo-Giorda, and A. Veneziani, *Numerical simulation of a susceptible–exposed–infectious space-continuous model for the spread of rabies in raccoons across a realistic landscape.* Journal of biological dynamics, 2013. **7**(sup1): p. 31-46.
67. Wilensky, U., *NetLogo [online: http://ccl. northwestern. edu/netlogo/]. Center for Connected Learning and Computer-based Modeling.* Northwestern University, Evanston, IL, 1999.
68. Eissing, T., et al., *A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks.* Frontiers in physiology, 2011. **2**: p. 4.
69. Shi, Z.Z., C.-H. Wu, and D. Ben-Arieh, *Agent-based model: a surging tool to simulate infectious diseases in the immune system.* Open Journal of Modelling and Simulation, 2014. **2014**.
70. Parker, J. and J.M. Epstein, *A distributed platform for global-scale agent-based models of disease transmission.* ACM Transactions on Modeling and Computer Simulation (TOMACS), 2011. **22**(1): p. 1-25.
71. Macal, C. and M. North. *Introductory tutorial: Agent-based modeling and simulation*. in *Proceedings of the Winter Simulation Conference 2014*. 2014. IEEE.
72. Dunham, J.B., *An agent-based spatially explicit epidemiological model in MASON.* Journal of Artificial Societies and Social Simulation, 2005. **9**(1).

73. Grignard, A., et al. *GAMA 1.6: Advancing the art of complex agent-based modeling and simulation*. in *International conference on principles and practice of multi-agent systems*. 2013. Springer.
74. Al-Salem, W., J.R. Herricks, and P.J. Hotez, *A review of visceral leishmaniasis during the conflict in South Sudan and the consequences for East African countries.* Parasites & vectors, 2016. **9**(1): p. 460.
75. Alvar, J., S. Yactayo, and C. Bern, *Leishmaniasis and poverty.* Trends in parasitology, 2006. **22**(12): p. 552-557.
76. Organization, W.H., *Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases.* 2010.
77. Postigo, J.A.R., *Leishmaniasis in the world health organization eastern mediterranean region.* International journal of antimicrobial agents, 2010. **36**: p. S62-S65.
78. Oryan, A. and M. Akbari, *Worldwide risk factors in leishmaniasis.* Asian Pacific journal of tropical medicine, 2016. **9**(10): p. 925-932.
79. Jain, K. and N. Jain, *Vaccines for visceral leishmaniasis: A review.* Journal of immunological methods, 2015. **422**: p. 1-12.
80. Tesh, R.B., *Control of zoonotic visceral leishmaniasis: is it time to change strategies?* The American journal of tropical medicine and hygiene, 1995. **52**(3): p. 287-292.
81. Guerin, P.J., et al., *Visceral leishmaniasis: current status of control, diagnosis, and treatment, and a proposed research and development agenda.* The Lancet infectious diseases, 2002. **2**(8): p. 494-501.
82. Dantas-Torres, F. and S.P. Brandão-Filho, *Visceral leishmaniasis in Brazil: revisiting paradigms of epidemiology and control.* Revista do Instituto de Medicina Tropical de São Paulo, 2006. **48**(3): p. 151-156.
83. Quinnell, R. and O. Courtenay, *Transmission, reservoir hosts and control of zoonotic visceral leishmaniasis.* Parasitology, 2009. **136**(14): p. 1915-1934.
84. Pigott, D.M., et al., *Global database of leishmaniasis occurrence locations, 1960–2012.* Scientific data, 2014. **1**: p. 140036.
85. Chappuis, F., et al., *Visceral leishmaniasis: what are the needs for diagnosis, treatment and control?* Nature reviews microbiology, 2007. **5**(11): p. 873.
86. Thimphu, B., *REVIEW OF THE DECISIONS AND RESOLUTIONS OF THE SIXTIETH WORLD HEALTH ASSEMBLY AND THE 120TH AND 121ST SESSIONS OF THE EXECUTIVE BOARD.* 2007.
87. Organization, W.H., *Report of a WHO informal consultation on liposomal amphotericin B in the treatment of visceral leishmaniasis, Rome, Italy, 16 April 2005.* 2007.
88. Alvar, J., et al., *Leishmaniasis worldwide and global estimates of its incidence.* PloS one, 2012. **7**(5): p. e35671.
89. Ostyn, B., et al., *Vector control by insecticide‐treated nets in the fight against visceral leishmaniasis in the Indian subcontinent, what is the evidence?* Tropical medicine & international health, 2008. **13**(8): p. 1073-1085.
90. Picado, A., et al., *Longlasting insecticidal nets for prevention of Leishmania donovani infection in India and Nepal: paired cluster randomised trial.* BMJ, 2010. **341**: p. c6760.
91. Da Silva, M. and M.I. Cano, *Frontiers in Parasitology - Molecular and Cellular Biology of Pathogenic Trypanosomatids*. 2017.
92. Werneck, G.L., *Visceral leishmaniasis in Brazil: rationale and concerns related to reservoir control.* Revista de saude publica, 2014. **48**(5): p. 851-856.

93. Lengeler, C., *Insecticide‐treated bed nets and curtains for preventing malaria.* The Cochrane Library, 2004.
94. Picado, A., et al., *Long-lasting insecticidal nets to prevent visceral leishmaniasis in the Indian subcontinent; methodological lessons learned from a cluster randomised controlled trial.* PLoS neglected tropical diseases, 2015. **9**(4): p. e0003597.
95. Poché, R.M., et al., *The role of Palmyra palm trees (Borassus flabellifer) and sand fly distribution in northeastern India.* Journal of Vector Ecology, 2012. **37**(1): p. 148-153.
96. Özbel, Y., C. Sanjoba, and Y. Matsumoto, *Geographical distribution and ecological aspect of sand fly species in Bangladesh*, in *Kala Azar in South Asia*. 2016, Springer. p. 199-209.
97. Jaffe, C.L., N. Rachamim, and R. Sarfstein, *Characterization of two proteins from Leishmania donovani and their use for vaccination against visceral leishmaniasis.* The Journal of Immunology, 1990. **144**(2): p. 699-706.
98. Kedzierski, L., Y. Zhu, and E. Handman, *Leishmania vaccines: progress and problems.* Parasitology, 2006. **133**(S2): p. S87-S112.
99. Amaral, V., et al., *Study of the safety, immunogenicity and efficacy of attenuated and killed Leishmania (Leishmania) major vaccines in a rhesus monkey (Macaca mulatta) model of the human disease.* Memorias do Instituto Oswaldo Cruz, 2002. **97**(7): p. 1041-1048.
100. Duthie, M.S., et al., *Strategic evaluation of vaccine candidate antigens for the prevention of Visceral Leishmaniasis.* Vaccine, 2016. **34**(25): p. 2779-2786.
101. Dye, C., *The logic of visceral leishmaniasis control.* The American journal of tropical medicine and hygiene, 1996. **55**(2): p. 125-130.
102. Courtenay, O., et al., *Infectiousness in a cohort of Brazilian dogs: why culling fails to control visceral leishmaniasis in areas of high transmission.* The Journal of infectious diseases, 2002. **186**(9): p. 1314-1320.
103. Ribas, L.M., et al., *Estimating the optimal control of zoonotic visceral leishmaniasis by the use of a mathematical model.* The Scientific World Journal, 2013. **2013**.
104. Organization, W.H., *Investing to Overcome the Global Impact of Neglected Tropical Diseases: Third Who Report on negLected Tropical Diseases 2015*. Vol. 3. 2015: World Health Organization.
105. Subramanian, A., V. Singh, and R.R. Sarkar, *Understanding Visceral Leishmaniasis Disease Transmission and its Control—A Study Based on Mathematical Modeling.* Mathematics, 2015. **3**(3): p. 913-944.
106. Biswas, S., *Mathematical modeling of Visceral Leishmaniasis and control strategies.* Chaos, Solitons & Fractals, 2017. **104**: p. 546-556.
107. Shimozako, H.J., J. Wu, and E. Massad, *Mathematical modelling for Zoonotic Visceral Leishmaniasis dynamics: a new analysis considering updated parameters and notified human Brazilian data.* Infectious Disease Modelling, 2017. **2**(2): p. 143-160.
108. Le Rutte, E.A., et al., *Elimination of visceral leishmaniasis in the Indian subcontinent: a comparison of predictions from three transmission models.* Epidemics, 2017. **18**: p. 67-80.
109. Costa, D.N., et al., *Culling dogs in scenarios of imperfect control: realistic impact on the prevalence of canine visceral leishmaniasis.* PLoS neglected tropical diseases, 2013. **7**(8): p. e2355.

110. Sevá, A.P., et al., *Canine-based strategies for prevention and control of visceral leishmaniasis in Brazil.* PloS one, 2016. **11**(7): p. e0160058.
111. Zou, L., J. Chen, and S. Ruan, *Modeling and analyzing the transmission dynamics of visceral leishmaniasis.* Mathematical Biosciences & Engineering, 2017. **14**(5-6): p. 1585-1604.
112. Agusto, F.B. and I.M. ELmojtaba, *Optimal control and cost-effective analysis of malaria/visceral leishmaniasis co-infection.* PloS one, 2017. **12**(2): p. e0171102.
113. Bi, K., et al., *A New Zoonotic Viseral Leishmaniasis Dynamic Transmission Model with Age-Structure.* PLoS neglected tropical diseases, 2018.
114. Werneck, G.L. and J.H. Maguire, *Spatial modeling using mixed models: an ecologic study of visceral leishmaniasis in Teresina, Piauí State, Brazil.* Cadernos de Saúde Pública, 2002. **18**(3): p. 633-637.
115. Werneck, G., et al., *Multilevel modelling of the incidence of visceral leishmaniasis in Teresina, Brazil.* Epidemiology & Infection, 2007. **135**(2): p. 195-201.
116. Thompson, R.A., et al., *Climatic and demographic determinants of American visceral leishmaniasis in northeastern Brazil using remote sensing technology for environmental categorization of rain and region influences on leishmaniasis.* The American journal of tropical medicine and hygiene, 2002. **67**(6): p. 648-655.
117. de Araújo, V.E.M., et al., *Relative risk of visceral leishmaniasis in Brazil: a spatial analysis in urban area.* PLoS neglected tropical diseases, 2013. **7**(11): p. e2540.
118. McNyset, K., *Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas.* Ecology of Freshwater Fish, 2005. **14**(3): p. 243-255.
119. Nieto, P., J.B. Malone, and M.E. Bavia, *Ecological niche modeling for visceral leishmaniasis in the state of Bahia, Brazil, using genetic algorithm for rule-set prediction and growing degree day-water budget analysis.* Geospatial health, 2006. **1**(1): p. 115-126.
120. González, C., et al., *Climate change and risk of leishmaniasis in North America: predictions from ecological niche models of vector and reservoir species.* PLoS neglected tropical diseases, 2010. **4**(1): p. e585.
121. Colacicco-Mayhugh, M.G., P.M. Masuoka, and J.P. Grieco, *Ecological niche model of Phlebotomus alexandri and P. papatasi (Diptera: Psychodidae) in the Middle East.* International journal of health geographics, 2010. **9**(1): p. 2.
122. Elnaiem, D.-E.A., et al., *Risk mapping of visceral leishmaniasis: the role of local variation in rainfall and altitude on the presence and incidence of kala-azar in eastern Sudan.* The American journal of tropical medicine and hygiene, 2003. **68**(1): p. 10-17.
123. Oshaghi, M., et al., *Application of predictive degree day model for field development of sandfly vectors of visceral leishmaniasis in northwest of Iran.* Journal of Vector Borne Diseases, 2009. **46**(4): p. 247.
124. Karagiannis-Voules, D.-A., et al., *Bayesian geostatistical modeling of leishmaniasis incidence in Brazil.* PLoS neglected tropical diseases, 2013. **7**(5): p. e2213.
125. Shimozako, H.J., J. Wu, and E. Massad, *The Preventive Control of Zoonotic Visceral Leishmaniasis: Efficacy and Economic Evaluation.* Computational and mathematical methods in medicine, 2017. **2017**.
126. Pontryagin, L.S., *On the zeros of some elementary transcendental functions.* Amer. Math. Soc. Transl, 1955. **2**(1): p. 95-110.

127. Pontryagin, L.S., *The mathematical theory of optimal processes and differential games.* Trudy Matematicheskogo Instituta imeni VA Steklova, 1985. **169**: p. 119-158.
128. Biswas, S., et al., *Optimal combinations of control strategies and cost-effective analysis for visceral leishmaniasis disease transmission.* PloS one, 2017. **12**(2): p. e0172465.
129. Miller, E., et al., *Quantifying the contribution of hosts with different parasite concentrations to the transmission of visceral leishmaniasis in Ethiopia.* PLoS neglected tropical diseases, 2014. **8**(10): p. e3288.
130. Stauch, A., et al., *Model-based investigations of different vector-related intervention strategies to eliminate visceral leishmaniasis on the Indian subcontinent.* PLoS neglected tropical diseases, 2014. **8**(4): p. e2810.
131. Zamir, M., G. Zaman, and A.S. Alshomrani, *Control strategies and sensitivity analysis of anthroponotic visceral leishmaniasis model.* Journal of biological dynamics, 2017. **11**(1): p. 323-338.
132. Boukhalfa, F., M. Helal, and A. Lakmeche, *Mathematical Analysis of Visceral Leishmaniasis Model.* Research in Applied Mathematics, 2017. **1**.
133. Gorahava, K.K., J.M. Rosenberger, and A. Mubayi, *Optimizing insecticide allocation strategies based on houses and livestock shelters for visceral leishmaniasis control in Bihar, India.* The American journal of tropical medicine and hygiene, 2015. **93**(1): p. 114-122.
134. Rock, K., et al., *Progress in the mathematical modelling of visceral leishmaniasis,* in *Advances in parasitology*. 2016, Elsevier. p. 49-131.
135. Topno, R.K., et al., *Asymptomatic infection with visceral leishmaniasis in a disease-endemic area in Bihar, India.* The American journal of tropical medicine and hygiene, 2010. **83**(3): p. 502-506.
136. Desjeux, P., *Leishmaniasis: current situation and new perspectives.* Comparative immunology, microbiology and infectious diseases, 2004. **27**(5): p. 305-318.
137. Palatnik-de-Sousa, C.B. and M.J. Day, *One Health: the global challenge of epidemic and endemic leishmaniasis.* Parasites & vectors, 2011. **4**(1): p. 197.
138. Gradoni, L., *A Brief Introduction to Leishmaniasis Epidemiology,* in *The Leishmaniases: Old Neglected Tropical Diseases*. 2018, Springer. p. 1-13.
139. Fukami, R. and H. Nishiura, *Examining the reservoir potential of animal species for $ $\varvec {\textit {Leishmania infantum}} $ $ Leishmania infantum infection.* Japan Journal of Industrial and Applied Mathematics, 2015. **32**(3): p. 661-673.
140. Travi, B., et al., *Canine visceral leishmaniasis: dog infectivity to sand flies from non-endemic areas.* Research in veterinary science, 2002. **72**(1): p. 83-86.
141. Espejo, L., S. Costard, and F. Zagmutt, *Modelling canine leishmaniasis spread to non-endemic areas of Europe.* Epidemiology & Infection, 2015. **143**(9): p. 1936-1949.
142. Antoniou, M., et al., *Increasing incidence of zoonotic visceral leishmaniasis on Crete, Greece.* Emerging infectious diseases, 2009. **15**(6): p. 932.
143. Hollingsworth, T.D., et al., *Quantitative analyses and modelling to support achievement of the 2020 goals for nine neglected tropical diseases.* Parasites & vectors, 2015. **8**(1): p. 630.
144. Gradoni, L., A. Bryceson, and P. Desjeux, *Treatment of Mediterranean visceral leishmaniasis.* Bulletin of the World Health Organization, 1995. **73**(2): p. 191.
145. Podaliri Vulpiani, M., et al., *Methods of control of the Leishmania infantum dog reservoir: state of the art.* Veterinary medicine international, 2011. **2011**.

146. Badaro, R., et al., *A prospective study of visceral leishmaniasis in an endemic area of Brazil.* Journal of Infectious Diseases, 1986. **154**(4): p. 639-649.
147. Gramiccia, M., et al., *The burden of visceral leishmaniasis in Italy from 1982 to 2012: a retrospective analysis of the multi-annual epidemic that occurred from 1989 to 2009.* Eurosurveillance, 2013. **18**(29): p. 20535.
148. Collin, S., et al., *Conflict and kala-azar: determinants of adverse outcomes of kala-azar among patients in southern Sudan.* Clinical Infectious Diseases, 2004. **38**(5): p. 612-619.
149. Zijlstra, E., et al., *Endemic kala-azar in eastern Sudan: a longitudinal study on the incidence of clinical and subclinical infection and post-kala-azar dermal leishmaniasis.* The American journal of tropical medicine and hygiene, 1994. **51**(6): p. 826-836.
150. Olliaro, P. and S. Sundar, *Anthropometrically derived dosing and drug costing calculations for treating visceral leishmaniasis in Bihar, India.* Tropical Medicine & International Health, 2009. **14**(1): p. 88-92.
151. Wang, J.-Y., et al., *Current epidemiological profile and features of visceral leishmaniasis in People's Republic of China.* Parasites & vectors, 2012. **5**(1): p. 31.
152. Terrosi, C., et al., *Age-dependent seroprevalence of Toscana virus in central Italy and correlation with the clinical profile.* Clinical and Vaccine Immunology, 2009. **16**(8): p. 1251-1252.
153. Lee, B.Y., et al., *The economic value of a visceral leishmaniasis vaccine in Bihar state, India.* The American journal of tropical medicine and hygiene, 2012. **86**(3): p. 417-425.
154. Reis, A.B., et al., *Systemic and compartmentalized immune response in canine visceral leishmaniasis.* Veterinary immunology and immunopathology, 2009. **128**(1-3): p. 87-95.
155. Costa, C.H., et al., *Competence of the human host as a reservoir for Leishmania chagasi.* The Journal of infectious diseases, 2000. **182**(3): p. 997-1000.
156. Hartemink, N., et al., *Integrated mapping of establishment risk for emerging vector-borne infections: a case study of canine leishmaniasis in southwest France.* PloS one, 2011. **6**(8): p. e20817.
157. Burattini, M.N., et al., *Modelling the dynamics of leishmaniasis considering human, animal host and vector populations.* Journal of Biological Systems, 1998. **6**(04): p. 337-356.
158. Martin, J.A., et al., *Births: final data for 2013.* 2015.
159. Helel, K.B., et al., *Risk factors for mortality of children with zoonotic visceral leishmaniasis in Central Tunisia.* PloS one, 2017. **12**(12): p. e0189725.
160. Belhadj, S., et al., *Leishmanioses viscérales et cutanées du nord. Etude rétroactive des cas diagnostiqués à l'hôpital la Rabta de Tunis.* Bulletin de la Société de Pathologie Exotique, 1996. **89**: p. 269-273.
161. ALBORZI, A., M. RASOULI, and A. SHAMSIZADEH, *Leishmania tropica–isolated patient with visceral leishmaniasis in southern Iran.* The American journal of tropical medicine and hygiene, 2006. **74**(2): p. 306-307.
162. Palatnik-de-Sousa, C.B., et al., *Impact of canine control on the epidemiology of canine and human visceral leishmaniasis in Brazil.* The American journal of tropical medicine and hygiene, 2001. **65**(5): p. 510-517.
163. Desjeux, P., *The increase in risk factors for leishmaniasis worldwide.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2001. **95**(3): p. 239-243.
164. Bora, D., *Epidemiology of visceral leishmaniasis in India.* National Medical Journal of India, 1999. **12**: p. 62-68.

165. Mittal, A. and N. Paragios. *Motion-based background subtraction using adaptive kernel density estimation*. in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. 2004. Ieee.
166. Botev, Z.I., J.F. Grotowski, and D.P. Kroese, *Kernel density estimation via diffusion.* The annals of Statistics, 2010. **38**(5): p. 2916-2957.
167. Esch, K.J. and C.A. Petersen, *Transmission and epidemiology of zoonotic protozoal diseases of companion animals.* Clinical microbiology reviews, 2013. **26**(1): p. 58-85.
168. Mahmoud, M.Z., *Assessment of visceral leishmaniasis consequences using ultrasound.* Open Journal of Radiology, 2014. **4**(02): p. 201.
169. Blayneh, K.W., et al., *Backward bifurcation and optimal control in transmission dynamics of West Nile virus.* Bulletin of mathematical biology, 2010. **72**(4): p. 1006-1028.
170. Stauch, A., et al., *Visceral leishmaniasis in the Indian subcontinent: modelling epidemiology and control.* PLoS neglected tropical diseases, 2011. **5**(11): p. e1405.
171. Jameson, A., W. Schmidt, and E. Turkel. *Numerical solution of the Euler equations by finite volume methods using Runge Kutta time stepping schemes*. in *14th fluid and plasma dynamics conference*. 1981.
172. Coren, S., *Do Dogs Dream?: Nearly Everything Your Dog Wants You to Know*. 2012: WW Norton & Company.
173. Widaa, S.O., et al., *Sandflies (Diptera: Psychodidae) in a focus of visceral leishmaniasis in White Nile, Sudan.* Memórias do Instituto Oswaldo Cruz, 2012. **107**(4): p. 470-475.
174. Garba, S.M., A.B. Gumel, and M.A. Bakar, *Backward bifurcations in dengue transmission dynamics.* Mathematical biosciences, 2008. **215**(1): p. 11-25.
175. Diekmann, O., J.A.P. Heesterbeek, and J.A. Metz, *On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations.* Journal of mathematical biology, 1990. **28**(4): p. 365-382.
176. Dietz, K., *The estimation of the basic reproduction number for infectious diseases.* Statistical methods in medical research, 1993. **2**(1): p. 23-41.
177. Chowell, G., M. Miller, and C. Viboud, *Seasonal influenza in the United States, France, and Australia: transmission and prospects for control.* Epidemiology & Infection, 2008. **136**(6): p. 852-864.
178. Faulde, M., et al., *Zoonotic cutaneous leishmaniasis outbreak in Mazar-e Sharif, northern Afghanistan: an epidemiological evaluation.* International Journal of Medical Microbiology, 2008. **298**(5-6): p. 543-550.
179. Faulde, M.K., G. Heyl, and M.L. Amirih, *Zoonotic cutaneous leishmaniasis, Afghanistan.* Emerging infectious diseases, 2006. **12**(10): p. 1623.
180. Faulde, M., et al., *High efficacy of integrated preventive measures against zoonotic cutaneous leishmaniasis in northern Afghanistan, as revealed by quantified infection rates.* Acta tropica, 2009. **110**(1): p. 28-34.
181. Hassard, B. and Y. Wan, *Bifurcation formulae derived from center manifold theory.* Journal of Mathematical Analysis and Applications, 1978. **63**(1): p. 297-312.
182. Mills, T.C., *Time series techniques for economists*. 1991: Cambridge University Press.
183. Box, G.E., et al., *Time series analysis: forecasting and control*. 2015: John Wiley & Sons.
184. Whitle, P., *Hypothesis testing in time series analysis*. Vol. 4. 1951: Almqvist & Wiksells.

185. Schaffer, C., *Selecting a classification method by cross-validation.* Machine Learning, 1993. **13**(1): p. 135-143.
186. Google. *Google Public Data*. 2017 [cited 2017 Oct 6th]; Available from: https://www.google.com/publicdata/explore?ds=d5bncppjof8f9_&met_y=sp_pop_totl&idim=country:SDN:SSD:SOM&hl=en&dl=en.
187. Bradley, T. and R. King, *The dog economy is global—but what is the world's true capital.* The Atlantic, 2012.
188. Dale, S. *World Pet Population Data at Mixed Bag Google Public Data*. 2016 [cited 2017 Oct 6th]; Available from: http://stevedalepetworld.com/world-pet-population-data-mixed-bag/.
189. Nascimento, B.W.L., et al., *Study of sand flies (Diptera: Psychodidade) in visceral and cutaneous leishmaniasis areas in central western of Minas Gerais state–Brazil.* Acta tropica, 2013. **125**(3): p. 262-268.
190. Poché, D., et al., *Bionomics of phlebotomine sand flies from three villages in Bihar, India.* Journal of Vector Ecology, 2011. **36**(s1).
191. Gebresilassie, A., et al., *Species composition of phlebotomine sand flies and bionomics of Phlebotomus orientalis (Diptera: Psychodidae) in an endemic focus of visceral leishmaniasis in Tahtay Adiyabo district, Northern Ethiopia.* Parasites & vectors, 2015. **8**(1): p. 248.
192. Gavgani, A.M., et al., *Effect of insecticide-impregnated dog collars on incidence of zoonotic visceral leishmaniasis in Iranian children: a matchedcluster randomised trial.* The Lancet, 2002. **360**(9330): p. 374-379.
193. Manson-Bahr, P., B. Southgate, and A. Harvey, *Development of kala-azar in man after inoculation with a Leishmania from a Kenya sandfly.* British Medical Journal, 1963. **1**(5339): p. 1208.
194. Campos, G.S., A.C. Bandeira, and S.I. Sardi, *Zika virus outbreak, Bahia, Brazil.* Emerging infectious diseases, 2015. **21**(10): p. 1885.
195. Musso, D., et al., *Potential sexual transmission of Zika virus.* Emerging infectious diseases, 2015. **21**(2): p. 359.
196. Kindhauser, M.K., et al., *Zika: the origin and spread of a mosquito-borne virus.* Bulletin of the World Health Organization, 2016. **94**(9): p. 675.
197. PAHO, P.A.H.O. *Zika Cumulative Cases*. 2018 [cited 2018 08/01]; Available from: https://www.paho.org/hq/index.php?option=com_content&view=article&id=12390&Itemid=42090&lang=en.
198. Duffy, M.R., et al., *Zika virus outbreak on Yap Island, federated states of Micronesia.* New England Journal of Medicine, 2009. **360**(24): p. 2536-2543.
199. de Paula Freitas, B., et al., *Ocular findings in infants with microcephaly associated with presumed Zika virus congenital infection in Salvador, Brazil.* JAMA ophthalmology, 2016. **134**(5): p. 529-535.
200. Fernandez, E. and M.S. Diamond, *Vaccination strategies against Zika virus.* Current opinion in virology, 2017. **23**: p. 59-67.
201. Lee, B.Y., et al., *The potential economic burden of Zika in the continental United States.* PLoS neglected tropical diseases, 2017. **11**(4): p. e0005531.
202. Karwowski, M.P., et al., *Zika virus disease: a CDC update for pediatric health care providers.* Pediatrics, 2016. **137**(5): p. e20160621.

203. Kucharski, A.J., et al., *Transmission dynamics of Zika virus in island populations: a modelling analysis of the 2013–14 French Polynesia outbreak.* PLoS neglected tropical diseases, 2016. **10**(5): p. e0004726.
204. Gao, D., et al., *Prevention and control of Zika as a mosquito-borne and sexually transmitted disease: a mathematical modeling analysis.* Scientific reports, 2016. **6**: p. 28070.
205. Isea, R. and K.E. Lonngren, *A Preliminary Mathematical Model for the Dynamic Transmission of Dengue, Chikungunya and Zika.* arXiv preprint arXiv:1606.08233, 2016.
206. Agusto, F., S. Bewick, and W. Fagan, *Mathematical model for Zika virus dynamics with sexual transmission route.* Ecological Complexity, 2017. **29**: p. 61-81.
207. Ding, C., N. Tao, and Y. Zhu. *A mathematical model of Zika virus and its optimal control*. in *Control Conference (CCC), 2016 35th Chinese*. 2016. IEEE.
208. Shah, N.H., Z.A. Patel, and B.M. Yeolekar, *Preventions and Controls on Congenital Transmissions of Zika: Mathematical Analysis.* Applied Mathematics, 2017. **8**(04): p. 500.
209. Tchuenche, J., et al., *Optimal control and sensitivity analysis of an influenza model with treatment and vaccination.* Acta biotheoretica, 2011. **59**(1): p. 1-28.
210. Bi, K., et al., *Modeling learning and forgetting processes with the corresponding impacts on human behaviors in infectious disease epidemics.* Computers & Industrial Engineering, 2018.
211. Martinuzzi, S., W.A. Gould, and O.M.R. Gonzalez, *Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data.* Landscape and Urban Planning, 2007. **79**(3-4): p. 288-297.
212. Andraud, M., et al., *Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches.* PloS one, 2012. **7**(11): p. e49085.
213. Chikaki, E. and H. Ishikawa, *A dengue transmission model in Thailand considering sequential infections with all four serotypes.* The Journal of Infection in Developing Countries, 2009. **3**(09): p. 711-722.
214. Maxian, O., et al., *Zika virus dynamics: When does sexual transmission matter?* Epidemics, 2017. **21**: p. 48-55.
215. de Castro Medeiros, L.C., et al., *Modeling the dynamic transmission of dengue fever: investigating disease persistence.* PLOS neglected tropical diseases, 2011. **5**(1): p. e942.
216. Bearcroft, W., *Zika virus infection experimentally induced in a human volunteer.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 1956. **50**(5).
217. Gourinat, A.-C., et al., *Detection of Zika virus in urine.* Emerging infectious diseases, 2015. **21**(1): p. 84.
218. Medford, A. and J.W. Vaupel, *Human lifespan records are not remarkable but their durations are.* PloS one, 2019. **14**(3): p. e0212345.
219. Dong, X., B. Milholland, and J. Vijg, *Evidence for a limit to human lifespan.* Nature, 2016. **538**(7624): p. 257.
220. Abbink, P., et al., *Protective efficacy of multiple vaccine platforms against Zika virus challenge in rhesus monkeys.* Science, 2016: p. aah6157.
221. Larocca, R.A., et al., *Vaccine protection against Zika virus from Brazil.* Nature, 2016. **536**(7617): p. 474-478.

222. Muthumani, K., et al., *In vivo protection against ZIKV infection and pathogenesis through passive antibody transfer and active immunisation with a prMEnv DNA vaccine.* npj Vaccines, 2016. **1**: p. 16021.
223. Diseases, N.I.o.A.a.I. *Phase 2 Zika Vaccine Trial Begins in U.S., Central and South America*. 2017 [cited 2018 8 Mar]; Available from: https://www.niaid.nih.gov/news-events/phase-2-zika-vaccine-trial-begins-us-central-and-south-america.
224. Lima, E.P., et al., *Insecticide resistance in Aedes aegypti populations from Ceará, Brazil.* Parasites & Vectors, 2011. **4**(1): p. 5.
225. Yakob, L. and T. Walker, *Zika virus outbreak in the Americas: the need for novel mosquito control methods.* The Lancet Global Health, 2016. **4**(3): p. e148-e149.
226. Phuc, H.K., et al., *Late-acting dominant lethal genetic systems and mosquito control.* BMC biology, 2007. **5**(1): p. 11.
227. Walker, T., et al., *The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations.* Nature, 2011. **476**(7361): p. 450.
228. Foy, B.D., et al., *Probable non–vector-borne transmission of Zika virus, Colorado, USA.* Emerging infectious diseases, 2011. **17**(5): p. 880.
229. Anderson, R.M., R.M. May, and B. Anderson, *Infectious diseases of humans: dynamics and control*. Vol. 28. 1992: Wiley Online Library.
230. Pontryagin, L.S., *Mathematical theory of optimal processes*. 1987: CRC Press.
231. Cruz, I.L., *Efficient evolutionary algorithms for optimal control*. 2002: sn].
232. Saerens, B., M. Diehl, and E. Van den Bulck, *Optimal control using pontryagin's maximum principle and dynamic programming*, in *Automotive Model Predictive Control*. 2010, Springer. p. 119-138.
233. Strekalovsky, A.S., *On global maximum of a convex terminal functional in optimal control problems.* Journal of global optimization, 1995. **7**(1): p. 75-91.
234. Kim, N., S. Cha, and H. Peng, *Optimal control of hybrid electric vehicles based on Pontryagin's minimum principle.* IEEE Transactions on Control Systems Technology, 2011. **19**(5): p. 1279-1287.
235. Moscato, P., *On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms.* Caltech concurrent computation program, C3P Report, 1989. **826**: p. 1989.
236. Mladenović, N. and P. Hansen, *Variable neighborhood search.* Computers & operations research, 1997. **24**(11): p. 1097-1100.
237. Lowe, R., et al., *The Zika Virus Epidemic in Brazil: From Discovery to Future Implications.* International journal of environmental research and public health, 2018. **15**(1): p. 96.
238. Bullard-Feibelman, K.M., et al., *The FDA-approved drug sofosbuvir inhibits Zika virus infection.* Antiviral research, 2017. **137**: p. 134-140.
239. Leander, R., S. Lenhart, and V. Protopopescu, *Optimal control of continuous systems with impulse controls.* Optimal Control Applications and Methods, 2015. **36**(4): p. 535-549.
240. Simeonov, P. and D. Bainov, *Stability with respect to part of the variables in systems with impulse effect.* Journal of mathematical analysis and applications, 1986. **117**(1): p. 247-263.

241. Morris, B. and J.W. Grizzle, *Hybrid invariant manifolds in systems with impulse effects with application to periodic locomotion in bipedal robots.* IEEE Transactions on Automatic Control, 2009. **54**(8): p. 1751-1764.
242. Yang, T. and L.O. Chua, *Impulsive stabilization for control and synchronization of chaotic systems: theory and application to secure communication.* IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 1997. **44**(10): p. 976-988.
243. Pellicano, C., et al., *Morphometric changes in the reward system of Parkinson's disease patients with impulse control disorders.* Journal of neurology, 2015. **262**(12): p. 2653-2661.
244. Korn, R., *Some applications of impulse control in mathematical finance.* Mathematical Methods of Operations Research, 1999. **50**(3): p. 493-518.
245. Gao, S., et al., *Analysis of an SIR epidemic model with pulse vaccination and distributed time delay.* BioMed Research International, 2007. **2007**.
246. Liu, Q., M. Zhang, and L. Chen, *State feedback impulsive therapy to SIS model of animal infectious diseases.* Physica A: Statistical Mechanics and its Applications, 2019. **516**: p. 222-232.
247. Yu, H., et al., *Effect of seasonality on the dynamical behavior of an ecological system with impulsive control strategy.* Journal of the Franklin Institute, 2011. **348**(4): p. 652-670.
248. Zou, L., Z. Xiong, and Z. Shu, *The dynamics of an eco-epidemic model with distributed time delay and impulsive control strategy.* Journal of the Franklin Institute, 2011. **348**(9): p. 2332-2349.
249. Wang, L., Y. Xie, and J. Fu, *The dynamics of natural mortality for pest control model with impulsive effect.* Journal of the Franklin Institute, 2013. **350**(6): p. 1443-1461.
250. He, Z.L., et al., *Nonlinear state-dependent feedback control strategy in the SIR epidemic model with resource limitation.* Advances in Difference Equations, 2017. **2017**(1): p. 1-18.
251. De la Sen, M., et al., *On a generalized time-varying SEIR epidemic model with mixed point and distributed time-varying delays and combined regular and impulsive vaccination controls.* Advances in Difference Equations, 2010. **2010**(1): p. 281612.
252. He, Z.L., et al., *Nonlinear state-dependent feedback control strategy in the SIR epidemic model with resource limitation.* Advances in Difference Equations, 2017. **2017**(1): p. 209.
253. Taynitskiy, V., E. Gubar, and Q. Zhu, *Optimal impulsive control of epidemic spreading of heterogeneous malware.* IFAC-PapersOnLine, 2017. **50**(1): p. 15038-15043.
254. Zeng, G.Z., L.S. Chen, and L.H. Sun, *Complexity of an SIR epidemic dynamics model with impulsive vaccination control.* Chaos, Solitons & Fractals, 2005. **26**(2): p. 495-505.
255. Verriest, E.I., F. Delmotte, and M. Egerstedt. *Optimal impulsive control for point delay systems with refractory period*. in *Proceedings of the 5-th IFAC Workshop on Time Delay Systems, Leuven, Belgium*. 2004.
256. Karnik, A. and P. Dayama. *Optimal control of information epidemics*. in *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012)*. 2012. IEEE.
257. Chatfield, C., *The analysis of time series: an introduction*. 2003: Chapman and Hall/CRC.

258. Kim, J., J. Kwon Lee, and K. Mu Lee. *Deeply-recursive convolutional network for image super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
259. Sutskever, I., J. Martens, and G.E. Hinton. *Generating text with recurrent neural networks*. in *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011.
260. Schuster, M. and K.K. Paliwal, *Bidirectional recurrent neural networks.* IEEE transactions on Signal Processing, 1997. **45**(11): p. 2673-2681.
261. Graves, A. and J. Schmidhuber, *Framewise phoneme classification with bidirectional LSTM and other neural network architectures.* Neural networks, 2005. **18**(5-6): p. 602-610.
262. McCluskey, C.C., *A model of HIV/AIDS with staged progression and amelioration.* Mathematical biosciences, 2003. **181**(1): p. 1-16.
263. Krasovskii, N.N., *Stability of motion*. Vol. 2. 1963: Stanford university press Stanford.
264. Pang, G. and L. Chen, *Periodic solution of the system with impulsive state feedback control.* Nonlinear Dynamics, 2014. **78**(1): p. 743-753.
265. Lever, J., M. Krzywinski, and N. Altman, *Points of significance: model selection and overfitting*. 2016, Nature Publishing Group.
266. Polgar, S., *Health and human behavior: areas of interest common to the social and medical sciences.* Current Anthropology, 1962. **3**(2): p. 159-205.
267. Shiffrin, R.M. and R.C. Atkinson, *Storage and retrieval processes in long-term memory.* Psychological Review, 1969. **76**(2): p. 179.
268. Engle, R.W., et al., *Working memory, short-term memory, and general fluid intelligence: a latent-variable approach.* Journal of experimental psychology: General, 1999. **128**(3): p. 309.
269. Ebbinghaus, H., *Memory: A contribution to experimental psychology (No. 3). University Microfilms*. 1913.
270. Wingfield, A. and D.L. Byrnes, *The psychology of human memory*. 2013: Academic Press.
271. Weltens, B. and A.D. Cohen, *Language attrition research: An introduction.* Studies in Second Language Acquisition, 1989. **11**(2): p. 127-133.
272. Bailey, C.D., *Forgetting and the learning curve: A laboratory study*. Management science, 1989. **35**(3): p. 340-352.
273. Badiru, A.B., *Computational survey of univariate and multivariate learning curve models.* IEEE transactions on Engineering Management, 1992. **39**(2): p. 176-188.
274. Reyna, V.F. and C.J. Brainerd, *Fuzzy-trace theory: An interim synthesis.* Learning and individual Differences, 1995. **7**(1): p. 1-75.
275. Carlson, J.G. and A.J. Rowe, *How much does forgetting cost.* Industrial Engineering, 1976. **8**(9): p. 40-47.
276. ELM'AGHRABY, S.E., *Economic manufacturing quantities under conditions of learning and forgetting (EMQ/LaF).* Production Planning & Control, 1990. **1**(4): p. 196-208.
277. Jaber, M.Y. and M. Bonney, *Production breaks and the learning curve: the forgetting phenomenon.* Applied mathematical modelling, 1996. **2**(20): p. 162-169.
278. Jaber, M.Y., H.V. Kher, and D.J. Davis, *Countering forgetting through training and deployment.* International Journal of Production Economics, 2003. **85**(1): p. 33-46.

279. Sikström, S. and M.Y. Jaber, *The power integration diffusion model for production breaks.* Journal of Experimental Psychology: Applied, 2002. **8**(2): p. 118.
280. Sikström, S. and M.Y. Jaber, *The Depletion–Power–Integration–Latency (DPIL) model of spaced and massed repetition.* Computers & Industrial Engineering, 2012. **63**(1): p. 323-337.
281. Slovic, P., *The perception of risk*. 2016: Routledge.
282. Zhao, S., et al., *Risk perception and human behaviors in epidemics.* IISE Transactions on Healthcare Systems Engineering, 2018. **8**(4): p. 315-328.
283. Chen, F.H., *Modeling the effect of information quality on risk behavior change and the transmission of infectious diseases.* Mathematical biosciences, 2009. **217**(2): p. 125-133.
284. Funk, S., et al., *The spread of awareness and its impact on epidemic outbreaks.* Proceedings of the National Academy of Sciences, 2009. **106**(16): p. 6872-6877.
285. Kiss, I.Z., et al., *The impact of information transmission on epidemic outbreaks.* Mathematical biosciences, 2010. **225**(1): p. 1-10.
286. Nahl, D. and D. Bilal, *Information and emotion: The emergent affective paradigm in information behavior research and theory*. 2007: Information Today, Inc.
287. Pessoa, L., *On the relationship between emotion and cognition.* Nature reviews neuroscience, 2008. **9**(2): p. 148-158.
288. Altizer, S., et al., *Social organization and parasite risk in mammals: integrating theory and empirical studies.* Annual Review of Ecology, Evolution, and Systematics, 2003. **34**(1): p. 517-547.
289. Lloyd-Smith, J.O., et al., *Superspreading and the effect of individual variation on disease emergence.* Nature, 2005. **438**(7066): p. 355-359.
290. Bansal, S., B.T. Grenfell, and L.A. Meyers, *When individual behaviour matters: homogeneous and network models in epidemiology.* Journal of the Royal Society Interface, 2007. **4**(16): p. 879-891.
291. Klovdahl, A.S., *Social networks and the spread of infectious diseases: the AIDS example.* Social science & medicine, 1985. **21**(11): p. 1203-1216.
292. Sahneh, F.D. and C. Scoglio, *May the best meme win!: New exploration of competitive epidemic spreading over arbitrary multi-layer networks.* arXiv preprint arXiv:1308.4880, 2013.
293. Fishbein, M., S.E. Middlestadt, and P.J. Hitchcock, *Using information to change sexually transmitted disease-related behaviors*, in *Preventing aids*. 1994, Springer. p. 61-78.
294. Hadeler, K.P. and C. Castillo-Chávez, *A core group model for disease transmission.* 1994.
295. Grassly, N.C. and C. Fraser, *Mathematical models of infectious disease transmission.* Nature Reviews Microbiology, 2008. **6**(6): p. 477-487.
296. Baddeley, A.D., *Human memory: Theory and practice*. 1997: Psychology Press.
297. Wixted, J.T. and E.B. Ebbesen, *On the form of forgetting.* Psychological science, 1991. **2**(6): p. 409-415.
298. Steimer, T., *The biology of fear-and anxiety-related behaviors.* Dialogues in clinical neuroscience, 2002. **4**(3): p. 231.
299. Reeves, B. and C.I. Nass, *The media equation: How people treat computers, television, and new media like real people and places*. 1996: Cambridge university press.
300. Coval, M.L., *Analysis of Hill interaction coefficients and the invalidity of the Kwon and Brown equation.* Journal of Biological Chemistry, 1970. **245**(23): p. 6335-6336.

301. Wakefield, M.A., B. Loken, and R.C. Hornik, *Use of mass media campaigns to change health behaviour.* The Lancet, 2010. **376**(9748): p. 1261-1271.
302. Bagnoli, F., P. Lio, and L. Sguanci, *Risk perception in epidemic modeling.* Physical Review E, 2007. **76**(6): p. 061904.
303. Yang, C. and U. Wilensky, *Netlogo epidem basic model.* Center for Connected Learning and ComputerBased Modeling Northwestern University Evanston IL, pages Evanston, IL, 2011.
304. Spencer, S. and P.W. Jones, *Time course of recovery of health status following an infective exacerbation of chronic bronchitis.* Thorax, 2003. **58**(7): p. 589-593.
305. Peiris, J.M., L.L. Poon, and Y. Guan, *Emergence of a novel swine-origin influenza A virus (S-OIV) H1N1 virus in humans.* Journal of Clinical Virology, 2009. **45**(3): p. 169-173.
306. Rello, J., et al., *Intensive care adult patients with severe respiratory failure caused by Influenza A (H1N1) v in Spain.* Critical care, 2009. **13**(5): p. R148.
307. Archer, B.N., et al., *Interim report on pandemic H1N1 influenza virus infections in South Africa, April to October 2009: epidemiology and factors associated with fatal cases.* 2009.
308. Plennevaux, E., et al., *Immune response after a single vaccination against 2009 influenza A H1N1 in USA: a preliminary report of two randomised controlled phase 2 trials.* The Lancet, 2010. **375**(9708): p. 41-48.
309. Mossong, J., et al., *Social contacts and mixing patterns relevant to the spread of infectious diseases.* PLoS medicine, 2008. **5**(3).
310. Del Valle, S.Y., et al., *Mixing patterns between age groups in social networks.* Social Networks, 2007. **29**(4): p. 539-554.
311. Read, J.M., et al., *Social mixing patterns in rural and urban areas of southern China.* Proceedings of the Royal Society B: Biological Sciences, 2014. **281**(1785): p. 20140268.
312. Hurwitz, E.S., et al., *Effectiveness of influenza vaccination of day care children in reducing influenza-related morbidity among household contacts.* Jama, 2000. **284**(13): p. 1677-1682.
313. Bridges, C.B., et al., *Effectiveness and cost-benefit of influenza vaccination of healthy working adults: a randomized controlled trial.* Jama, 2000. **284**(13): p. 1655-1663.
314. Osterholm, M.T., et al., *Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis.* The Lancet infectious diseases, 2012. **12**(1): p. 36-44.
315. Gibson, C., *Population of the 100 largest cities and other urban places in the United States: 1790-1990*. 1998: US Bureau of the Census Washington, DC.