Investigating disease persistence in host vector systems: dengue as a case study

by

Bilal Ahmed Usmani

A thesis submitted for the degree of

Doctor of Philosophy (Ph.D.)

Institute of Biodiversity, Animal Health and Comparative Medicine (IBAHCM)

College of Medical, Veterinary and Life Sciences (MVLS)

University of Glasgow

June 2016

Abstract

The investigation of pathogen persistence in vector-borne diseases is important in different ecological and epidemiological contexts. In this thesis, I have developed deterministic and stochastic models to help investigating the pathogen persistence in host-vector systems by using efficient modelling paradigms. A general introduction with aims and objectives of the studies conducted in the thesis are provided in Chapter 1. The mathematical treatment of models used in the thesis is provided in Chapter 2 where the models are found locally asymptotically stable. The models used in the rest of the thesis are based on either the same or similar mathematical structure studied in this chapter. After that, there are three different experiments that are conducted in this thesis to study the pathogen persistence. In Chapter 3, I characterize pathogen persistence in terms of the Critical Community Size (CCS) and find its relationship with the model parameters. In this study, the stochastic versions of two epidemiologically different host-vector models are used for estimating CCS. I note that the model parameters and their algebraic combination, in addition to the seroprevalence level of the host population, can be used to quantify CCS. The study undertaken in Chapter 4 is used to estimate pathogen persistence using both deterministic and stochastic versions of a model with seasonal birth rate of the vectors. Through stochastic simulations we investigate the pattern of epidemics after the introduction of an infectious individual at different times of the year. The results show that the disease dynamics are altered by the seasonal variation. The higher levels of pre-existing seroprevalence reduces the probability of invasion of dengue. In Chapter 5, I considered two alternate ways to represent the dynamics of a host-vector model. Both of the approximate models are investigated for the parameter regions where the approximation fails to hold. Moreover, three metrics are used to compare them with the Full model. In addition to the computational benefits, these approximations are used to investigate to what degree the inclusion of the vector population in the dynamics of the system is important. Finally, in Chapter 6, I present the summary of studies undertaken and possible extensions for the future work.

Contents

List of Figures

List of Tables

Dedicated to my family

Acknowledgements

I would like to start-off by thanking Almighty ALLAH (SWT), the Most Gracious and the Most Merciful for blessing me to write this thesis. Secondly, I would like to express my sincere gratitude to my supervisor Prof. Daniel T. Haydon for his continuous support during my PhD study. I am indebted for his patience, motivation, and immense knowledge. His guidance helped me all the time in research and writing of this thesis. He taught me how to question thoughts and express ideas at a higher level. I have learned a lot from him and could not have imagined having a better supervisor and mentor for my PhD study.

I would sincerely like to pay my regards to Dr. Rebecca Mancy. I have been amazingly fortunate to have her as my second supervisor. I am always thankful for her help in technical matters as well as her moral support which cannot be quantified using any unit! Her support helped me stay sane through these difficult years. Her insightful comments and constructive criticisms at different stages of my research were thought-provoking and they helped me focus my ideas.

Besides my PhD supervisors, I would like to thank Louise Matthews for her help and support in issues related to the mathematical sections of Chapter 5. I would like to extend my thanks to Jamie and Rodney for many helpful discussions and ideas which helped me to widen my research from various perspectives. Special thanks is for Jamie who persuaded me to use the tau-leap algorithm.

I thank my fellow office mates for the stimulating discussions. In particular, I would like to thank Iain, Davidae, Robin, Julio, Becca, Yi-Hsiu, Dan, Asha, Laurie and Maude. I had a very good time with them during my stay at Glasgow. I am also thankful to Lorna and Florence for chocolates and helping me resolve the administrative issues.

I would like to thanks Sophie, Roman and Jan for many helpful discussions during the statistics laboratory sessions. It was really fun demonstrating with you all!

Last but not the least, I would like to thank my family: my parents, my uncles and to my brothers and sisters for supporting me spiritually throughout writing this thesis and my life in general. I would like to pay a special tribute to my wife and in-laws for their moral support during the course of my PhD.

Declaration

This thesis, and the work contained within it, is conducted by Bilal Ahmed Usmani. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The information derived from the work of others has been acknowledged in the text and a list of references is provided in the bibliography.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

_____

CHAPTER 1

Introduction

Introduction

1.1   Vector-borne diseases

Disease ecology focuses typically on the study of infectious disease within a population. It focuses on the interaction, behaviour and ecology of hosts with the biology of pathogens and draws on ideas from ecology, medicine, genetics, immunology and epidemiology. Pathogens or parasites are the agents that act as a medium for transmitting disease. They are transferred either directly by hosts or by contact between hosts, through different mediums including air, water or soil; or via biting arthropods (mostly mosquitoes and ticks, often referred to as vectors) (Kilpatrick and Altizer, 2012). Pathogens can infect single or multiple host species and can alter the within-host dynamics as well as the dynamics of the host population. Understanding the transmission and spread of the pathogen over space and time and its influence upon the host population forms the core of disease ecology.

Among the many different types of diseases, Vector-Borne Disease (VBD) are diseases spread in the host population by vectors. Vector-borne infections are defined in "Stedman's medical dictionary for the health professions and nursing" (Stedman, 2005) as:

> *"Class of infections transmitted by an insect or animal vector. The vector may merely be a passive carrier of the infectious agent, but many kinds of infectious agents undergo a stage in biological development in the vector. The vector, as well as the human host, is essential to the survival of the infectious agent."*

As defined by Magori and Drake (2013), the dynamics of VBD have three distinctive features: (i) marked seasonality, as the life cycle and vectorial capacity of most of the vectors is influenced by environmental factors; (ii) explosive outbreaks, especially if the pathogen is introduced into a naive environment; and (iii) sporadic annual outbreaks as vectors require a specialized habitat and successful amplification of disease, which depends upon different abi-

otic factors. In unfavourable situation, the occurrence of the disease in this case is infrequent or at irregular intervals.

Vector-borne diseases infect more than a billion people a year and kill more than one million per annum (Butler, 2013). Half of the world's population is currently at risk of these diseases and their prevalence is influenced by the geographical location, socio-economic status and living standards of the communities they circulate in (WHO, 2004). According to WHO, they account for 17% of the estimated global burden of all diseases. Malaria is the most lethal, causing 627,000 deaths in 2012 alone (Butler, 2013) whereas Dengue is the fastest growing, with a 30-fold increase in disease incidence over the last 50 years (WHO, 2014). In addition to affecting humans and animals, vector-borne diseases cause great loss to plants (McKirdy et al., 2002). They have the potential to cause enormous economic harm when livestock and crops become diseased and in the most extreme cases, limit local and global trade (Institute of Medicine, 2008). For example the Bluetongue virus, spread by midges, can severely harm livestock, particularly sheep, resulting trade and economic loss. Rift valley fever which can also affect humans, also exerts a heavy economic toll, with economic losses stemming from death and aborting among infected livestocks and treatment of infected humans.

1.2 Population and disease persistence

In population ecology, scientists are often interested in whether a community or species persists in the long or the short term and how changing ecological conditions alter persistence dynamics. The dynamics affecting the density or size of a population are important in studying the population ecology. This section provides an overview of persistence in the context of infectious disease ecology and its relation to host-vector models.

In epidemiological theory, greater attention has been paid to disease invasion thresholds as compared to persistence thresholds, due in large part to complications arising from the stochastic nature of persistence. Work on disease invasion is, therefore, well developed. In contrast, there are a range of definitions and theories for pathogen persistence and endemicity in a population [for example, see Castle and Gilligan (2012) and Nåsell (2005)]. As categorized by Mancy, the operational definitions of persistence can be grouped into three main categories, associated with studies using deterministic models, stochastic models, or a data driven approach (Mancy, 2015).

In the current work, persistence is characterized as continued existence of either infected or infectious individuals in either the host or vector population until a pre-defined target time

[see Mancy et al. (2013) for a similar approach]. The population dynamics of both host and vectors may be altered by the transmission and maintenance of the pathogen. On the other hand, the dynamics of the pathogen depend on those of the host and vector due to the interaction among the infected and infectious classes in both populations. Therefore, the nature of vector-borne disease requires a deep understanding of the ecology of hosts and vectors, in addition to considering the pathogens' ability to survive in the environment. Finally, the structure of the host population plays an important role in the spread of the disease. Many recent diseases have emerged from interactions between complex ecological communities that involve multiple hosts and parasites. It is suggested that disease prevalence may be altered by changing the structure of the host population (Collinge and Ray, 2006).

1.3 Mathematical modelling in infectious diseases

Mathematical techniques involved in understanding and forecasting the spread of infectious diseases draw on techniques developed from numerous areas like dynamical systems, stochastic processes, numerical computing and optimization theory. Typically, building an infectious disease model requires information about demographic or biological characteristics of the host and the pathogen and whether pathogen transmission is dependent upon the physical surroundings. The distribution, heterogeneity and structure of the host population are sometimes important to take account of realistic behaviour in the modelling structure. These models can be linked to observational studies or can be used to test different hypotheses or to forecast patterns of disease prevalence (Magori and Drake, 2013).

Early models of infectious disease were disease-specific, often deterministic in nature, and focused on addressing problems like finding the peak and final size of the epidemic and explaining the effect of vaccination on disease spread considering a large and homogeneously mixed population. They were then generalized to account for more realistic disease patterns by including contact heterogeneity, setting up multiple populations in a community and allowing seasonal variations [see (Reiner et al., 2013) for an overview of the mathematical models developed so far for mosquito-borne pathogen transmission]. Other generalizations of simple models were stochastic epidemic models used to answer additional questions like the probability of a major outbreak and the persistence time of the disease [see (Ditlevsen and Samson, 2013) for an overview of stochastic models]. A detailed summary of mathematical models used for modelling different types of host-vector systems is provided in the beginning of Chapters 2-5.

1.4 Introduction to dengue disease

This section provides an overview of dengue disease and dengue fever, a symptom caused by dengue virus. The maps highlighting the global distribution of dengue are shown in Figures 1.1, 1.2 and 1.3 to portray the distribution of the disease. Then the geographical location, habitat, and entomology of the dengue vectors is presented. The main factors that may impact on the transmission and persistence dengue are discussed at the end.
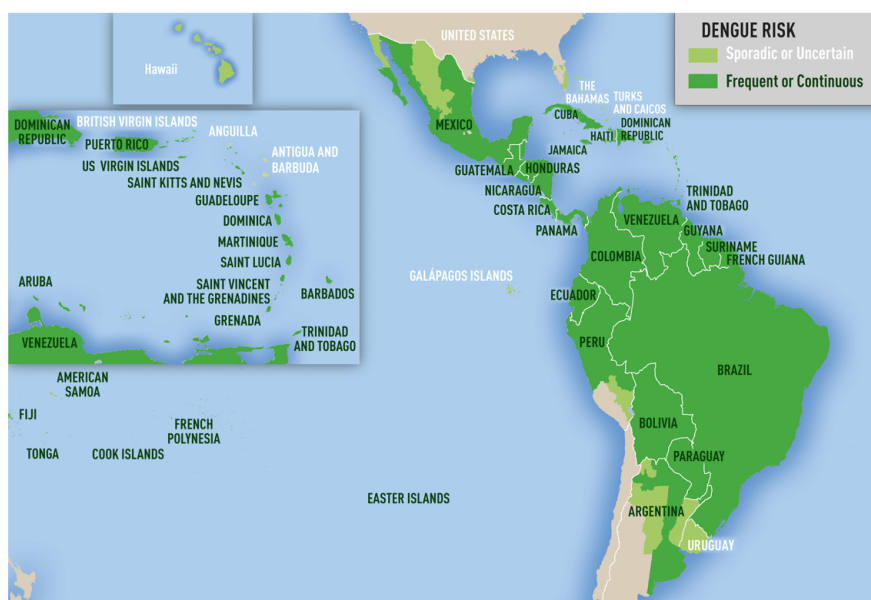
MAP 3-1. DISTRIBUTION OF DENGUE IN THE AMERICAS AND THE CARIBBEAN[1][2]

[1] Risk areas are shown on a national level except for where evidence exists of different risk levels at sub-national regions. Areas that are too small to be seen on the regional maps are labeled in white or gray depending on their risk categorization.

[2] Based on surveillance data, official reports, published research, and expert opinion, including data from Brady et al. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. PLoS Negl Trop Dis 6(8): e1760 doi: 10.1371/journal.pntd.0001760 (2012). It was compiled by the CDC Dengue Branch in collaboration with the University of Oxford.

Figure 1.1: Distribution of dengue in the Americas and the Caribbean. (http://wwwnc.cdc.gov)

Dengue virus is a member of the Flaviviridae family of viruses having four distinct serotypes commonly knowns as (DENV1-4), where DENV stands for the dengue virus. The global incidence of dengue has grown dramatically in the recent decades. It is estimated that about 2.5 billion of the worlds population is at risk of dengue disease and it is endemic in more than 100 countries (De Benedictis et al., 2003; Erlanger et al., 2008; Simmons et al., 2012; WHO, 2009). Figure 1.1, 1.2 and 1.3 (*taken from* http://wwwnc.cdc.gov *at the end of 2015*) shows

dengue risk maps developed in 2012 which portray the possible global spread of the disease in coming years. Dengue is an infectious tropical disease whose transmission mechanism falls in three distinct and intersecting spheres (human, mosquito and virus sphere) and is spread by the bite of a mosquito from the *Aedes* family, chiefly by *Aedes aegypti* and *Aedes albopictus* (Kyle and Harris, 2008). The latter mosquito is considered a dengue maintenance vector in many parts of Asia and Europe.
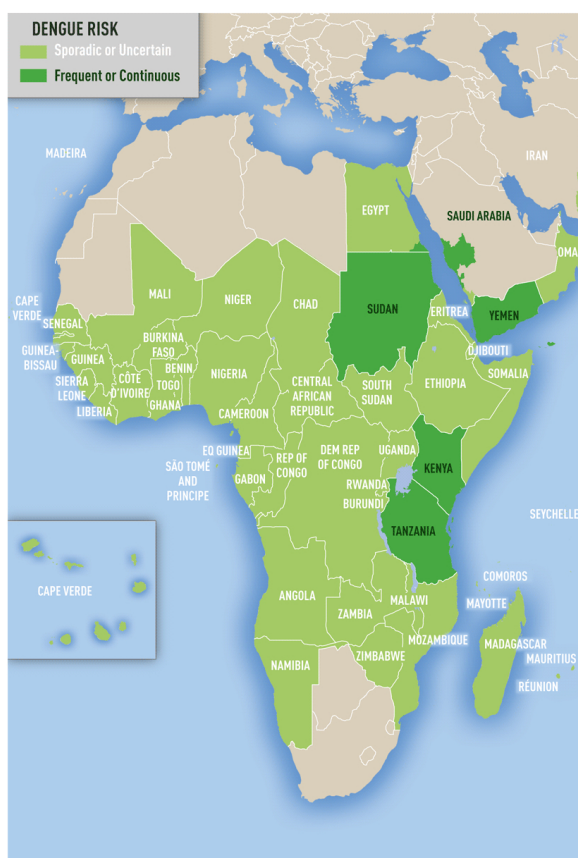
MAP 3-2. DISTRIBUTION OF DENGUE IN AFRICA AND THE MIDDLE EAST[1] [2]

[1] Risk areas are shown on a national level except for where evidence exists of different risk levels at sub-national regions. Areas that are too small to be seen on the regional maps are labeled in white or gray depending on their risk categorization.

[2] Based on surveillance data, official reports, published research, and expert opinion, including data from Brady et al. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. PLoS Negl Trop Dis 6(8): e1760 doi: 10.1371/journal.pntd.0001760 (2012). It was compiled by the CDC Dengue Branch in collaboration with the University of Oxford.

Figure 1.2: Distribution of dengue in Africa and the Middle East. (http://wwwnc.cdc.gov)

The global expansion of dengue in the twentieth century is believed to have occurred during World War II when infected people took the virus to Pacific areas. Transport of goods (tires, vehicles, etc.) used in the course of war helped establish vector populations in many parts of the world (Esteva and Vargas, 1999). The dissemination of the disease was enhanced after the

war as human population growth lead to poor sanitization, domestic water storage issues, and over crowding. These problems arise due to population growth provided favourable conditions for the breeding of *Aedes aegypti* and dengue virus, especially in parts of Asia where there are more places for water to stand. Since then, ever increasing use of food containers (e.g. tinned food and plastic utensils), used or discarded plastic products and discarded tyres have served as potential breeding sites for the mosquitoes. The world reporting map for dengue, (http://www.healthmap.org/dengue/) shows the dengue incidence in the world until the end of November 2015 (Figure 1.4).
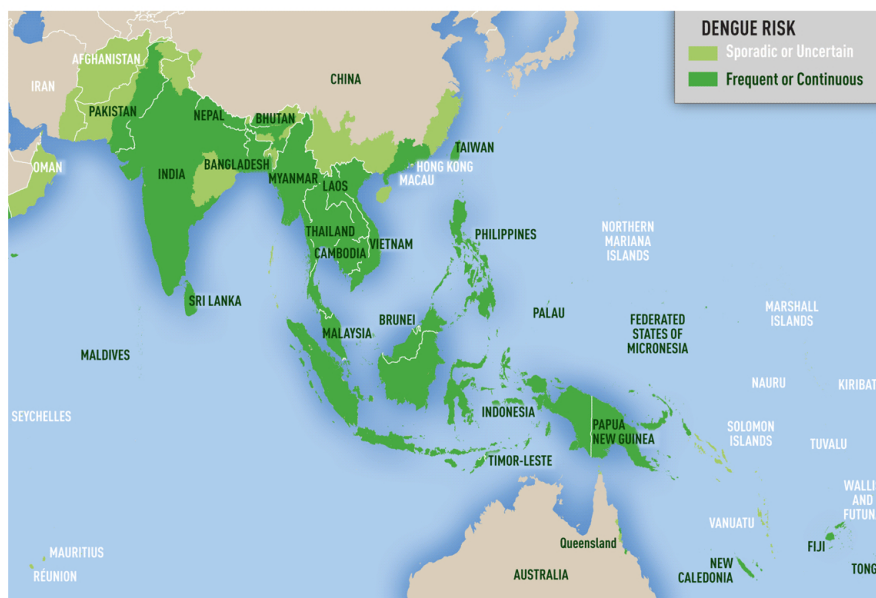
MAP 3-3. DISTRIBUTION OF DENGUE IN ASIA AND OCEANIA[1][2]

[1] Risk areas are shown on a national level except for where evidence exists of different risk levels at sub-national regions. Areas that are too small to be seen on the regional maps are labeled in white or gray depending on their risk categorization.

[2] Based on surveillance data, official reports, published research, and expert opinion, including data from Brady et al. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. PLoS Negl Trop Dis 6(8): e1760 doi: 10.1371/journal.pntd.0001760 (2012). It was compiled by the CDC Dengue Branch in collaboration with the University of Oxford.

Figure 1.3: Distribution of dengue in Asia and Oceania. (http://wwwnc.cdc.gov)

Dengue fever (DF), a symptom caused by DENV induces a severe pain and flu-like illness in humans with a small proportion of patients developing Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) which can be fatal in some cases, especially for children. In 1999, the WHO divided the illnesses caused by DENV in two categories: dengue and severe dengue (WHO, 2009). Dengue includes three phases: a febrile phase, a critical phase and a recovery phase. The febrile phase includes high fever, dehydration, body aches and
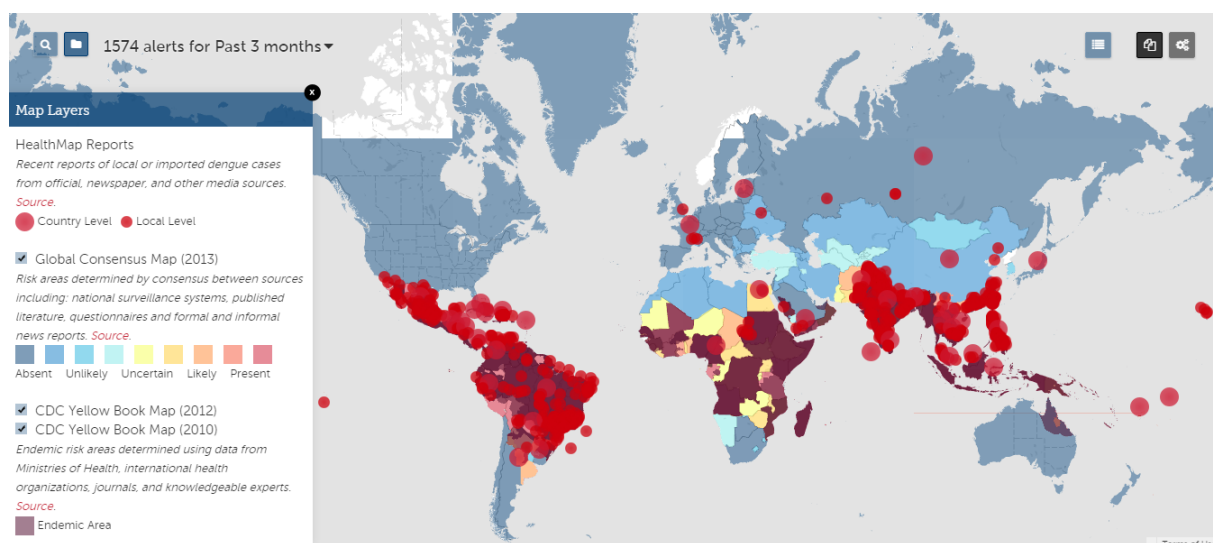
Figure 1.4: Dengue reporting map of the World for 2015. (http://www.healthmap.org/dengue/)

neurological problems. The critical phase involves haemorrhage and organ impairment, in addition to low blood platelet count. The last phase, recovery, includes re-absorption of fluids. Severe dengue occurs as a result of shock due to plasma leakage, severe bleeding or when severe organ impairment takes place. The old terms DHF and DSS are still commonly used to describe dengue fever.

The mosquito vector becomes infected by taking a blood meal from an ill person suffering from dengue disease. In diseased mosquitoes, the viral infection establishes in their organs, especially in the salivary glands and central nervous system. The virus alters the feeding behaviour of the mosquitoes and it is reported that both probing and blood meal timings are increased in the infected mosquitoes as compared to uninfected ones (Platt et al., 1997). This results in multiple attempts to feed during the gonothropic cycle–which comprises of blood feeding, egg maturation and ovipoisiton.

Infection by DENV requires an incubation period in both mosquitoes and humans. The infected mosquito is able to transmit the dengue virus after the process of virus maturation commonly known as the extrinsic incubation period (EIP). The estimated extrinsic incubation period has been shown to vary with respect to serotype and temperature (Chen and Hsieh, 2012). Intrinsic incubation period (IIP) is the time required for virus maturation in host. The average intrinsic incubation period is estimated to be 5.9 days. After that humans become viremic (virus in the blood stream) and are infectious to the vector (Chan and Johansson, 2012), which will spread the infection to another host. Recovery from infection caused by one serotype provides life-long immunity to that serotype but due to antigenic diversity of dengue, patients have limited cross-immunity against remaining serotypes. To date, there is

no vaccine for dengue fever so vector control is the most efficient way to control and minimize dengue disease.

1.5 Vectors of dengue virus

There are two main vectors of dengue virus, (i) *Aedes aegypti* and, (ii) *Aedes albopictus*. Although the WHO reports describe two more vectors from the *Aedes* genus for dengue transmission, *Aedes polynesiensis* and *Aedes scutellaris* (WHO, 2009), their role is very limited in the spread of dengue. Both *Aedes aegypti* and *Aedes albopictus* have similar life cycles, beginning with eggs laid in standing water that hatch to become larvae then grow to pupal stage. Finally the mosquito emerge as adults (winged form) from the water. All development stages are sensitive to climate and influenced by the suitability of the habitat. Both mosquitoes are weak-fliers, live below an altitude of 1000 metres, bite during early morning and evening (WHO, 2009) and able to live in natural as well as man-made environments. Adults can reproduce immediately after hatching. The blood feeding patterns of both species show that they both almost exclusively feed on humans, with very few cases of multiple feedings on cats, dogs and swine.This feeding behaviour is reported in different studies (Kamgang et al., 2010; Ponlawat and Harrington, 2005; Valerio and Marini, 2010). Females take a blood meal within the first 2-3 days following emergence, which is vital for the development of eggs. Both species usually take more than one blood meal on multiple persons during their gonothropic cycle (Paupy et al., 2010). Size and survival of adults, length of gonothropic cycle, and the speed of virus replication depends heavily on the temperature (Barbazan et al., 2010; Focks et al., 2000; Yang et al., 2009). The next section briefly discusses the geographical location, habitat, and entomology of both vectors.

1.5.1 The vector *Aedes aegypti*

Aedes aegypti is the primary vector for DENV transmission and is found through tropical and subtropical region of the Americas, Africa, Asia, as well as the south Eastern U.S., the Indian Ocean Islands, Northern Australia (Kraemer et al., 2015) and Europe (Medlock et al., 2012). It is transported via increasing trade around the world. There is a great deal of literature related to the behaviour and the habitat of *Aedes aegypti* since it is a well studied vector of several infectious diseases such as Chikungunya, Yellow fever as well as Dengue (Paupy et al., 2010) and (Christophers, 1960). The species is highly adaptable to new environments and tends to live near human hosts. *Aedes aegypti* feed almost exclusively on humans and many major disease epidemics in the world are caused and maintained by this species. It

can't survive in low temperatures, except in sheltered sites. With the recent invasion of *Aedes albopictus*, *Aedes aegypti* have been pushed into different parts of the world in search of suitable habitat. This migration, may play a role in the dramatic increase in the incidence of dengue disease in many parts of the world where dengue was not previously present. Figure 1.5 shows a female *Aedes aegypti*.

Figure 1.5: *Aedes aegypti* (Taken from http://rdontheroad.wordpress.com)

Aedes aegypti has a very limited flight range so the eggs are likely to be found in the vicinity of the hosts. Their flight range was determined in mark-release-recapture experiments where they showed maximum dispersal distance of 500 metres and dispersal distance was independent of sex of the mosquitoes. Almost 75% of mosquitoes are found at the place where they emerge as an adult. Therefore, it was suggested that people are responsible for dengue virus transmission between communities (Harrington et al., 2005). Moreover Adams and Kapan (2009) considered frequency dependent biting in a metapopulation and concluded that the pathogen is maintained at reservoirs of infection due to distribution of mosquito population and variability in human travelling patterns. Both studies highlighted the inability of this species to carry virus at long distances.

Figure 1.6: *Aedes albopictus* (Taken from http://glacvcd.org)

1.5.2   The vector *Aedes albopictus*

Aedes albopictus (Figure 1.6), known as the Asian tiger mosquito is found in eastern Asia, India, Japan, and several islands in the Pacific (Australasia). It has also spread in Italy and other regions in the Mediterranean basin, as well as parts of Africa, Brazil, Central America, the Caribbean, and most of the United States (East coast and the Midwest). Due to its global spread it is included among the 100 world's most invasive species. It is assumed that this vector is transmitted around the world via trade (especially tires and lucky bamboo plant). Due to its superior interspecies competition, resistance of eggs for cold weather, higher survival rate, flexible breeding biology, adaptation in natural and artificial habitats / environments, and broad host range, it is displacing *Aedes aegypti*.

It was initially considered to have less dengue competence than *Aedes aegypti*, but recent studies documented its role as a maintenance vector (that maintains pathogen transmission) for dengue outbreaks, especially in the absence of *Aedes aegypti* in different parts of the world including Europe [see (Enserink, 2008; European Centre for Disease Prevention and Control, 2011; Gratz, 2004; Medlock et al., 2012; Paupy et al., 2010)]. Richards et al. (2006) reported that it can be a principal vector for dengue under right circumstances, especially in the absence of *Aedes aegypti* . According to laboratory studies,*Aedes albopictus* can act as a transmission vector for at least 22 different arboviruses including all four serotypes of

dengue (Gratz, 2004). Its potential role as a dengue vector is receiving much attention in dengue free temperate regions of the world where it has established its colonies. It is an opportunistic feeder having strong anthrophilic nature. Multiple feeding in one gonothropic cycle and biting at a rate of 30-48 bites per hour (Cancrini et al., 2003) greatly enhances its ability to transmit virus in humans. Although being exophilic (tends to inhabit outdoors), *Aedes albopictus* rarely disperses above 300 metres. In mark-release-recapture experiments (Marini et al., 2010), most adults are found within 50-100 metres range of their emergence site and in general the entomology literature, their flight range broadly falls within 200-400 metres.

1.6 Possible factors affecting dengue transmission

In the current times, dengue is endemic in more than 100 countries. Its spread is mainly due to poor vector control, climate change and overpopulation which leads to many health and environmental issues. Sometimes the economic and political condition of a region leads to the disease spread (for example, the wars and related instability of the government leave little or no interest for the fund allocation for combating the disease). In countries where dengue is endemic, employing biological and chemical control independently is not a very successful approach in stopping disease transmission (Simmons et al., 2012). The main reasons include the limitations and the resulting environmental impact of these approaches. For example, the release of infectious genetically modified male mosquitoes in the wild female population was found more effective than vector control based on insecticide use, (Alphey N, Alphey L, Bonsall MB (2011) **A model framework to estimate impact and cost of genetics-based sterile insect methods for dengue vector control**. PLoS One 6: e25384), but the authors ignore the effect of seasonality, which is essential for the variation in the dengue incidence during a year. Similarly, the introduction of larvivorous fish (Poecilia reticulata) into water storage containers is proven successful in Cambodia, but as a community-based vector control tool. (Seng CM et al. **Community-based use of the larvivorous fish Poecilia reticulata to control the dengue vector Aedes aegypti in domestic water storage containers in rural Cambodia**. Journal of Vector Ecology, 2008, 33:139144. doi:10.3376/1081- 1710(2008)33[139:CUOTLF]2.0.CO;2 pmid:18697316 , Seng CM et al. **The effect of long-lasting insecticidal water container covers on field populations of Aedes aegypti (L.) mosquitoes in Cambodia**. Journal of Vector Ecology, 2008, 33:333341. doi:10.3376/1081-1710-33.2.333 pmid:19263854). Erlanger et al. (2008) argue that dengue vector control is effective in reducing vector populations when interventions use a community-based, integrated approach tailored to the local eco-epidemiological

and socio-cultural settings.

Due to the continuous increase in the global health burden of dengue, scientists are keen to better understand the persistence of dengue virus and vector. Vectors are either establishing new territories or re-establishing previously infected geographic regions causing frequent epidemics despite control efforts. One of the main challenges of modelling dengue dynamics is related to the persistence of virus and vectors in the inter-epidemic time and in extreme weather. Since both *Aedes aegypti* and *Aedes albopictus* live in close proximity to human hosts, they are able to use dark and shaded areas in man-made environments to mitigate the effects of extreme weather. The eggs of *Aedes albopictus* are capable of over-wintering by dipause, and can resist freezing temperatures (Medlock et al., 2012) and (European Centre for Disease Prevention and Control, 2012). These adaptations and the continued spread and recolonization of geographic regions by vectors aid the maintenance of mosquitoe population during adverse climate but are not enough to characterize disease persistence in a region.

Dengue virus is maintained in a population in an endemic state or by continual reintro-ductions/epidemics. There can be many factors affecting the transmission of dengue virus; broadly, transmission by vectors is favoured by the tropical temperature (Chen and Hsieh, 2012), type of virus strain (Anderson and Rico-Hesse, 2006) and suitability of habitat (Jansen and Beebe, 2010). Transmission in humans is dependent upon the host demography, spatial distribution in a region and frequency of commuting (Adams and Kapan, 2009; Andraud et al., 2012; Liebman et al., 2012; Stoddard et al., 2009). The factors discussed in this section are those which are commonly found in the literature for modelling dengue. Some of these are included in Chapter 4, which is dedicated for the mathematical modelling of dengue transmission. Based upon the complexity arise in modelling and research questions of interest, the impact of factors like seasonality, migration of an infectious human to the naive population and the effect of seroprevalence on the re-introduction of dengue are explored in that Chapter. In the next sections, a brief introduction of some of the factors affecting the transmission of dengue in both host or vector populations is given.

1.6.1   Vectorial capacity of *Aedes* mosquitoes

Vectorial capacity (C) is defined as the number of new infections disseminated per case per day by a vector. It is a broad term encompassing vector competence; which can be thought of as the ability of a vector to transmit disease. Vectorial capacity depends on the vector density (relative to its vertebrate host), frequency of blood meals, daily survival probability and extrinsic incubation period of virus. Mathematically it can be defined as a function of

the above variables and can be regarded as an efficiency measure for vector-borne disease transmission. Its formula is (Smith and McKenzie, 2004):

$$C = \frac{ma^2 b_H b_V p^n}{-ln(p)}.$$

Where $m = V/H$ is the relative density of vectors (V) with respect to hosts (H), a is the proportion of vectors feeding on a host divided by the length of gonotrophic cycle in days, $b_H(b_V)$ is the per bite probability of transmission from the vector (human) to human (vector), p is the daily survival rate of vectors, and n is the extrinsic incubation period. In the above formula, vector competence is the product of b_H and b_V. Vectorial capacity is favoured by high bite rate and vector density with respect to hosts and decreased by higher vector mortality rate and longer incubation period in vectors. Since infection requires dual contact between host and vector, bite rate appears squared to reflect the dual transmission of infection.

Aedes aegypti is considered a highly effective species in acquiring, maintaining and transmitting the virus. Historically it is associated with many dengue outbreaks throughout the world. The vectorial capacity of *Aedes aegypti* for different dengue strains is estimated by Anderson and Rico-Hesse (2006) which indicates a more virulent Southeast Asian (SEA) genotype of dengue which displaces the America (AM) genotype of dengue serotype 2 in several countries. Viral replication in the midgut of *Aedes aegypti* of the former serotype was significantly higher resulting in a 2- to 65-fold increase in the vectorial capacity and is more likely to cause the severe form of dengue. On the other hand, evidence mounts for the role of *Aedes albopictus* as a principal vector for dengue virus. In Central Africa, Paupy et al. (2010) conducted a study which portrays *Aedes albopictus* as the major vector for dengue virus as it was repeatedly found to be infected with DENV in Libreville, Gabon, Africa where no infected *Aedes aegypti* were detected. In an outbreak of dengue in Hawaii, Effler et al. (2005) observed the presence of *Aedes albopictus* in all the affected communities whereas *Aedes aegypti* were not detected whereas Vega-Rua et al. (2013) found *Aedes albopictus* with unexpectedly high susceptibility and high efficiency of transmission for some dengue strains in the Southeast of France. Evidence as to whether it is a main vector or not remains equivocal, it is however evident that *Aedes albopictus* possess almost a similar threat level for dengue outbreak as *Aedes aegypti* in many parts of the world.

Virus ingestion amount and type of virus strain are important factors which influence the vectorial capacity. A study hypothesizing that endemic DENV strains are more efficient at infecting urban populations of both *Aedes aegypti* and *Aedes albopictus* as compared to ancestral sylvatic (wild type) DENV strains (Moncayo et al., 2004). Different but high susceptibility levels (94% and 69%) are reported in both species in case of epidemic/endemic DENV 2 strain that significantly reduced infection levels compared to the sylvatic strain.

Virus dissemination of dengue strains are similar for both vectors. Richards et al. (2012) conducted experiments in the Florida Keys measuring the rates of infection, virus dissemination and transmission of virus in *Aedes aegypti* and *Aedes albopictus*. Similar levels of infection, dissemination and competence levels are found in both vectors.

1.6.2 Seasonality

Seasonality has major effects for the spread and persistence of dengue. Mosquito populations are bounded inside favourable seasonal conditions like humidity, rainfall amount, wind speed and temperature ranges. In particular, temperature and rainfall are the main variables that account for the change in Extrinsic Incubation period (EIP) of the mosquitoes. In most regions in the world temperature greater than $35\,^{\circ}\text{C}$ reduces the survival probability of the adult mosquitoes (Focks et al., 2000). In Saifur et al. (2012), it was reported that heavy rainfall decreases dengue transmission due to population losses of *Aedes* aquatic population. Most epidemics breaks are observed in the middle or at the end of the rainfall season as the ratio of mosquito-to-human is usually high in most of the tropical countries during that season. However, seasonality alone cannot determine the mosquito abundance and dengue occurrence as *Aedes* mosquitoes can create their own micro-environments due to close association to humans. Climate factors use average values which can differ for swarms of mosquitoes living in diverse local environments. In Puerto Rico, Johansson et al. (2009) found a positive and statistically significant relationship between monthly changes in (i) temperature and (ii) precipitation, with monthly changes in dengue transmission and further concluded that spatial heterogeneity can vary this relationship.

Temperature has a major effect on the mosquito life cycle. Under normal conditions, they cannot survive in extremely low and high temperatures. Temperature can also effect the gonothropic cycle and size of female mosquitoes as well as hatching of the mosquito eggs in to adults (Focks et al., 2000). In the same study Focks et al. (2000) have shown that the ideal temperature for the transmission of dengue virus falls between $20\,^{\circ}\text{C}$ to $35\,^{\circ}\text{C}$, with temperature greater than $35\,^{\circ}\text{C}$ eliminating the possibility of adult population existence as aggregate survival of eggs and larval and pupal stages of mosquitoes are insufficient. Richards et al. (2012) measure vector competence for both *Aedes* species at two different laboratory temperatures ($28\,^{\circ}\text{C}$ and $30\,^{\circ}\text{C}$) and observe optimal transmission at $28\,^{\circ}\text{C}$.

The vector population is affected by changing seasons in many ways. In the wet season, more dengue cases are reported compared to in the dry / winter seasons and Harrington et al. (2005) reported *Aedes aegypti* covering greater dispersal distance in rainy season. Rainfall

and wet season also affects the life cycle and transmission efficiency of the *Aedes* mosquitoes but their relation is unclear in these studies. As remarked in a review paper, the relation between rainfall and the mosquito abundance is dependent on the lifestyle of the localities and their primary sites of water storage (Jansen and Beebe, 2010). The amount of rainfall after a certain level can cause rain water to stand at places within the human habitats. This provides an excellent breeding ground for the mosquitoes and hence an increased number of dengue cases. So in localities with poor sanitary conditions and lack of good administration, there are more chances of dengue outbreaks. Hii et al. (2012) reported that heavy rainfall creates abundant outdoor breeding sources for *Aedes* in the long run, but dry spells in some settings trigger an increase in water storage containers which can serve as breeding habitats. The effects of seasonality are implicitly modelled in the seasonal model of dengue in later Chapter.

1.6.3   Human movement and demography

As mentioned is section 1.3, the short flight range of both *Aedes aegypti* and *Aedes albopictus* is clearly documented (Harrington et al., 2005). An obvious question is how dengue can spread and persist in many parts of the world with vectors that are weak fliers. It has been established and argued by several scientists that humans are the central source of resurgence and maintenance of vector-borne pathogens, particularly for dengue virus (Harrington et al., 2005; Stoddard et al., 2009) and Adams and Kapan (2009). Rapid and unplanned urbanization, distribution of community birth / death rate and age , extensive commuting nationally and internationally, and dramatic redistribution of populations in cities are some of the factors emerging as major sources of short and long term dengue spread and persistence.

The effects of human movement are clearly reported in an extensive study by Harrington et al. (2005). Mark-release recapture experiments were performed on *Aedes aegypti* dispersal in two countries Puerto Rico and Thailand for eleven years. Results showed that people rather than mosquitoes are the major cause of dengue virus dissemination within and in between communities. The importance of human movement and demography on the spread of dengue is also highlighted in (Kyle and Harris, 2008). The movement of an infectious human in naive human or naive vector populations is investigated in Chapter 4 for exploring dengue outbreak dynamics.

1.6.4   Immunity and cross immunity of hosts

Immunity in the context of dengue refers to the ability of the body to resist a re-infection due to the development of an antigen-specific antibodies. When a human becomes infected, its body launches the immune response which in turn neutralizes the pathogen by producing antigen specific antibodies. Most diseases are generated by one strain of a spectrum of closely related pathogens. When one serotype invades the population it provides temporary immunity to the other strains. This immunity eventually wanes after some time, making the host susceptible to reinfection with another strain (Feng and Velasco-Hernández, 1997).

In the course of a dengue infection the human body produces antibodies against the disease which usually last lifelong. The homologous Immunoglobin G (IgG) antibodies, produced by memory B cells serve this purpose. In addition, the human body also provides temporary or short lived partial cross immunity against the remaining serotypes. This period can last from some months to a couple of years depending upon the serotype. After that, a secondary infection can lead to what is called Antibody Dependent Enhancement (ADE). ADE occurs when cross-reactive heterotypic IgG antibodies generated by a prior infection wane to levels that no longer neutralize the heterotypic virus and instead of preventing infection, the binding of antibodies to virus at subneutralizing concentrations can lead to enhanced viral replication resulting in more severe dengue (Wearing and Rohani, 2006). In the same study, the authors include all serotypes of dengue in their model and introduced seasonal variation in the recruitment of vectors. They explore the effects of temporary cross-immunity, ADE, and variation in serotype virulence on persistence and eradication on certain serotypes while establishing their effects on transmission and mortality of vectors. Results are compared with a long term dengue clinical study on Thai children and they suggest that to generate infection time series that corresponds with the data, a combination of seasonal variation in the vector demography and a short lived period of cross immunity is sufficient. Serotype extinction due to vector competence is explained in Anderson and Rico-Hesse (2006). They prove that more viremic SEA dengue strains replace less virulent AM dengue strains causing different immune responses in humans.

1.6.5   Transovarial route of transmission

Initially it was considered that DENV only transmits between humans and mosquitoes (horizontal transmission) but now there is a growing evidence of vertical transmission of dengue virus in vectors from different parts of the world. In India, Joshi et al. (2002) documented transovarial transmission until the seventh generation of *Aedes aegypti* where first generation

17

emerges from hundred percent DENV 3 positive parents. Reporting more infected females in numbers than males, they found high larval mortality and low fertility in infected mosquitoes when compared to controls. They suggested that mosquitoes may act as a potential reservoir for dengue virus, but that rate of vertical transmission found in the field seems low for a heterogeneous population. The decrease in number of eggs hatching into adults as transovarial transmission increases is reported by Joshi and Sharma (2001) where they indicate that via this route, virus persists in optimal numbers of best-selected individuals by virtue of their genetic superiority. Conversely, vertically acquired infection in mosquitoes also acts as a biological control for their population. In Oaxaca, Mexico, Günther et al. (2007) found strong support for vertical transmission in mosquitoes for DENV 2-4 viruses. In Brazil, at Pampulha region of Belo Horizonte, authors used minimum infection rate (MIR) to confirm vertical infection in *Aedes albopictus* for DENV 2 (Cecílio et al., 2009), and in Fortaleza, Ceara region, the natural evidence of the vertical transmission of DENV 2 and DENV 3 is reported in both species (Martins et al., 2012) using MIR. In Surabaya, Malaysia, (Mulyatno et al., 2012) found transovarially infected *Aedes aegypti* during wet and dry seasons for DENV 1 and DENV 2 viruses and concluded that this route plays an important role in the virus maintenance in nature and in humans, especially in the rainy seasons. The efficacy of the vertical transmission is discussed in (Adams and Boots, 2010) by using a mathematical model. They concluded that the role of vertical transmission requires more evidence to understand its impact on dengue persistence.

As for vertical transmission of dengue virus in humans, there are some clinical studies reporting dengue transmission in infants from infected pregnant mothers but its occurrence is very rare (Chye et al., 1997; Fatimil et al., 2003) and (Tan and Rajasingam, 2008). Non-vector methods of dengue transmission in humans are discussed in Chen and Wilson (2004) in which authors review different publications highlighting needlestick injuries, bone marrow transplantation, and intrapartum and vertically acquired infection as a source of dengue virus transmission.

1.6.6 Host preference of vectors

In general mosquitoes take blood meals from different hosts ranging from primates and bovids to rabbits and mice. The understanding of the frequency of contact between mosquitoes and hosts is critical for the transmission dynamics of any arthropod borne disease, such as dengue. For humans, it is a common experience that mosquitoes bite some people more than others. This heterogeneous biting mechanism is a result of factors including body size, time spend indoors and the residential status of the individual (Harrington et al., 2014; Liebman

et al., 2014). Understanding the feeding behaviour of the mosquito can help develop targeted mosquito repellents and that could potentially save many lives.

Out of all available hosts, *Aedes aegypti* and *Aedes albopictus* prefer to feed on humans (Kamgang et al., 2010; Richards et al., 2006). These mosquitoes show a considerable variability in the biting behaviour and selection of human hosts. For example *Aedes aegypti* tends to feed on young adults and males. Sometimes the hosts can be selective. De Benedictis et al. (2003) have shown that three people accounted for 56% of meals in 22 houses in Florida, Puerto Rico. Like *Aedes aegypti*, *Aedes albopictus* exhibits a strong host preference; Richards et al. (2006) show that out of 40 human blood meal samples collected from *Aedes albopictus* mosquitoes in North Carolina, 80% of meals are from a single human. They remark that since *Aedes albopictus* fed predominantly on mammalian hosts (83% of the samples), this species can be a potential vector for disease transmission among mammals.

1.7 Research aims

The broader aims for this study are to investigate pathogen persistence in host-vector systems by using efficient modelling framework. This aim is achieved by conducting three different studies in Chapters 3, 4 and 5. In each piece of work, different nested compartmental models for host and vector populations are considered for a smooth methodological transition and better result comparison. These frameworks are used for answering several questions based on the processes and structure of the system under consideration in that specific study: (i) examining pathogen persistence using stochastic systems in host population, termed as the Critical Community Size (CCS), and make a conjecture for the existence of such persistence in the vector population; (ii) finding the determinants of CCS; (iii) investigating the impact of seasonality on the persistence of dengue; (iv) construction of efficient stochastic models that approximate the dynamics of the full host-vector model; (v) finding the Quasi-Stationary Distribution (QSD) for a host-vector system using the stochastic framework, as a counterpart of deterministic endemic equilibrium. In all of the studies conducted in this work, the main theme is to quantify pathogen persistence in terms of host-population size in host-vector systems and using a minimalist modelling framework to achieve this goal. As a result, the concepts related to modelling the host-vector system and pathogen persistence reappear in different parts of the thesis. Moreover, the term persistence used in this work refers to the pathogen's persistence in both host and vector populations. The particular research questions addressed in each chapter are described at the beginning of that study and the overall contribution of studies conducted in each chapter is presented in the final chapter of the thesis.

1.8 Thesis overview and contribution

The central theme of this thesis is to develop deterministic and stochastic models to investigate pathogen persistence in host-vector systems. Therefore, this study requires me to quantify the pathogen persistence and develop, apply and extend the existing modelling methods for this population structure. Models are constructed by using a minimalist approach and efficient algorithms are written to explore the relationship between the community size and the extinction of disease. As a motivational application, a stochastic dynamic seasonal host-vector model for dengue is constructed by using the analysis developed. A chapter is devoted to developing methods to approximate the population dynamics of vectors as this work strongly focuses on the construction of simple and powerful models. The overall work undertaken in this study broadly falls in the areas of infectious disease ecology and mathematical biology. The research areas in which the current study has potential contribution are shown in Figure 1.7. The thesis comprises of a series of semi-independent chapters which are written in 'extended paper form'. Background and important core concepts are discussed at the beginning of the present chapter, followed by the introduction of dengue disease and possible factors which have an effect of the transmission cycle. Chapter 1 acts as an preamble to the rest of the work and provides the necessary background information to the reader. Chapter 2 can be thought as an introductory chapter to mathematical models with standard techniques used to explore the dynamics of host-vector systems in general. As mentioned in the research aims, the next three chapters are the main studies of this thesis. Each main chapter addresses a specific scientific problem and starts with the literature review, followed by a section describing the models and methodology employed. Findings and main results are documented and commented in each chapter in the results and discussion section. Although, writing a thesis in this way results in repetition of theoretical concepts, which are mostly present in the sections describing the models, the reader can progress rapidly through these sections of the thesis. Table 1.1 provides a quick overview of the work undertaken in each chapter and a brief summary of chapters is provided in the following sections.

Chapter 2: Persistence thresholds and stability in deterministic models for host-vector systems

Chapter 2 deals with the mathematical modelling of host-vector systems. The chapter starts with an overview of work performed using compartmental models in the deterministic settings. A brief overview of the Ross Macdonald model and its assumptions is presented. Shortcomings of this model are addressed in the development of subsequent models. In order to assess the dynamic behaviour of host-vector system in the context of the Ross Macdonald framework, two models are constructed. The First model has host immunity (*SIR for hosts*

Chapter	Main areas / Questions	Models used	Main findings
Chapter 2	Investigating the deterministic dynamics of Ross-Macdonald based models, defining the epidemic threshold for host and vector	Ross-Macdonald model with immune class in host RM_{SIR} & incubation in both populations RM_{SEIR}.	Models are found locally asymptotically stable and solutions reach asymptotically to the trivial equilibrium and endemic equilibrium. R_0 is both invasion and persistence threshold.
Chapter 3	Estimating pathogen persistence in host-vector systems using stochastic models. How pathogen persistence relates to the parameters of the model?	RM_{SIR} & RM_{SEIR}	Determinants of CCS for host-vector systems are found in terms of the parameters of the model. The most important determinants are birth / death and latency rate of vectors and clearance rate of hosts as well as the number of recovered humans R^* and $\frac{R_0}{R_0-1}$.
Chapter 4	Modelling persistence using a stochastic version of Ross Macdonald dengue models with seasonality	RM_{SEIR}^s	The impact of introducing an infectious individual at different times of year and the peak of infectious individuals is clearly determined by the seasonal conditions.
Chapter 5	Investigating the approximations of the full host-vector model that have less computational costs.	Vector Proxy (VP) & Reservoir model	Both models approximated the host-vector dynamics well. Composite parameters were estimated analytically. Approximation breaks down when the behaviour of is far from equilibrium.

Table 1.1: A brief summary of the work done in different chapters.
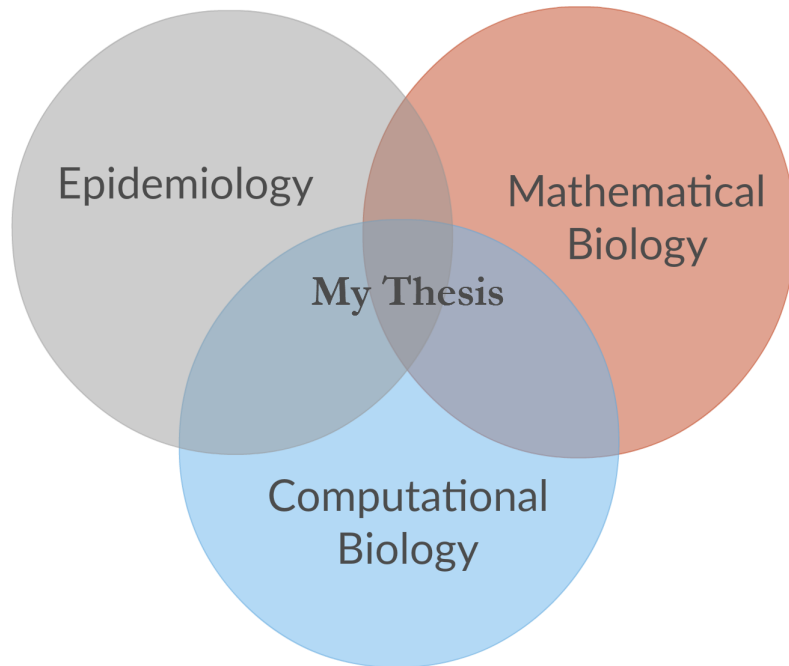
Figure 1.7: Venn diagram showing the main areas of my thesis.

and SI for vectors), and is termed RM_{SIR} and the second model, termed RM_{SEIR} has an exposed class in both host and vector populations. Values of the parameters are selected from literature corresponding to those of dengue with very low host-to-vector and vector-to-host transmission rates. The basic reproductive number R_0, along with R_0^{HV} for hosts and R_0^{VH} for vectors is derived for both models. Models are analysed for linear stability and the Disease Free Equilibrium (DFE) and Endemic Equilibrium (EE) for models with host and vector demography are found locally asymptotically stable. R_0 is shown to be the threshold value and the model shows bifurcation in behaviour at $R_0 = 1$, i.e, infection will invade if $R_0 > 1$ and vice versa. This chapter lays the theoretical mathematical foundation for the work developed later in the thesis and discusses the stability analysis for both Ordinary Differential Equations (ODEs) based models. Moreover, this work concentrates only on the deterministic behaviour of the models and can be seen as the first step in using the classical way of analysing the problem of pathogen persistence using ODEs.

Chapter 3: Determinants of long-term pathogen persistence in host-vector systems

This chapter gives an in-depth analysis of the stochastic version of the models developed

in the previous chapter. The aim is to look at the stochastic dynamics of both RM_{SIR} and RM_{SEIR} and hence investigate the determinants of the pathogen persistence in host and vector populations. The persistence is measured in relation to the CCS, which associates the pathogen persistence in stochastic models with the population size. Stochastic simulations using the parameter space from the previous chapter are carried out to find the probability of extinction $P(E)$ of the disease. Then $P(E)$ which is obtained by running simulations until a targeted time of 25 years is used to estimate the CCS within the host population. To find the stochastic fade-out of disease, variants of the Gillespie algorithm are used for carrying out the stochastic realizations. The CCS is found to be 1.3 million hosts for the baseline model RM_{SIR} and was reduced to less than half (0.6 million hosts) in RM_{SEIR}, i.e., the inclusion of latent periods have dramatic impact on the persistence of dengue virus in hosts. To further investigate the association between parameters of the models and CCS, sensitivity analysis was performed and general linear models were used to quantify the relationship between CCS and the model parameters. The work undertaken in this chapter provides a computational framework for relating persistence with the host population and estimating the formulation of the CCS using the core parameter values or their simple algebraic combinations. From an population ecology and mathematical biology perspective, the work done provides a novel approach of finding an algebraic formula for CCS in host-vector systems. From a computational biology perspective, this study involves developing fast and efficient variants of Gillespie algorithms to aid the estimation of CCS.

Chapter 4: Modelling persistence using Ross Macdonald dengue model with seasonality

Chapter 4 further develops the host-vector model, RM_{SEIR} developed in Chapter 2, for modelling transmission dynamics of dengue and investigating the persistence of the virus in the host and vector population. The model in this chapter has a seasonally dependent birth rate $\delta_b(t)$ for vectors, considered to be mediated by temperature, rainfall and humidity. The new model is termed as RM_{SEIR}^s, the superscript s denoting seasonality. The birth rate of the vectors varies throughout the year during the wet (favourable) and dry (unfavourable) seasons. Both the deterministic and stochastic versions of the model are presented in this chapter. The parameter space used in this work corresponds to that of dengue as measured in empirical studies. Twelve different seasonal points during a year were chosen as starting points for the deterministic and stochastic versions of the model. The derivation of different analytic measures including seasonal reproductive numbers $R_{0|t_0}$, $R_{t|t_0}$, and $R_{t|t_0,S_h}$ along with the probabilities of invasion $P_{Inv|I_v=1,t_0}$, and $P_{Inv|I_h=1,t_0}$ were presented. The date of arrival of an infectious individual effects the timing and the distribution of the infectious humans. If introduced in an unfavourable season, the outbreak takes longer to attain the

peak value of I_h because they have a longer time to evolve before the next unfavourable season (Otero and Solari, 2010). In other experiment, it was shown that the increase in the seroprevalence levels of hosts reduced the probability of invasion. The probability of persistence of dengue infection during one year was greater in unfavourable seasons, as the change in mosquito population alters the transmission mechanism. The distribution of time to extinction t_e was effected by the seasons. In favourable season, t_e was less after higher peaks of infectious individuals. The peak of infectious humans I_h and the time to attain the peak value was dominated by the seasonal fluctuation in the vector population. From a mathematical biology perspective, the work undertaken in this chapter includes the derivation of the analytical forms of different basic reproductive numbers and probabilities of invasion. From an epidemiological perspective, the time evolution of the basic reproductive ratios in the seasonal model was investigated and the relationship between the probability of invasion and seroprevalence levels in humans was explored. At higher seroprevalence levels, decreasing $R_{t|t_0}^{VH}$ can be helpful in bringing $R_{t|t_0} < 1$, which can be achieved by decreasing the rate of transmission from an infectious mosquito.

Chapter 5: Direct transmission models to represent host-vector systems

One of the main focuses of the studies conducted in this thesis is the development of mathematical models using a 'minimalist' approach. In this chapter, an argument about explicitly incorporating the vector population in modelling dengue transmission is developed by constructing two alternatives to the host-vector modelling structure. The baseline host-vector model RM_{SIR} is modified as (i) An SIR model with a latent class L that acts as a proxy for the effect of vectors in the host population that allows for a delay in transmission. (ii) An SIR model that contains a 'Pool' or reservoir of infection P which infects the host population. In this case the reservoir represents the population of the infectious vectors. Parameter space where the models performed well were identified. Models were thoroughly tested in biological scenarios that leads to the breakdown of the approximation. The approximated models are validated by comparing the stochastic trajectories, CCS measures and the Quasi-stationary distributions (QSD) to the corresponding compartments in RM_{SIR}. From a mathematical biology perspective, this work includes the construction, validation and identification of the areas of parameter space where the models serve as acceptable approximations to RM_{SIR}. It also includes analytical derivation of unknown rates in both models with their biological explanation. From a computational biology perspective, the main novel contribution of this chapter is the presentation of robust and alternative schemes for estimating pathogen persistence in host-vector systems. These findings have the potential for significant contributions to the real-world applications; the models are easier to use analytically and can be used to answer a wide range of questions compared to the full host-vector model.

Chapter 6: Conclusion and Future directions

This Chapter concludes the thesis. This chapter starts with an overview of the findings from the main studies of the thesis. Then a section is devoted to highlight important conclusions and discusses the potential future research arising from this project.

CHAPTER 2

Persistence thresholds and stability in deterministic models for host-vector systems

Persistence thresholds and stability in deterministic models for host-vector systems

2.1 Introduction

Vector-Borne Diseases (VBDs) are diseases spread in the host population by insects or vectors. They include malaria, dengue fever, yellow fever, Chagas disease, Rift Valley fever, and Chikungunya. VBDs are transmitted to humans, plants and animals by vectors. It is worth studying the transmission mechanism of VBDs because of their rapid spread and persistence across the globe (WHO, 2014). Over the last five decades, many mosquito-borne human illnesses have emerged in different parts of the world whereas malaria and dengue have re-emerged in Asia and Americas (Gubler, 1998). As a result, their impact over the economy, ecology and public health are increasing with time. In many parts of the world diseases like malaria and dengue show endemic behaviour, causing recurrent outbreaks (Simmons et al., 2012; World Health Organization, 2014). Understanding the biology and ecology of pathogens, hosts, vectors, and their environment is crucial for the development of novel and effective intervention and mitigation measures (Institute of Medicine, 2008).

Modelling provides a cost-effective approach to address problems of invasion and persistence in epidemiology. The foundations of using compartmental models in epidemiology were laid more than a century ago. In early 1900 Ross formulated the seminal model for malaria transmission between humans and mosquitoes (Ross, 1911) using a set of Ordinary Differential Equations (ODEs) to represent the rate of change of individuals in host and vector populations. This model was revisited by Lotka (1912). In 1927 and 1932, basic compartment mathematical models were proposed by Kermack and McKendrick (1927) and Kermack and McKendrick (1932) which divide the total population of individuals into healthy (Susceptible; S), sick (Infected; I) and immune (Recovered; R) individuals. The rate of change of the number of individuals in each class is represented in the form of ODEs having defined transmission and recovery rates. This form of modelling is still widely used today and provides

valuable information about the mechanisms of disease transmission.

Recently, Reiner et al. (2013) presented a review of mathematical models used for mosquito-borne pathogen transmission from 1970 to 2010. They found that most of the disease transmission models closely resemble to the Ross Macdonald model. A review of deterministic and stochastic compartmental models was provided by Nishiura (2006), which was centred on dengue control. The author estimated R_0 using different techniques and discussed the impact of different factors upon dengue transmission. He further suggested that the interaction between field professionals and theoretical modellers provides meaningful insights from dengue data. Koella (1991) used simple mathematical models for the understanding of malaria transmission to help to plan control strategies. Johansson et al. (2011) conducted a detailed study of the models having compartmental structure. These authors reviewed the mathematical approaches from 1972 to 2010 and used them for the estimation of the basic reproductive number R_0 for assessing the critical vaccination fraction of the population. Non-spatial and deterministic approaches for modelling dengue are reviewed by Andraud et al. (2012). They suggested a multi-serotype host-vector model with the combination of vector-control and vaccination strategies for areas where the pathogen is persistent.

There are different approaches taken to model VBD transmission using the Ross-Macdonald compartmental framework. For dengue disease, these approaches include the work of Esteva and Vargas (1999) that assessed the effects of human demography. They identified three threshold parameters: R_0, R_1 R. The first threshold parameter, R_0 was conditioned on the existence and stability of an endemic equilibrium, R_1 is related to the behaviour of the number of infectious humans and R controls the growth of the host population density. The same authors Esteva and Vargas (2000) studied the impact of vertical and mechanical transmission (*after an interrupted meal by a mosquito on an infectious person*) routes. Vertical transmission was found to favour the persistence by dramatically increasing the endemic proportion of infectious vectors, especially in endemic areas with low host population. However, the proportion of vertically infected mosquitoes in the vector population is critical for the efficacy of this transmission route. In contrast to the above study, Adams and Boots (2010) questioned the efficacy of vertically acquired infection in vectors. They argued that the rates of vertically acquired infection reported in most of the empirical studies are very low to have a profound impact on the persistence of dengue virus in the vector population. The long term persistence of the dengue virus is not possible with these rates as the virus is rapidly lost with every generation of mosquitoes.

In this chapter, a simple non-seasonal deterministic model RM_{SIR} is constructed on the basis of Ross-Macdonald's modelling framework. Here the subscript SIR refers to the compart-

ments of the host population. The model has immunity incorporated into the host population whereas the vector population has only two classes (i) susceptible, S and (ii) infectious I. This model is later extended by adding incubation compartments E in the host and vector populations. The second model is termed as RM_{SEIR}. These models are not entirely new, rather they are a variant of models studied by different authors (Andraud et al., 2012; Nishiura, 2006). The intention here is to use simple model structures to generate persistence patterns for endemic disease and identify the primary determinants of pathogen persistence via deterministic modelling. Both of these models are epidemiologically different, so comparing the dynamics of a disease in these models is an interesting problem in its own right. In the rest of the chapter, both models are parametrized using same or similar quantities. This is done to help comparing the output generated by these models. In most of the work reported above, comparing the results from epidemiologically different models for a single disease is not done and this chapter is dedicated for the comparison of deterministic results obtained from RM_{SIR} and RM_{SEIR}.

The rest of the chapter is arranged as follows: the parameters and assumptions of both models are explained in detail in subsequent sections. Next, both models are investigated sequentially for local asymptotic stability. The invasion threshold R_0 is discussed and population-level reproductive numbers are identified by separating R_0 into R_0^{VH} and R_0^{HV}. This chapter surveys the behaviour of both models in deterministic settings and compare the variation in results in relation to modification in modelling structure. The main contribution of this chapter is being a mathematical preamble for the rest of the thesis. Therefore, this chapter serves as a mathematical foundation for next three chapters where these models are extended to include stochastic effects.

2.2   Ross Macdonald model with immunity (RM_{SIR})

The model and description used in this section i.e., RM_{SIR} is closely related to the model proposed by Lloyd et al. (2007) and the generic 'single-serotype' dengue model presented in the review articles of Nishiura (2006) and Andraud et al. (2012). Over the course of this section, the deterministic models for host and vector populations are represented. After that the 'ingredients' of the model are explained in detail. The basic features and assumptions of the model are then presented. At the end of this section, the model is examined for local stability analysis at the disease-free equilibrium and at endemic equilibrium.

Following conventions from the start of previous section 2.1, the Ross Macdonald compartmental framework with immunity is made up of a Susceptible-Infected-Recovered (S I R)

system of equations for host population dynamics and a Susceptible-Infected ($S\,I$) system for the mosquito population. There is no recovered / immune class for mosquitoes as once they are infected, they remain infected for the rest of their life. This model was originally developed for malaria but can be adapted to represent many host-vector infection dynamics. It consists of the host population, the size of which is denoted by H. Susceptible hosts are denoted by S_h and I_h denotes infected hosts. The recovered class is denoted by R_h. The closure assumption of the host population leads to $H = S_h + I_h + R_h$. The rate of change of the number of susceptible, infective and recovered hosts is expressed using the following set of equations

$$
\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= \Lambda H - kpI_v\left(\frac{S_h}{H}\right) - \gamma S_h \\
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= kpI_v\left(\frac{S_h}{H}\right) - (\xi + \gamma)I_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \xi I_h - \gamma R_h.
\end{aligned}
\tag{2.1}
$$

Vectors in the model are either susceptible S_v or infected I_v. Since the vector population V is closed, $V = S_v + I_v$. Only female mosquitoes require blood meals during their gonothropic cycle for egg production so the model excludes male mosquitoes and V accounts for roughly half of the total vector population. Table 2.1 provides an overview of the parameters used in the model.

$$
\begin{aligned}
\frac{\mathrm{d}S_v}{\mathrm{d}t} &= \delta V - kqS_v\left(\frac{I_h}{H}\right) - \delta S_v \\
\frac{\mathrm{d}I_v}{\mathrm{d}t} &= kqS_v\left(\frac{I_h}{H}\right) - \delta I_v.
\end{aligned}
\tag{2.2}
$$

Symbol	Explanation	Value used	Reference
k	Bite per mosquito per day, in days^{-1}	0.5	Lloyd et al. (2007)
p	Transmission probability from an I_v to a S_h	0.2	Lloyd et al. (2007)
q	Transmission probability from an I_h to a S_v	0.15	Lloyd et al. (2007)
α	Vector-to-host transmission rate, in days^{-1}	kp	-
β	Host-to-vector transmission rate, in days^{-1}	kq	-
ξ	Average recovery rate of hosts, in days^{-1}	0.1428	Adams and Boots (2010)
$\Lambda = \gamma$	Birth / death rate of hosts, in days^{-1}	4.215×10^{-5}	estimated
δ	Birth / death rate of vectors, in days^{-1}	0.125	Adams and Boots (2010)

Table 2.1: Parameters used in RM_{SIR}. Here the average birth rate Λ of hosts is equal to average death rate γ. The average host life expectancy is set to $\frac{1}{65}$ years, which is used in days^{-1} in the model.

Equation sets 2.1 and 2.2 present the host-vector system for RM_{SIR}. Here $\frac{S_h}{H}$ and $\frac{I_h}{H}$ represents the susceptible and infected proportion of the host population. Assumption of

non-varying host and mosquito population yields the rate of change of the population sizes $\frac{dH}{dt} = \frac{dV}{dt} = 0$. The schematic diagram of the model is presented in Figure 2.1 whereas the parameters included in above equations are explained in Table 2.1. All parameters are positive constants i.e., α, β, ξ, δ, Λ, and γ and $\in \mathbb{R}_+$. Here $\alpha = k \times p$, where k is the bite rate per mosquito per day and p is the probability of transmission from an infectious vector to a susceptible host. Similarly, $\beta = k \times q$, where q is the probability of transmission from an infectious host to a susceptible vector.
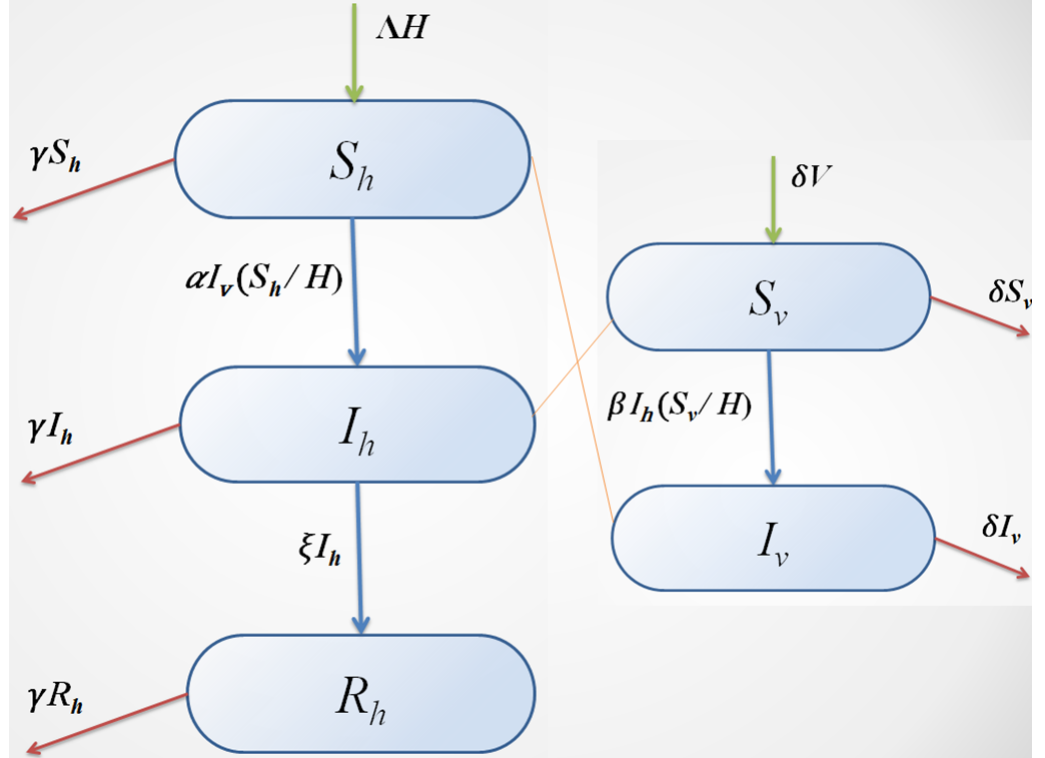
Figure 2.1: Schematic Diagram of the Ross Macdonald model with Host Immunity. Here green arrow represents the birth and red arrows represents death in all compartments. Blue arrows show the flow of individuals from one compartment to other. Light orange lines between population show transmission.

The contact rate α can be interpreted as the expected number of bites a vector makes per unit time which leads to $\left(\alpha I_v \frac{S_h}{H}\right)$. This expression shows the population rate of bites that create new infections, conditioned upon an infectious vector biting a susceptible individual of the host population. It is assumed that the average infective individual will make sufficient contacts (i.e., there are sufficient hosts to bite) to transmit the disease and the transmission rate is independent of the population size (i.e, frequency dependent transmission, which is commonly used for host-vector systems). A similar analogy is applied for defining β by swapping the roles of susceptible proportion of hosts to infected proportion $\left(\frac{I_h}{H}\right)$ and infectious vector I_v to susceptible S_v. The inverse of the average birth and death rate of

vector, $\left(\frac{1}{\delta}\right)$ also gives the average infectious period of the vector, as it is assumed that once infected, a vector remains infected till death. The average host clearance rate ξ is taken as $\frac{1}{7}$ days^{-1}. This is due to the fact that most of the host remains infected roughly around a week and a very small proportion of individuals, who develop severe complications of the disease, remain ill beyond this time period. To study the long-term dynamics of an endemic disease requires that host demography be included in the model, as this is a pre-requisite for any possibility of disease persistence. The birth and death rates in both populations scale with the population size and are constant to keep H and V the same over the passage of time. It was assumed that there is no external re-introduction of disease in any population. The time scale for demographic turnover in hosts is very long as compared to the infectious period of the disease, but is equal in the case of vectors as they do not recover.

The above model can be thought as a simplified non-seasonal form of a dengue model, since two parameters δ and ξ are taken from Adams and Boots (2010). It is anticipated that the parametrization of the model represents an endemic disease. The description of the basic reproductive number of RM_{SIR} and assumptions of the model are presented in later sub-sections.

2.2.1 The basic reproductive numbers

The basic reproductive number R_0 is regarded as one the of most important concepts developed in a mathematical treatment of disease spread. In epidemiology, R_0 is the number of individuals infected by a single infected individual during his or her entire infectious period, in an entirely susceptible population (Heffernan et al., 2005).

The basic reproductive number R_0 for RM_{SIR} is obtained from the introduction of one infective host in an entirely susceptible vector population or from the introduction of one viremic vector. There are different techniques to derive R_0 and this work employs the next-generation matrix method, where the spectral radius of the next generation matrix gives the basic reproductive number (Diekmann and Heesterbeek, 2000). This method is particularly useful for estimating R_0 in multiple populations as different categories of individuals take part in disease transmission. The basic reproductive number R_0 of the model under consideration is derived in appendix A section A.1.1.

$$R_0 = \sqrt{R_0^{VH} \times R_0^{HV}} = \frac{\alpha}{\delta} \times \frac{\beta V}{(\xi + \gamma)H} = \sqrt{\frac{\alpha \beta V}{(\xi + \gamma)\delta H}}. \tag{2.3}$$

In the above expression, R_0^{VH} denotes the average number of hosts directly infected by the introduction of a single infective vector into an entirely susceptible host population. Since the transmission rate from an infected vector to a susceptible host is α and the average life time of a mosquito is $\frac{1}{\delta}$, R_0^{VH} is the product of these parameters. Similarly, R_0^{HV} denotes the average number of vectors that become directly infected upon the introduction of a single infectious host into an entirely susceptible vector population. Following the same argument as above, R_0^{HV} is $\beta \frac{V}{H} \times \frac{1}{(\xi+\gamma)}$.

The number $R_0 = 1$ represents the threshold condition for both invasion and persistence in the deterministic model. It is important to note that the product $\alpha \times \beta$ present in the expression 2.3 results the square of the bite rate; a specific feature of the Ross Macdonald model which indicates a two-step life-cycle of infection. In general, transmission of the virus is favoured by high densities of mosquitoes that bite frequently and hindered by the death and quick recovery of hosts, along with the high mortality rate of vectors.
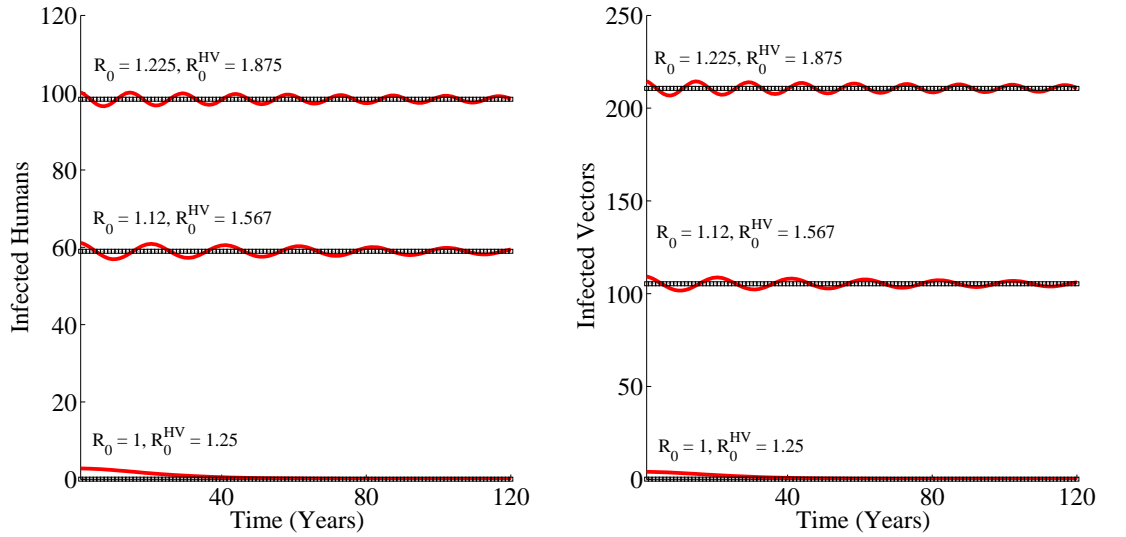
Figure 2.2: The change in equilibrium values of infected hosts and vectors for different values of R_0. Horizontal lines are equilibrium points of infected compartments at the corresponding values of R_0 from 1 (bottom) to 1.225 (top). The Infectious populations drops to zero at $R_0 = 1$ and persists at other values. Host population size (H) is 1×10^5 individuals and the ratio $\frac{V}{H}$ is altered to obtain the desired value of R_0. R_0^{VH} has a constant value of 0.8 in all the plots as it only depends upon the fixed parameters α and δ, whereas R_0^{VH} is varied according to the ratio $\frac{V}{H}$. The formula for basic reproductive number is $R_0 = \sqrt{R_0^{VH} \times R_0^{HV}}$, where $R_0^{VH} = \frac{\alpha}{\delta}$ and $R_0^{HV} = \frac{\beta V}{(\xi+\gamma)H}$. Constant parameters are $k = 0.5$, $p = 0.2$, $q = 0.15$, $\xi = \frac{1}{7}$, $\delta = \frac{1}{8}$ and $\Lambda = \gamma = 4.5 \times 10^{-5}$.

It is straightforward to infer that if $R_0 < 1$, every infected individual, on average, produces

less than one infected individual and disease will ultimately die out in both populations, so that the system will reach the disease free state. If $R_0 > 1$, then, on average, more than one secondary infection will be produced and the pathogen will invade the population. In other words, R_0 reflects the stability of the Disease Free Equilibrium (DFE) as when this number is less than one, we can predict that the pathogen will be cleared from the population.

R_0 is the product of R_0^{HV} and R_0^{VH} which are the reproductive numbers for vector and host populations respectively. R_0^{HV} is a positive linear function of vector-to-human population ratio $\frac{V}{H}$, so it effects the persistence of dengue by changing the basic reproductive number, though R_0^{VH} remains fixed. Either R_0^{HV} or R_0^{VH} can take a value less than one, the condition being that their product should be greater than 1. This is the threshold condition of the existence of the endemic state. As pointed out by Lloyd et al. (2007), the term $\frac{V}{H}$ in R_0^{HV} creates an asymmetry in invasion thresholds for hosts, if V is much larger than H. The relation between these reproductive numbers and disease persistence in shown in Figure 2.2. This plot is drawn by adjusting $\frac{V}{H}$ ratios in such a way that give $R_0 = 1, 1.12$ and 1.225. The left hand side shows the time evolution of infectious humans I_h and right hand side presents the time evolution of I_v at different values of R_0. The model RM_{SIR} is numerically integrated by adding some individuals to the endemic equilibrium values of the infectious compartments of both populations. The same number of individuals is subtracted from the number of susceptible individuals at endemic equilibrium to keep the population sizes, H and V constant. Black horizontal lines are the equilibrium points of the infected compartments. At $R_0 = 1$ (bottom), the Infectious populations drops to zero and persists at $R_0 = 1.12$ (middle) and 1.225 (top). The host population size (H) is 1×10^5 individuals and R_0^{VH} has a fixed value 0.8. It is important to note that unless specified, the vector-to-host population ratio $\frac{V}{H}$ is kept fixed at six vectors per host for the analysis performed on RM_{SIR} in the rest of the Chapter.

The deterministic solution showing the time evolution of host and vector populations is given in Figure 2.3. The numerical integration is performed while initializing the solver by taking 14 individuals from S_v^*, where $*$ denotes the number of individuals at the endemic equilibrium. These individuals are added to I_v^* to keep the size of the vector population constant. The same is done for the host population, but for five humans, so that $I_h^*(0) = I_h^* + 5$ and $S_h^*(0) = S_h^* - 5$. The oscillations in the number of individuals decay to reach the stable endemic equilibrium. These patterns are explained in terms of R_0 by (Esteva and Vargas, 1998). If $R_0 > 1$ and initial fraction of susceptible individuals $S = \frac{S_h}{H} + \frac{S_v}{V}$ satisfies $R_0 S > 1$, then the susceptible proportion will decrease and the infected proportion $I = \frac{I_h}{H} + \frac{I_v}{V}$ increases to a peak and then decreases. Transmission halts when there are no sufficient susceptible present in the populations and starts again when $R_0 S > 1$, due to the birth of new susceptible individuals. This process is repeated until there are secondary smaller epidemics and thus
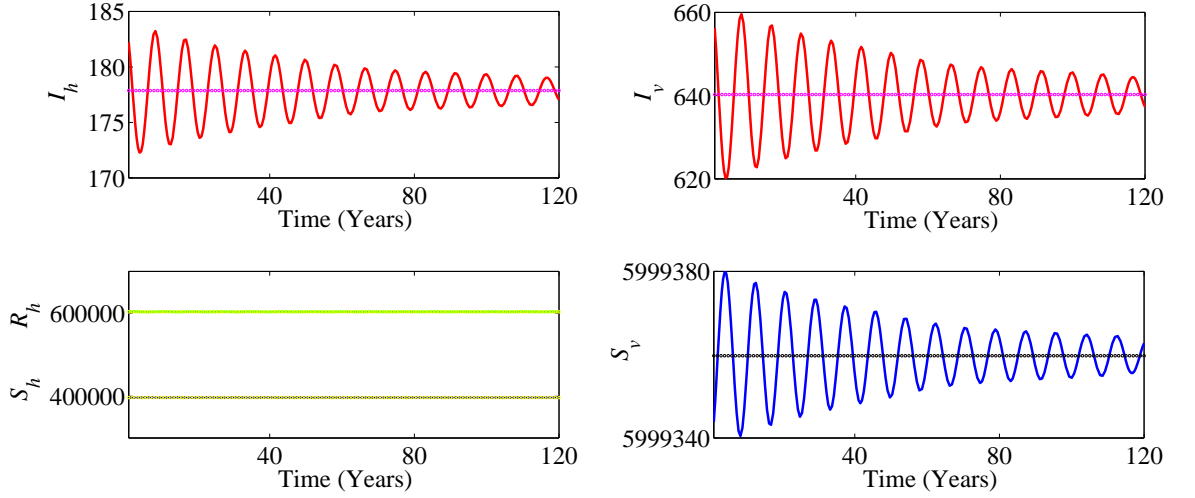
Figure 2.3: The convergence of host and vector compartments to the equilibrium points. Compartments in both populations reached the stable endemic equilibrium via damped oscillations. From top row: Temporal evolution of (i) Left: Human infectious compartment (I_h) and (ii) Right: vector Infectious compartment (I_v) . In bottom row, the susceptible (S_h) and recovered (R_h) compartments of hosts are at left side and Susceptible vectors (S_v)are on the right panel. Horizontal lines represent the deterministic equilibrium points. Initial conditions of the solver at detailed in the section 2.2.1. Host population size (H) is 1×10^6 individuals and vector population size is six times the host population. $R_0 = \sqrt{c \times \frac{V}{H}} = 1.56$ where $c = \frac{\alpha\beta}{(\xi+\gamma)\delta}$ and $\frac{V}{H} = 6$. Constant parameters are same as in Figure 2.2.

the solution relaxes to the endemic equilibrium. If $R_0 < 1$, the system approaches the DFE asymptotically. As the susceptible proportion is increased by constant birth rate, in the long run all recovered individual will die out leaving the population consisting of only susceptible individuals.

2.2.2 Assumptions of the model

The assumptions of this model are as follows:

- The human population size, H is constant as the birth rate is equal to the death rate. Humans are divided into susceptible S_h, infectious I_h, and recovered / immune R_h classes. Humans gain life-long immunity after recovering form the dengue infection.

- The population size of mosquitoes, V, is constant and mosquitoes are divided into susceptible $S_v = V - I_v$, and infectious, I_v classes. Birth and death rate are same

and new-born mosquitoes are considered susceptible. Once infected, vectors remain infectious for life.

- There is no seasonal fluctuation in the recruitment and death of vector population.

- The rate at which mosquitoes bite humans is proportional to the number of mosquitoes but independent of the number of people. Transmission of infection in humans involve two biting events from the vector.

- Humans are assumed to be the only competent hosts and no bites are assumed on bovine, cattle and wild and domestic animals.

- The infectious period of hosts and vectors is exponentially distributed with expectation $\frac{1}{\xi}$ and $\frac{1}{\delta}$, respectively.

- Both hosts and vectors are assumed to become infectious immediately following infection, without an incubation period.

- Populations of host and vector are large and well-mixed or homogeneous (with frequency-dependent transmission).

2.2.3 Equilibrium points and local stability analysis

By taking into account the population closure assumption for hosts and vectors, recovered human and susceptible vector classes are written as: $R^* = H - S_h^* - I_h^*$ and $S_v^* = V - I_v^*$, providing three equations from the equation set 2.1 and 2.2 for further analysis. This system of three equations admits two distinct equilibrium points: (i) A Disease Free Equilibrium point (DFE) denoted by $E^0(H, 0, 0)$, and (ii) an Endemic Equilibrium point (EE), denoted by $E^*(S_h^*, I_h^*, I_v^*)$.

The DFE can be interpreted as if there are no infected individuals in both host and vector populations. At the DFE, both populations will consist of only susceptible individuals as there will be no recovered human as a result of the demographic of turn-over of the host population. In models with lifelong immunity, the EE can be thought as the state where disease is present in some groups of the population. The endemic equilibrium of a system is sustained by continuous recruitment of susceptible individuals that become infectious. In deterministic dynamics, the EE is referred as the endemic stationary state of the system. The endemic equilibrium points of equations 2.1 and 2.2 are derived using basic algebra. The calculations are verified by using MATLAB $^\circledR$ Symbolic Math Toolbox (MATLAB, 2014).

$$S_h^* = \frac{H(\gamma H R_0 + \alpha V)}{R_0(\alpha V + \gamma H)}$$

$$I_h^* = \frac{\delta \gamma H^2 (R_0 - 1)}{\beta(\alpha V + \gamma H)} \tag{2.4}$$

$$R_h^* = \frac{\xi \delta H^2 (R_0 - 1)}{\beta(\alpha V + \gamma H)}$$

$$S_v^* = \frac{V(\alpha V + \gamma H)}{(\alpha V + \gamma H R_0)}$$

$$I_v^* = \frac{\delta \gamma H (\xi + \gamma)(R_0 - 1)}{\alpha\Big(\delta(\xi + \gamma) + \beta\gamma\Big)} \tag{2.5}$$

By substituting values from Table 2.1 and taking the values of H and V same as described in Figure 2.3, $R_0 = 1.6$, $S_h^* = 396985$, $I_h^* = 178$, $R_h^* = 602836$, $S_v^* = 5999360$, and $I_v^* = 640$. At equilibrium, seroprevalence level in hosts is 60% and 0.02% of the human population is infectious. In vectors. 0.011% of the population is infectious at endemic equilibrium. These endemic equilibrium points are biologically feasible if $R_0 \geq 1$ as the number of individuals in each compartment are bound to be positive. If $R_0 < 1$, then compartments have negative number of individuals.

In the following subsections and in section 2.3.1, the following result is used to establish the stability of the models: If all eigenvalues are real, negative, and distinct, the system is stable (Boyce and DiPrima, 1986). Where possible, the coefficients of the characteristic equations are shown in terms of R_0, and eigenvalues are obtained by solving the characteristic equation.

2.2.3.1 Local Stability of disease free equilibrium

As discussed above, the dynamical system RM_{SIR} admitted two distinct equilibrium points. In many studies related to compartmental systems, the basic reproductive number was found as a threshold which exhibited a bifurcating behaviour between both disease equilibrium states [for example, see Rodriguesa et al. (2012); Yang et al. (2010) and Esteva and Vargas (1998) for vector-borne diseases]. In the following two subsections, the local asymptotic stability analysis of both equilibria has been presented where the model behaviour in the disease-free and endemic state was investigated with respect to R_0. The goal was to investigate whether the disease free equilibrium point E^0 was locally asymptotically stable when $R_0 < 1$. Moreover at $R_0 > 1$, E^0 became unstable and the endemic equilibrium point E^* was locally asymptotically stable. For host and vector systems in equations 2.1 and 2.2, the next paragraph addresses the local stability of the DFE, i.e, $E^0 (H, 0, 0)$.

The Jacobian matrix $J\left(S_h, I_h, I_v\right)$ of above mentioned system is:

$$J\left(S_h, I_h, I_v\right) = \begin{pmatrix} -(\alpha\frac{I_v}{H} + \gamma) & 0 & -\alpha\frac{S_h}{H} \\ \alpha\frac{I_h}{H} & -(\xi + \gamma) & \alpha\frac{S_h}{H} \\ 0 & \beta\frac{S_v}{H} & -\delta \end{pmatrix} \tag{2.6}$$

As explained in section 2.2.3, the number of infective hosts and vectors is zero at the DFE. Evaluating the above Jacobian matrix at DFE, by setting $I_h = I_v = 0$, $S_h = H$ and $S_v = V$ in above matrix yields

$$J\left(H, 0, 0\right) = \begin{pmatrix} -\gamma & 0 & -\alpha \\ 0 & -(\xi + \gamma) & \alpha \\ 0 & \beta\frac{S_v}{H} & -\delta \end{pmatrix} \tag{2.7}$$

The local stability is governed by the eigenvalues of the matrix $J\left(H, 0, 0\right)$. By using the relation $|\lambda I_n - J\left(H, 0, 0\right)| = 0$, where I_n is the identity matrix having same order as $J\left(H, 0, 0\right)$ the following matrix is obtained.

$$\left|\lambda I_3 - J\left(H, 0, 0\right)\right| = \begin{pmatrix} \lambda + \gamma & 0 & \alpha \\ 0 & \lambda + (\xi + \gamma) & -\alpha \\ 0 & -\beta\frac{V}{H} & \lambda + \delta \end{pmatrix}$$

Expanding the determinant and equation the expression to zero will give the characteristic equation whose roots are the eigenvalues of $J\left(H, 0, 0\right)$

$$(\lambda + \gamma)\left((\lambda + \delta)(\lambda + \xi + \gamma) - \alpha\beta(\tfrac{V}{H})\right) = 0. \tag{2.8}$$

One root is $\lambda_1 = -\gamma$ and for the other two roots, the quadratic part of equation 2.8 can be simplified as:

$$\lambda^2 + K\lambda + L = 0 \tag{2.9}$$

where $\quad K = \xi + \delta + \gamma,$ and $\quad L = \delta(\xi + \gamma)(1 - R_0).$

The eigenvalues from equation 2.9 have negative real parts if and only if all the coefficients are positive (Routh-Hurwitz criterion for second-order polynomial). Here K is positive, and L is positive only when $R_0 < 1$. The negative real parts of the eigenvalues guarantees the local asymptomatic stability. Therefore, $E^0\left(H, 0, 0\right)$ is locally asymptotically stable. The other two roots are complex conjugates:

$$\lambda_{2,3} = -\frac{\xi + \delta + \gamma}{2} \pm \frac{\sqrt{(\xi + \delta + \gamma)^2 - 4\delta(\xi + \gamma)(1 - R_0)}}{2} \tag{2.10}$$

It can be concluded immediately that at $R_0 > 1$, the disease-free equilibrium $E^0 (H, 0, 0)$ becomes unstable as the sign of L becomes negative. In deterministic settings this means that the pathogen will persist after invasion. The plot of λ_2 and λ_3 in Figure 2.4 shows the stability of the eigenvalues. In this plot, R_0 is changed by altering the $\frac{V}{H}$ ratio, by increasing H and keeping V fixed. It is interesting to note that as the population size H goes up, the value of R_0 decreases. As R_0 approaches 1, λ_2 approaches to zero and λ_3 becomes more and more negative. For lower values of R_0, both eigenvalues converges to $\lambda_2 = -0.1$ and $\lambda_3 = -0.2$ at bigger population sizes.
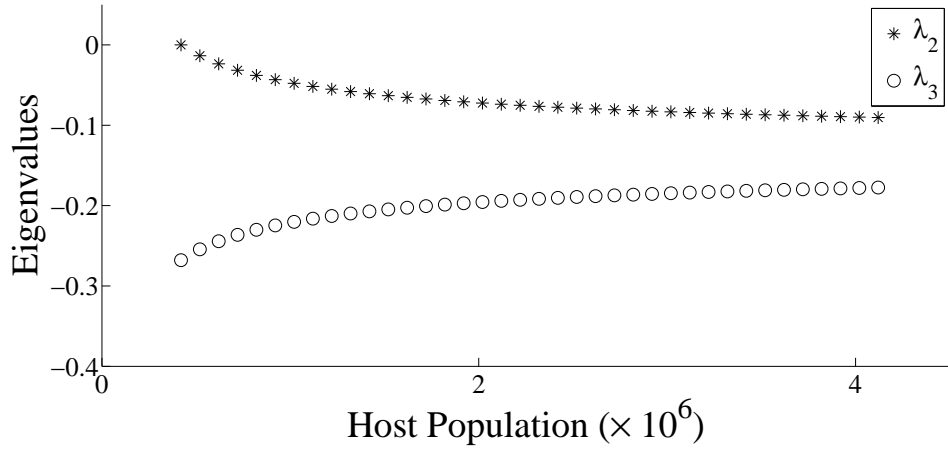
Figure 2.4: Change in the eigenvalues λ_2 and λ_3 as the host population (H) is altered. The vector population (V) is kept fixed. Since $R_0 = c \times \frac{V}{H}$, the value of the basic reproductive number R_0 drops from 1 to 0.1 with the gradual increase in H. It is important to note that when R_0 starts to increase past one, the eigenvalue λ_2 becomes positive. This results in the negative value of the constant term L which appears in equation 2.9.

2.2.3.2 Local Stability of endemic equilibrium

For the endemic equilibrium $E^* (S_h^*, I_h^*, I_v^*)$, the Jacobian matrix is:

$$J\left(S_h^*, I_h^*, I_v^*\right) = \begin{pmatrix} -(\alpha \frac{I_v^*}{H} + \gamma) & 0 & -\alpha \frac{S_h^*}{H} \\ \alpha \frac{I_v^*}{H} & -(\xi + \gamma) & \alpha \frac{S_h^*}{H} \\ 0 & \beta \frac{S_v^*}{H} & -\delta \end{pmatrix} \tag{2.11}$$

As previously, the characterstic equation is used to find the eigenvalues; this is obtained by using the relation $|\lambda I_3 - J(S_h^*, I_h^*, I_v^*)| = 0$. The resulting matrix is:

$$\left|\lambda I_3 - J\left(S_h^*, I_h^*, I_v^*\right)\right| = \begin{pmatrix} \lambda + (\alpha\frac{I_h^*}{H} + \gamma) & 0 & \alpha\frac{S_h^*}{H} \\ -\alpha\frac{I_v^*}{H} & \lambda + (\xi + \gamma) & -\alpha\frac{S_h^*}{H} \\ 0 & -\beta\frac{S_v^*}{H} & \lambda + \delta \end{pmatrix}$$

By solving $|\lambda I_3 - J\left(S_h^*, I_h^*, I_v^*\right)| = 0$, following equation is formed.

$$
\begin{aligned}
\lambda^3 &+ \lambda^2\left(\xi + 2\gamma + (\alpha\frac{I_v^*}{H})\right) \\
&+ \lambda\left(\delta(\xi+\gamma) + (\xi+\delta+\gamma)(\gamma + \alpha\frac{I_v^*}{H}) - \alpha\beta\frac{S_h^*}{H}\frac{S_v^*}{H}\right) \\
&+ \left((\delta(\xi+\gamma)(\gamma + \alpha\frac{I_v^*}{H}) - \alpha\beta\gamma\frac{S_h^*}{H}\frac{S_v^*}{H}\right) = 0
\end{aligned}
\tag{2.12}
$$

As the rate of change is zero at the equilibrium points, the following substitutions from equation set 2.1 and 2.2 are used in equation 2.12

$$\alpha\left(\frac{S_h^*}{H}\right) = (\xi+\gamma)\frac{I_h^*}{I_v^*}$$

$$\beta\left(\frac{S_v^*}{H}\right) = \delta\frac{I_v^*}{I_h^*}$$

The final equation is of the form:

$$\lambda^3 + K\lambda^2 + L\lambda + M = 0 \tag{2.13}$$

where $\quad K = \xi + \delta + 2\gamma + \alpha\left(\frac{I_v^*}{H}\right), \quad L = (\xi+\delta+\gamma)\left(\gamma + \alpha\left(\frac{I_v^*}{H}\right)\right)$

and $\quad M = \alpha\delta\,(\xi+\gamma)\left(\frac{I_v^*}{H}\right)$

The eigenvalues of equation 2.13 have negative real parts if and only if: (i) all the coefficients (K, L and M) are positive, and (ii) $KL > M$ (Routh-Hurwitz criterion for third-order polynomials). The first condition is quite straightforward as the coefficients consist of either positive parameter values or values from the host and vector population which cannot be negative by definition. For the second condition, after putting the values of parameters from Table 2.1 satisfies $KL > M$, i.e, $KL - M = 6.4796 \times 10^{-6}$. Therefore the endemic equilibrium state $E^*\left(S_h^*, I_h^*, I_v^*\right)$ is locally asymptotically stable.

The coefficients K, L, M can also be represented in terms of R_0 by replacing the value of I_v^* from equation 2.5 into equation 2.13. After some algebraic manipulation and assuming

($\Lambda = \gamma$), K, L and M can be written as a function of R_0. In general, the product of KL leads to more terms with squared values as compared to M.

$$K = (\xi + \delta + \gamma) + \left(\frac{\delta\gamma(\xi + \gamma)(R_0 - 1)}{\delta(\xi + \gamma) + \beta\gamma} \right)$$
$$L = (\xi + \delta + \gamma) \left(\gamma + \gamma\frac{\delta(\xi + \gamma)(R_0 - 1)}{\delta(\xi + \gamma) + \beta\gamma} \right)$$
$$M = \delta\gamma (\xi + \gamma) \left(\frac{\delta(\xi + \gamma)(R_0 - 1)}{\delta(\xi + \gamma) + \beta\gamma} \right).$$

2.3   Ross Macdonald model with incubation (RM_{SEIR})

In many vector borne diseases, there is a certain amount of time required for the virus to replicate itself inside the vector or the host. Therefore, when the virus levels reach a certain threshold in a vector, the virus reaches the salivary glands and is transmitted via infected bite. The incubation periods for DENV is estimated by Chan and Johansson (2012). Similarly after taking an infectious bite, it takes time for the virus to replicate inside the host before making its way into the blood stream in order to make a vector sick. In host-vector models, the incubation period of pathogen inside a host or a vector plays a vital role in different aspects of disease propagation and spread (Keeling and Grenfell, 1998). The model discussed previously assumed that when virus is injected into the blood stream of a susceptible, that individual is able to transmit the disease immediately.

In dengue disease, the time required for virus to replicate in the host, before reaching the blood stream to be release while getting an infectious bite, is termed the Intrinsic Incubation Period (IIP). Similarly, the time required for virus to replicate inside a vector, to be able to reach in the salivary glands for transmission during the infectious bite, is termed the Extrinsic Incubation Period (EIP). How the addition of these two compartments in host vector system affect the dynamics of dengue, as compared to the previous model RM_{SIR} will be an interesting area to explore. The description of the model is presented below.

The dynamics of the Ross Macdonald model developed in section 2.2 with the addition of exposed / latent parameters are incorporated by using two parameters (σ) and (ρ). The Ross Macdonald model with incubation (RM_{SEIR}) has four compartments in the host population and three in the vector population; one additional in each population (E_h and E_v respectively) denoting the incubation / latent period. The host population (H) is now represented by a Susceptible-Exposed-Infected-Recovered ($S\,E\,I\,R$) model where $H = S_h + E_h + I_h + R_h$. The
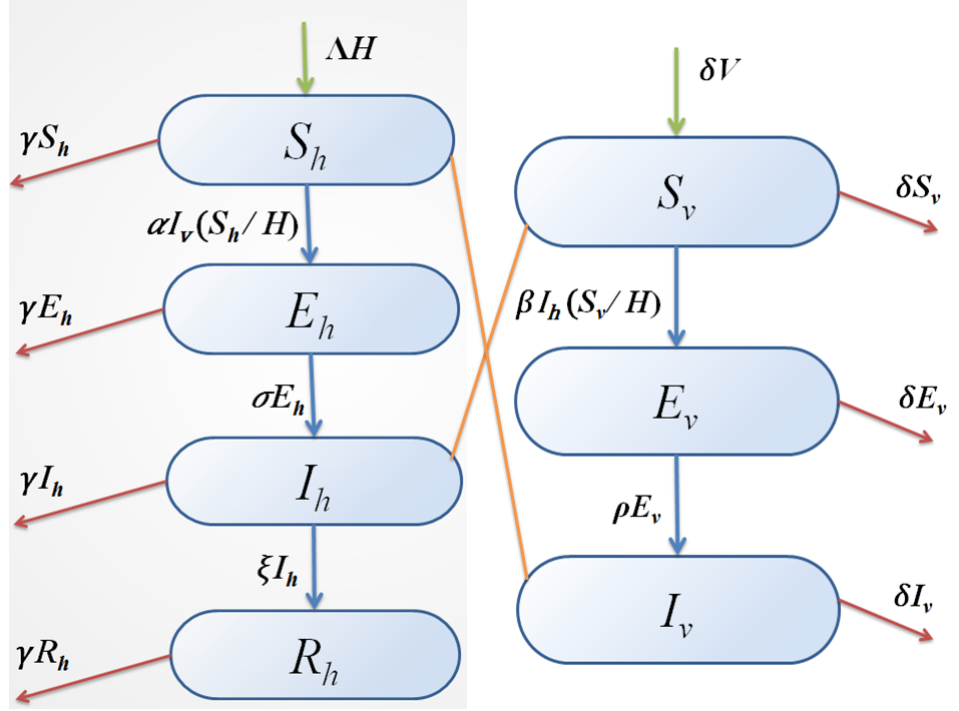
Figure 2.5: Schematic Diagram of the Ross Macdonald model with immunity in hosts and an incubation period in both populations. Line are coloured in the same way as Fig. 2.1

host birth and death rate are set equal so the population is closed.

The host model is represented by the following non linear, time-varying equations:

$$
\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= \Lambda H - \alpha I_v \left(\frac{S_h}{H} \right) - \gamma S_h \\
\frac{\mathrm{d}E_h}{\mathrm{d}t} &= \alpha I_v \left(\frac{S_h}{H} \right) - (\sigma + \gamma) E_h \\
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= \sigma E_h - (\xi + \gamma) I_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \xi I_h - \gamma R_h.
\end{aligned}
\tag{2.14}
$$

The vector population (V) is represented by a Susceptible-Exposed-Infected ($S\,E\,I$) model, where $V = S_v + E_v + I_v$. As mentioned previously in section 2.2, this model contains only female mosquitoes. Figure 4.1 shows the schematic diagram of the model and the equations for vectors are given as follows:
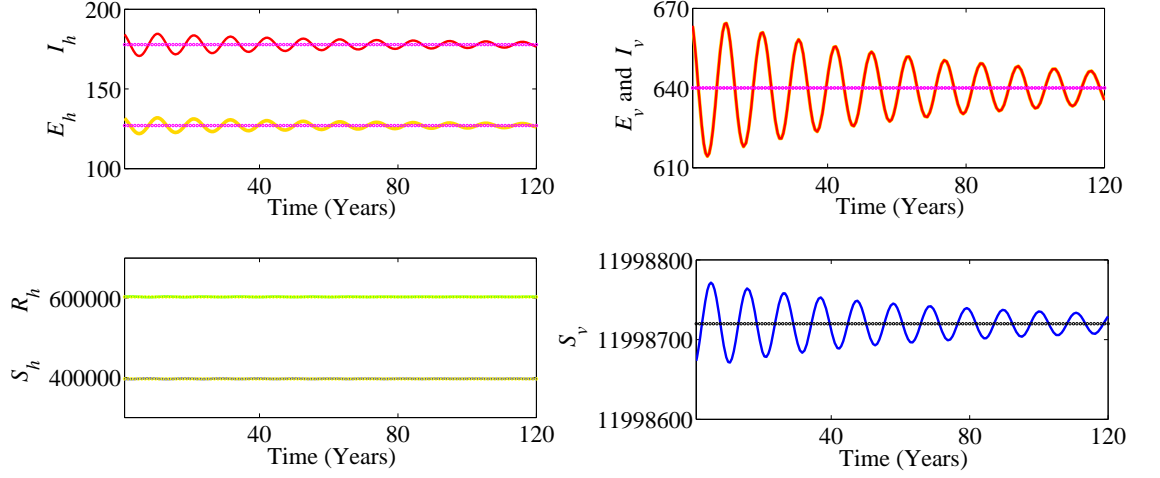
Figure 2.6: The deterministic solution of RM_{SEIR}. In the top row, from left to right, the figures shows latent and infectious individuals in host (E_h; *bottom line* and I_h; *top line*) and vector (E_v and I_v) compartments. In the bottom row, the first figure from the left shows susceptible (S_h) and recovered (R_h) individuals. The second figure shows the number of susceptible vectors (S_v). In all the plots, horizontal lines denotes the stable deterministic endemic equilibrium state. Host population size (H) is 1×10^6 individuals and vector to host ratio $\frac{V}{H}$ was set to twelve to have same value of R_0 as in RM_{SIR}. Model compartments show a cyclical behaviour to approach the endemic equilibrium state. Here all the parameters are the same as in Figure 2.3 and the average latent period in hosts and vectors are $\sigma = \frac{1}{5}$ and $\rho = \frac{1}{8}$ respectively.

$$\frac{\mathrm{d}S_v}{\mathrm{d}t} = \delta V - \beta S_v \left(\frac{I_h}{H} \right) - \delta S_v$$
$$\frac{\mathrm{d}E_v}{\mathrm{d}t} = \beta S_v \left(\frac{I_h}{H} \right) - (\rho + \delta)E_v \tag{2.15}$$
$$\frac{\mathrm{d}I_v}{\mathrm{d}t} = \rho E_v - \delta I_v.$$

The value of R_0 is (see Appendix A.1.2 for derivation):

$$R_0 = \sqrt{R_0^{VH} \times R_0^{HV}} = \frac{\alpha \rho}{\delta(\rho + \delta)} \times \frac{\beta \sigma V}{(\xi + \gamma)(\sigma + \gamma)H} = \sqrt{\frac{\alpha \beta \sigma \rho V}{\delta(\xi + \gamma)(\rho + \delta)(\sigma + \gamma)H}} \tag{2.16}$$

Here the basic reproductive number of the host population R_0^{VH} and vector population R_0^{HV} are explained in the same fashion as in section 2.2.1. The additional term $\frac{\rho}{\rho + \delta}$ in R_0^{VH} is the probability that a vector will survive the exposed state (E_v) and move to infectious compartment (I_v). In R_0^{VH}, $\frac{\sigma}{\sigma + \gamma}$ is the probability that a host will survive the exposed state (E_h) to enter the infectious state (I_h). R_0 acts as a general threshold for both populations and there is disease extinction when this number is less than 1 and either invasion and persistence

of disease, if this number is greater than 1. R_0^{HV} is enhanced by the factor $\frac{V}{H}$ if the number of vectors are greater than number of hosts. As the birth / death rate of humans, γ is very low, the basic reproductive number R_0^{HV} in RM_{SEIR} is not much different than in RM_{SIR}.

The numerical solution of RM_{SEIR} is given in Figure 2.6. The parameter values are the same as in Table 2.1 and the average incubation rate in hosts and vectors are $\sigma = \frac{1}{5}$ and $\rho = \frac{1}{8}$ respectively. The system of ODEs was integrated by initialization with 10% more latent and infected individuals compared with the equilibrium populations of hosts and vectors. These additional individuals are discounted from the susceptible host population. The solution for this model exhibits damped oscillations towards the endemic equilibrium state, similar to that was observed in Figure 2.3. R_0 is kept the same as previously by setting the vector to host ratio $\frac{V}{H}$ to twelve. These is a huge disparity between the number of E_h and infectious I_h in the host population whereas these two classes have the same count in the vector population. This is because the incubation and recovery rates are equal for the vector population but not for the humans.

2.3.1 Equilibrium points and local stability analysis

The system 2.14 and 2.15 admits two distinct equilibrium points, as observed in the previous model, (i) a DFE denoted by $E^0\,(H, 0, 0, 0, V, 0, 0)$, and (ii) an Endemic Equilibrium point (EE), denoted by $E^*\,(S_h^*, E_h^*, I_h^*, R_h^*, S_v^*, E_v^*, I_V^*)$. The endemic equilibrium points are:

$$
\begin{aligned}
S_h^* &= H - I_h^* - E_h^* - R_h^* \\
E_h^* &= \frac{(\xi + \gamma)}{\sigma} I_h^* \\
I_h^* &= \frac{\gamma}{\xi} R_h^* \\
R_h^* &= \frac{\xi H\,(\alpha\beta\sigma\rho V - \delta(\xi + \gamma)(\rho + \delta)(\sigma + \gamma)H)}{\beta\,(\xi + \gamma)\,(\sigma + \gamma)\,(\alpha\rho V - \gamma H\,(\rho + \delta))} \\
S_v^* &= V - E_v^* - I_v^* \\
E_v^* &= \frac{\delta}{\rho} I_v^* \\
I_v^* &= \frac{\beta\rho V I_h^*}{\left(\delta H + \beta I_h^*\right)(\rho + \delta)}
\end{aligned}
\tag{2.17}
$$

The solution is obtained by hand and verified in similar fashion as done in section 2.2.3. Here the analytical value of R_h^* is obtained first. The rest of the values are dependent upon

R_h^*. As stated before in section 2.2.3, these equilibrium points are biologically feasible when $R_0 > 1$.

These deterministic equilibrium points define the steady stationary state of the system. The DFE for RM_{SEIR} denotes the state of the system reached over passage of time when there are only susceptible individuals in the population. Similarly EE denotes the steady state of RM_{SEIR}. In the following two subsections, the local stability analysis of equilibrium points, presented in equation set 2.17 is performed. The local stability analysis is done on time-invariant baseline parameters, as discussed in section (2.2), with the addition of $\sigma = \frac{1}{5}\,\text{days}^{-1}$ and $\rho = \frac{1}{8}\,\text{days}^{-1}$. By putting the values used in plotting Figure 2.6, $R_0 = 1.6$, $S_h^* = 397068$, $I_h^* = 178$, $E_h^* = 127$ and $R_h^* = 602627$, $S_v^* = 11998720$, $E_v^* = 640$ and $I_v^* = 640$.

By comparing above values to the numerical estimate of the equilibrium points for RM_{SIR} in section 2.2.3, the susceptible proportion of hosts in RM_{SEIR} is slightly greater. This may leads to a major difference in the peak and timings of the epidemics in the long run for both models. This difference can be better studied using the stochastic settings. Moreover, the same number of infectious hosts are present at endemic equilibrium in both models, in addition to having the infected population E_h^* in RM_{SEIR}. Both of these factors shows that the infection in the host population of RM_{SEIR} is more stable. In the vector population, the number of individuals in I_v^* and E_v^* are equal as $\frac{\rho}{\rho+\delta} = 0.5$ at the steady state. This corresponds from the fact that the vector-to-host ration is set twice in RM_{SEIR} to keep R_0 same in both models. As in the host population, the infection in the vector population at the deterministic endemic equilibrium is supported from the E_v^* class.

2.3.1.1 Local stability of the disease free equilibrium

The system consisting of Eq. 2.14 and Eq. 2.15 has seven equations. Using the conditions mentioned in the previous paragraph, one compartment from both population can be dropped as $R_h = H - S_h - I_h - E_h$ and $E_v = V - S_v - I_v$. The Jacobian can be written as:

$$J\Big(S_h, E_h, I_h, S_v, I_v\Big) = \begin{pmatrix} -(\alpha\frac{I_v}{H} + \gamma) & 0 & 0 & 0 & -\alpha\frac{S_h}{H} \\ \alpha\frac{I_v}{H} & -(\sigma + \gamma) & 0 & 0 & \alpha\frac{S_h}{H} \\ 0 & \sigma & -(\xi + \gamma) & 0 & 0 \\ 0 & 0 & -\beta\frac{S_v}{H} & -(\beta\frac{I_h}{H} + \delta) & 0 \\ 0 & 0 & 0 & 0 & -\delta \end{pmatrix}.$$

$$\tag{2.18}$$

At the DFE, the number of infectious and infective individuals is zero and the number of

susceptible individuals is equal to the total population and there are no recovered individuals. Putting $E_h = I_h = E_v = I_v = 0$, $S_h = H$ and $S_v = V$ in the above matrix yields,

$$J\left(H,0,0,V,0\right) = \begin{pmatrix} -\gamma & 0 & 0 & 0 & -\alpha \\ 0 & -(\sigma+\gamma) & 0 & 0 & \alpha \\ 0 & \sigma & -(\xi+\gamma) & 0 & 0 \\ 0 & 0 & -\beta\frac{V}{H} & -\delta & 0 \\ 0 & 0 & 0 & 0 & -\delta \end{pmatrix}$$

The characteristic polynomial $\lambda I_5 - J\left(H,0,0,V,0\right)$ is used to find the eigenvalues of the disease-free state. By taking the determinant of the polynomial and equating it to zero, the eigenvalues of the system are acquired. Since this matrix is of order 5×5, the method of Laplacian expansion is used to reduce its order by one row and column at a time.

$$\lambda I_5 - J\left(H,0,0,V,0\right) = \begin{pmatrix} \lambda+\gamma & 0 & 0 & 0 & \alpha \\ 0 & \lambda+(\sigma+\gamma) & 0 & 0 & -\alpha \\ 0 & -\sigma & \lambda+(\xi+\gamma) & 0 & 0 \\ 0 & 0 & \beta\frac{V}{H} & \lambda+\delta & 0 \\ 0 & 0 & 0 & 0 & \lambda+\delta \end{pmatrix}$$

Expanding by the first column yields a fourth order determinant

$$\lambda I_5 - J\left(H,0,0,V,0\right) = (\lambda+\gamma)\begin{pmatrix} \lambda+(\sigma+\gamma) & 0 & 0 & -\alpha \\ -\sigma & \lambda+(\xi+\gamma) & 0 & 0 \\ 0 & \beta\frac{V}{H} & \lambda+\delta & 0 \\ 0 & 0 & 0 & \lambda+\delta \end{pmatrix} \quad (2.19)$$

Expanding the determinant in Eq. 2.19 by first column yields two determinants of order 3×3.

$$\left(\lambda+(\sigma+\gamma)\right)\begin{pmatrix} \lambda+(\xi+\gamma) & 0 & 0 \\ \beta\frac{V}{H} & \lambda+\delta & 0 \\ 0 & 0 & \lambda+\delta \end{pmatrix} + \sigma\begin{pmatrix} 0 & 0 & -\alpha \\ \beta\frac{V}{H} & \lambda+\delta & 0 \\ 0 & 0 & \lambda+\delta \end{pmatrix}$$

The determinant of the matrix at the right hand side is zero, so the final characteristic equation obtained from the Laplacian expansion of left hand matrix can be written as:

$$(\lambda+\gamma)\left(\lambda+(\sigma+\gamma)\right)\left(\lambda+(\xi+\gamma)\right)(\lambda+\delta)^2 = 0 \quad (2.20)$$

The eigenvalues obtained from Eq. 2.20 are all negative which guarantees the asymptotic stability of the system (Boyce and DiPrima, 1986).

Equation 2.20 is obtained by taking expressions from five out of seven compartments of RM_{SEIR}, mentioned in equation 2.14 and equation 2.15. The compartments which were excluded are R_h and E_v. The above system of equations can also be solved by dropping R_h and S_v. This new reduced system of equation leads to the following characteristic equation.

$$(\lambda + \gamma)(\lambda + \delta)\left((\lambda + (\sigma + \gamma))(\lambda + (\xi + \gamma))(\lambda + (\rho + \delta))(\lambda + \delta) - \alpha\beta\sigma\rho\frac{V}{H}\right) = 0 \quad (2.21)$$

Equation 2.21 can be simplified to

$$\lambda^4 + K\lambda^3 + L\lambda^2 + M\lambda + N = 0 \quad (2.22)$$

where K, L, M and N can be written as:

$$K = \xi + \sigma + \rho + 2(\gamma + \delta)$$
$$L = (\xi + \gamma)(\sigma + \gamma) + \delta(\rho + \delta) + (\xi + \sigma + 2\gamma)\delta(\rho + \delta)$$
$$M = \delta(\rho + \delta)(\xi + \sigma + 2\gamma) + (\xi + \gamma)(\sigma + \gamma)(\delta + (\rho + \delta))$$
$$N = \delta(\xi + \gamma)(\sigma + \gamma)(\rho + \delta)(1 - R_0).$$

For a fourth-order polynomial, the Routh-Hurwitz stability criteria imposes three conditions: (i) all the coefficients must be positive ($K, L, M > 0$), (ii) $KL > M$, and (iii) $KLM > M^2 + K^2N$. Here all the coefficients of equation 2.22 are positive, and N can remain positive only if the value of R_0 is less than 1. When $R_0 = 1$, one of the eigenvalues becomes zero as $N = 0$ and the DFE becomes unstable (Boyce and DiPrima, 1986). Therefore for $R_0 > 1$, this condition fails and $R_0 > 1$ acts as threshold condition required for the existence of disease free equilibrium state of RM_{SEIR}.

2.3.1.2 Local stability of the endemic equilibrium

Proceeding the same way as in the previous section, the Jacobian matrix at equilibrium points of the compartments can be written as:

$$J\left(S_h^*, E_h^*, I_h^*, S_v^*, I_v^*\right) = \begin{pmatrix} -(\alpha\frac{I_v^*}{H} + \gamma) & 0 & 0 & 0 & -\alpha\frac{S_h^*I_v^*}{H} \\ \alpha\frac{I_v^*}{H} & -(\sigma + \gamma) & 0 & 0 & \alpha\frac{S_h^*I_v^*}{H} \\ 0 & \sigma & -(\xi + \gamma) & 0 & 0 \\ 0 & 0 & -\beta\frac{S_v^*}{H} & -(\beta\frac{I_h^*}{H} + \delta) & 0 \\ 0 & 0 & 0 & 0 & -\delta \end{pmatrix}$$

$$(2.23)$$

The characteristic equation for finding the eigenvalues at deterministic equilibrium point can be obtained by solving the following matrix equation.

$$\lambda I_5 - J\left(S_h^*, E_h^*, I_h^*, S_v^*, I_v^*\right) =$$
$$\begin{pmatrix} \lambda + (\alpha\frac{I_v^*}{H} + \gamma) & 0 & 0 & 0 & \alpha\frac{S_h^* I_v^*}{H} \\ -\alpha\frac{I_v^*}{H} & \lambda + (\sigma + \gamma) & 0 & 0 & -\alpha\frac{S_h^* I_v^*}{H} \\ 0 & -\sigma & \lambda + (\xi + \gamma) & 0 & 0 \\ 0 & 0 & \beta\frac{S_h^*}{H} & \lambda + (\beta\frac{I_h^*}{H} + \delta) & 0 \\ 0 & 0 & 0 & 0 & \lambda + \delta \end{pmatrix}$$
$$(2.24)$$

Solution of this system yields a fifth order equation, here represented in factor form.

$$\left(\lambda + \alpha\frac{I_v^*}{H} + \gamma\right)\left(\lambda + \beta\frac{I_h^*}{H} + \delta\right)(\lambda + (\sigma + \gamma))(\lambda + (\xi + \gamma))(\lambda + \delta) = 0 \qquad (2.25)$$

All the eigenvalues from characteristic equation, i.e, the values of λ are negative. Moreover, if $R_0 < 1$, then the fractions $\frac{I_h^*}{H}$ and $\frac{I_v^*}{H}$ becomes negative, which is invalid by the definition of the size of the populations. At $R_0 = 1$, $\frac{I_h^*}{H} = \frac{I_v^*}{H} = 0$ and this equation transforms into equation 2.20 which is not defined for the endemic region (see the argument for the value of R_0 to be substituted for N in the previous section). Thus the endemic equilibrium will always be stable for $R_0 > 1$. The explicit dependency on the value of R_0 can be shown by substituting the values of $\alpha\frac{I_v^*}{H}$ and $\beta\frac{I_h^*}{H}$ as undertaken previously in section 2.2.3.2.

2.4 Conclusion and discussion

This chapter is devoted to examining the deterministic models for host-vector systems. The aim here is to look at the mathematical behaviour of simple models and use the analysis undertaken in this chapter as a foundation for the next chapters. Two models, one baseline RM_{SIR} (SIR for hosts and SI for vectors) and a second having incubation periods RM_{SEIR} ($SEIR$ for hosts and SEI for vectors) are developed and analysed. In order to avoid over-parametrization and to include the minimum number of equations, these models include one viral serotype with one species of vector and do not include aquatic / premature stages of the vector. Host and vector population sizes in both models are held constant over time, by

having the same birth and death rate. The equilibrium points of both models are checked for stability and found to be locally asymptotically stable.

The basic reproductive number R_0 is found to be a threshold parameter between the disease free state and endemic state for both models. The basic reproductive ratio is split in to two different population level reproductive numbers, $R_0^{VH} = \frac{\alpha}{\delta}$ for hosts and $R_0^{HV} = \frac{\beta V}{(\xi + \gamma) H}$ for the vector population in RM_{SIR}. The population level reproductive numbers in the latter model are derived by using the same analogy as for the baseline model. In RM_{SEIR}, R_0^{VH} is multiplied by the factor $\frac{\sigma}{\sigma + \gamma}$ which denotes the probability that a host will survive from the exposed state to become infectious. In similar fashion, the basic reproductive number for the vector population R_0^{HV} is multiplied by the factor $\frac{\rho}{\rho + \delta}$. By substituting the values from Table 2.1, $\frac{\sigma}{\sigma + \gamma} \approx 1$ and $\frac{\rho}{\rho + \delta} = 0.5$. This implies that the probability that an infected host becomes infectious is almost 100% whereas 50% of the infected vectors die before becoming infectious. So R_0^{HV} of both models is nearly the same and R_0^{VH} in $R_0^{RM_{SEIR}}$ is multiplied by a factor of 0.5. In order to compare the results of both models on the basis of having equal basic reproductive numbers, the vector-to-host ratio is set to be twelve in RM_{SEIR}. At the endemic equilibrium, the infectious components in both models, i.e, I_h and I_v have the same number of individuals. In RM_{SEIR} at equilibrium, the host population has 0.0127% of the individuals in the latent compartment whereas the number of incubating vectors is same as the number of infectious vectors (see Figure 2.6). As R_0 is the threshold that bifurcates the DFE and EE at $R_0 = 1$, this denotes that the DFE of RM_{SEIR} is more stable and RM_{SIR} is more stable at the endemic equilibrium.

The number $R_0 = 1$ is both invasion and persistence threshold in the deterministic model but it is worthwhile to mention that although its definition is broad, there are different methods used for the estimation of R_0 (see Heffernan et al. (2005) for an overview), therefore different diseases cannot be compared unless the method employed to estimate R_0 is the same. If R_0 is to be a set as a threshold and it measures the average secondary cases caused by a single infectious individual in an otherwise susceptible population, then the existence of endemic equilibrium should be conditioned upon $R_0 > 1$ (Li et al., 2011). These authors further pointed out that the issues such as backward bifurcations and spatial structure of the population limit the use of R_0 as a persistence threshold. A clear indication of which method is used to estimate R_0 along with the underlying assumptions is important to compare results from different studies. In addition to these issues, it does not provide a time dependent measurement of the spread of the disease. Moreover, the depletion of susceptible individuals is not included in the formulation of R_0.

In this chapter, I have compared the results from two epidemiologically different host-vector

models. For this purpose, models are parametrised in such a way that most of the quantities remain same or similar. The basic reproductive number R_0 is kept same in both models by varying the vector-to-host ratio from 6 (RM_{SIR}) to 12 (RM_{SEIR}). The results related to R_0 and stability at DFE and EE are comparable to the models considered in (Esteva and Vargas, 1998, 2000; Pinho et al., 2010). The second model RM_{SEIR} is not much found in dengue literature but is important for other non-seasonal diseases that have incubation period in both host and vectors. In this case, including RM_{SEIR} can be important for generating the realistic patterns of the disease.

Although deterministic models provide an overview of the dynamical evolution of system, these models are unable to answer the questions related to the persistence and extinction of the disease. Therefore, the models considered here lack the 'real-world' behaviour of any host-vector disease due their deterministic nature. The solution trajectories always followed the same path if the model is solved using the same initial conditions and parameter values. In the real world, completely different results can be seen over time, even if all the initial conditions are same. The parameter values taken here are suitable for a scenario with endemic circulation of the pathogen with no sudden depletion of susceptible individuals which halts transmission. So the parameter values chosen in both models are not intended for examining detailed epidemiological patterns of dengue, but for looking at endemic persistence of dengue virus in host-vector settings over longer periods.

The basic question addressed in this thesis surrounds the persistence of a pathogen in a host-vector system which cannot be fully addressed by using a deterministic framework. Proper study of disease extinction from a population requires a stochastic framework of modelling. The current chapter has laid the theoretical foundations by proving that the model developed are mathematically sound for additional analyses. For further investigation, the next chapter deals with the stochastic counterpart of these models.

CHAPTER 3

Determinants of long-term pathogen persistence in host-vector systems

Determinants of long-term pathogen persistence in host-vector systems

3.1 Introduction

Vector-borne diseases (VBDs) are an important subset of multi-host pathogens. These diseases infect more than a billion people a year and kill more than one million per annum (Butler, 2013). The work of Bartlett (1957, 1960) relates the size of the human population with the fade-out of measles dynamics in USA and UK. The persistence of measles was observed as an increasing function of the population size. For vector-borne disease, the need to understand the mechanisms that relate the pathogen persistence and host population size for the VBDs is important for population ecologists and epidemiologists. As vector control is sometimes the most applicable strategy for the elimination of the VBDs (Jansen and Beebe, 2010), the work done in the current study focuses on the investigation of the factors which affect pathogen persistence and functional form of these relations. The objective is to establish the relationship between the size of the population and the persistence of the pathogen in host-vector systems.

The models used to investigate the relationship are the Ross Macdonald model with host immunity (RM_{SIR}) and Ross Macdonald model with latent periods and immunity (RM_{SEIR}). The key aim is to establish to what extent persistence of a pathogen in the host population, formally measured here as the threshold Critical Community Size (CCS), depends on the parameters of the host-vector system. For measuring probability of disease extinction in human and vector populations, individual-based models were simulated using tau-leap approximations of the Gillespie Algorithm. These were used to find the CCS for both models. To further investigate the association between parameters of the models and CCS, sensitivity analysis was performed and linear regression was used to quantify the relationship between CCS and its determinants. A detailed introduction to persistence thresholds and CCS follows.

3.2   Persistence thresholds in host-vector system

This section begins with an overview of persistence thresholds and their relationship to pathogen persistence. Then, the development of the CCS concept in the population biology and epidemiological literature is presented, followed by a description of the ecological factors affecting it. Towards the end, the concept of CCS is extended to the study of persistence in vector-borne diseases. The ideas related to pathogen persistence and thresholds presented in the following subsections are mainly in the context of vector-borne diseases. The focus of the discussion here is to construct a foundation for estimating persistence thresholds in interacting vector and host populations, with an emphasis on mosquito borne diseases.

3.2.1   Persistence thresholds

In mathematical modelling, a threshold is viewed as the point in a dynamical system where a quantitative change in parametrization leads to qualitative change in the system behaviour (Deredec and Courchamp, 2003). In theoretical epidemiology, the main threshold used in deterministic models is denoted by the basic reproductive ration, R_0 which enforces a boundary in parameter space at 1. R_0 is the average number number of individuals infected by a single infected individual in an entirely susceptible population (Heffernan et al., 2005). For a fully mixed population and simple disease dynamics, when $R_0 \leq 1$, any existing infection in the population will vanish whereas if $R_0 > 1$, the population can support a positive endemic infection level in the deterministic system. Thus R_0 decides the fate of the infection in the communities if the procedure of modelling is deterministic.

In the stochastic setting, *time to extinction* is used as a threshold that separates populations that can retain infection longer from the populations in which disease extinction occurs quickly. As described by Nåsell (2011), time to extinction depends upon the initial number of infected individuals. If the initial conditions are sampled from the Quasi-Stationary Distribution (QSD) then the extinction times follow an exponential distribution whose expected value can be used to provide a measure of persistence. The quasi-stationary state can be thought of as a stochastic counterpart of the deterministic equilibrium state. From the perspective of endemic infections, the remaining time to extinction of infection can be found if the system is in quasi-stationary state (e.g., see Nåsell (2002) for details).

Mancy (2015) described a way of estimating the persistence threshold until a particular time horizon and estimating the critical size of a community required to support a population that persists with a given probability, to at least this time (usually 50%). In the context

of infectious diseases, assessing persistence thresholds offers a possible way to seek measures that includes running public health campaigns and vaccination strategies. For example, consider an outbreak of a disease in a certain locality. The government departments and policy planners would need to allocate funds for public awareness and vaccination. In this case, it would be important to consider whether vaccinating a small community is really justified or not (see Beyer et al. (2012) for example). Due to their applications in above mentioned areas, persistence thresholds are estimated for a range of different classical and meta-population models [for example (Nåsell, 2005), (Lloyd-Smith et al., 2005) and (Hanski and Ovaskainen, 2000)].

3.2.2 Critical community size (CCS)

In the seminal work of Bartlett (Bartlett (1957) and Bartlett (1960)), the persistence of measles in England and US was observed as an increasing function of population size. In densely populated areas, fewer disease fade-outs were observed, providing a measure of the threshold for persistence of the pathogen in a population, which Bartlett called Critical Community Size (CCS). This concept has become one of the central measures of pathogen persistence and is considered an intrinsic property of epidemiological dynamics in transmission mechanisms for homogeneous or heterogeneous populations.

In the literature, CSS has been defined in various ways, but is intended to reflect the smallest number of individuals in a population where disease persists without reintroductions from an external source [Haydon et al. (2002) and Viana et al. (2014)]. Many diseases display cyclic dynamics. Populations in which a pathogen predominantly persists can be separated from smaller communities where frequent extinctions are observed between major epidemics. Stated another way, the concept of CCS characterises the relationship between population size and probability of extinction during inter-epidemic periods. It was found that the incidence of measles scales linearly with host population size (Grenfell et al., 2002) as the chain of transmission remains uninterrupted between the troughs of epidemic oscillations.

The long-term persistence of disease is not guaranteed after successful invasion to the population. In the absence of external forcing, pathogens go extinct due to two processes: (*a*) epidemic fadeout after a major epidemic that creates susceptible bottlenecks, caused by a too high basic reproductive number leading to depletion of the susceptible pool before demographic turnover and (*b*) endemic fade-out caused by too low basic reproductive number that creates transmission bottlenecks causing interruption in the chain of transmission leading to the stochastic extinction of the infection. CCS can be modified by many factors such as the

distribution of waiting times for infection and latency, heterogeneity in transmission and the spatial distribution of the population, existence of a reservoir, demographic turn-over, age-structured transmission levels and seasonal fluctuation (Keeling and Grenfell, 1997), (Keeling and Grenfell, 1998), (Lloyd, 2001), (Lloyd et al., 2007), (Conlan et al., 2010) and (Peel et al., 2014).

All of the aforementioned factors can alter the threshold required by a pathogen to persist in a population and hence, CCS. However, the notion of CCS lacks 'completeness' and requires further theoretical definition. According to Conlan et al. (2010), comparing CCS between studies is difficult, mainly because of the chosen measure of persistence and assumptions of the stochastic models employed to estimate it. There are multiple issues in the current definition of CCS, including those that are conceptual, operational, inferential and measurement related. Although a pervasive concept in human and wildlife epidemiology (Peel et al., 2014) and now being broadly used as a general term for any population threshold for disease persistence (Lloyd-Smith et al., 2005), it is difficult of operationalize (Viana et al., 2014). The concept of CCS is unable to capture some epidemiological characteristics, like connectedness of subpopulations and the terms 'fade-out' and 'major' are vague in the classical definition (Mancy, 2015). CCS also provides no information about the distribution around persistence probabilities, since typically a 50% probability of pathogen persistence is used to estimate the population size (Mancy, 2015). Finally, it is problematic to estimate CCS by definition, as it requires several instances of fade-out in communities of different sizes, which are unlikely to occur within the time frame of a study.

Nasell published a series of papers estimating the mean time to extinction using different compartment models (Nåsell, 1999, 2002) and derived analytic expressions for the quasi-stationary distribution. In more recent work (Nåsell, 2005), CCS was defined as "that value of N for which the probability of extinction after waiting for one quasi-period T_0 equals 0.5", where quasi-period refers to the length of time between two local maxima. The waiting time in Bartlett's work starts after a major epidemic Bartlett (1957). In contrast to Bartlett, the waiting time in Nåsell (2005) starts from an initial distribution which is approximately quasi-stationary. This indicates that the Nasell's approach is suitable for studying persistence thresholds when the disease is endemic in the population. By using the definition of persistence threshold presented in Nåsell (2005) and making some additional assumptions, the CCS can be derived as a function of R_0 (Nåsell, 2005).

For vector-borne diseases, the concepts of persistence thresholds and CCS remain relatively unexplored (Swinton et al., 2002) and only a few studies address persistence and extinction in host-vector systems [e.g. (Deredec and Courchamp, 2003),(Lloyd et al., 2007) and (De

Castro and Bolker, 2004)]. The effect of spatial heterogeneity for dengue is considered in (de Castro Medeiros et al., 2011) with a stochastic cellular automata model in a meta-population. Results suggest that the low transmission rates of the virus with connected localized human population structure can help sustain the virus for extended periods, even with a moderate vector-to-host ratio. de Castro Medeiros et al. (2011) modified the previous findings which states that dengue virus can be maintained by low house-indices (number of larvae or pupae per house) (Newton and Reiter, 1992). Adams and Boots (2010) used mathematical models for observing the effect of transovarial transmission on the persistence of dengue. Their model shows that the rates of vertical transmission required for the maintenance of dengue virus in inter-epidemic periods need to be higher than reported by the laboratory and field studies.

In contrast to most of the systems studied for CCS, infection in host-vector settings is mainly transmitted between host and vector populations and usually there is no direct host-to-host transmission. This induces a paired distribution of disease, where number of infected individuals in the two populations are correlated to each other. Moreover, as demographic turnover of the host and vector vary with the infectious and recovery rates, dynamics of such systems can be quite sensitive to the parameter values and their combinations. Identifying the determinants of pathogen persistence in host-vector systems is a difficult but important problem. Proper understanding of persistence can lead to better understanding of extinction dynamics for many host-vector disease and, as a result, to interventions and control.

3.3   Introduction to stochastic models

In contrast to deterministic solutions, stochastic models produce a range of outcomes as the infection and recovery of individuals varies randomly. As a result, model realizations follow different trajectories with different repetitions even with the same parameter values and initial conditions. Since the population size can only take integer values, stochastic disease extinction may occur even if the basic reproductive number R_0 is greater than 1, as a sequence of random events can drive the infection to zero; for example, too many individuals can recover before transmitting the disease to any susceptible.

This class of models is used to investigate stochastic fade-out and extinction of disease in a population, which is not possible in deterministic settings as infective numbers can drop to extremely low values ($<< 1$) in the inter-epidemic trough, only to bounce back showing artificial persistence of the disease. As mentioned by Lloyd, pathogen persistence is most naturally studied in an event-driven stochastic systems (Lloyd, 2001). These features of stochastic models are useful; in particular for ecologists and epidemiologists for policy making

where elimination and eradication of a pathogen is of interest.

3.4   Introduction to sensitivity analysis

The results obtained from mechanistic models face a main source of uncertainty which is related to the assumed parameter values in the model. The parameter values may not represent the exact phenomena or process which gives the observed output, introducing variability in the model's prediction of resulting dynamics. The statistical inference from mechanistic models in disease dynamics is sometimes very complicated as the model fits lack the information necessary to quantify the influence of individual parameters. In dynamical systems, learning about the influence of the parameters on a model's behaviour is of much interest. This requires measuring a quantitative relationship between the dependence of model's output and changing parameter values.

The sensitivity S of outcome ϕ to the value of parameter θ is the partial rate of change in ϕ with respect to θ, while holding all other parameters constant:

$$S = \frac{\partial \phi}{\partial \theta}.$$

Parameters have different units, so it is preferable to employ the proportional response to a proportional perturbation, formally known as elasticity. The elasticity E of outcome ϕ to the value of parameter θ is

$$E = \frac{\theta}{\phi} \frac{\partial \phi}{\partial \theta} = \frac{\partial \log \phi}{\partial \log \theta}.$$

In modelling infectious diseases, sensitivity and elasticity analyses can be used for understanding the relative importance of different mechanisms in generating observed patterns and identifying the core parameters. Results from this type of analysis help planning the measurement of parameters that are most influential and targeted for intervention and policy making. In addition, sensitivity analysis can help reduce the parameter space of complex models (Marino et al., 2008). As a result, only the most important parameters are considered that also helps reducing the cost of surveys and field studies.

3.5   Models and methods

This section provides an overview of the models used and the methodology employed during the rest of the study. Structured in three sub-sections, it opens with an introduction to the

stochastic versions of RM_{SIR} and RM_{SEIR}. A brief description of the tau-leap algorithm and detail of the initial conditions of the simulations constitutes the second part. Finally, the last section is dedicated to a concise outline of the techniques used for the sensitivity analysis of the parameters and the Partial Rank Correlation Coefficients (PRCC) technique. Details related to the tau-leap algorithm, sensitivity analysis and PRCC are discussed in detail in the appendices (B.1, B.3 and B.4) respectively.

3.5.1 Stochastic version of Ross-Macdonald models

In the following, a simple host-vector model RM_{SIR}, (SIR in hosts and SI in vectors) is used as a framework for investigating the persistence of a pathogen. The model is parametrized for dengue, although it is acknowledged that it is a very simple approximation of the complicated process which governs dengue dynamics. It is anticipated that a simple model helps uncovering the determinants of CCS in host-vector system. The work is further extended to a model with the inclusion of the latent class in both populations, here termed as RM_{SEIR}. This is done for two different reasons, the first being that dengue disease has a period of incubation in hosts and vectors and the second being that the SIR and $SEIR$ models are the most commonly investigated models in epidemiology for directly transmitted diseases and their treatment in host-vector systems is an interesting problem in its own right. Table 3.1 shows the parameters and their interpretation used in both models.

Symbol	Explanation	Value used
α	Vector-to-host transmission rate, days^{-1}	0.1
β	Host-to-vector transmission rate, days^{-1}	0.075
$\frac{1}{\xi}$	Average infectious period in hosts, days	7
Λ, γ	Birth / death rate of hosts, days^{-1}	4.215×10^{-5}
$\frac{1}{\sigma}$	Average latent period in hosts, days	5
δ	Birth / mortality rate of vectors, days^{-1}	0.125
$\frac{1}{\rho}$	Average latent period in vectors, days	8

Table 3.1: Parameters used in the models. Here the average birth rate Λ of hosts is equal to the average death rate γ. The average host life expectancy is set to $\frac{1}{65}$ years, which is expressed in days^{-1} in the model. The parameters representing the latent periods in host and vectors are used only in RM_{SEIR}.

The deterministic rates are replaced by transition rates in the stochastic version of RM_{SIR} and RM_{SEIR}. Transition rates affect the number of individuals entering or leaving the compartments. The durations of latency and infection follow an exponential distribution which means that there is no 'typical' duration of these events. In other words, a few individuals

Event	Effect	Transition Rate
Host :		
(i) Birth of susceptible host	$S_h \uparrow$	ΛH
(ii) Infection of susceptible host	$S_h \downarrow \,\&\, I_h \uparrow$	$\alpha I_v S_h / H$
(iii) Recovery of infected host	$I_h \downarrow \,\&\, R_h \uparrow$	ξI_h
(iv) Natural death of a host	$S_h \downarrow,\, I_h \downarrow,\, R_h \downarrow$	$\gamma S_h,\, \gamma I_h,\, \gamma R_h$
Vector :		
(i) Birth of susceptible vector	$S_v \uparrow$	δV
(ii) Infection of susceptible vector	$S_v \downarrow \,\&\, I_v \uparrow$	$\beta S_v I_h / H$
(iii) Death of a vector	$S_v \downarrow \,\&\, I_v \downarrow$	$\delta S_v,\, \delta I_v$

Table 3.2: The events in the stochastic RM_{SIR} model are shown as stochastic transition rates. Here the subscript h denotes host population and v denotes vector population. The direction of the arrow in second column *'Effect'* denotes either addition or subtraction of an individual to and from a compartment.

can remain in these states for a very long time, whereas most of them remains in these states for a shorter time. The exponential distribution is chosen for mathematical convenience because of its 'memory-less' property. The current state of the stochastic system is system is fully defined by the number of individuals present in each state and information of past events in not considered. Stochastic events are summarized in Tables 3.2 and 3.3. Births and deaths in both populations are considered as distinct events which results in the random walk fluctuation in the total population sizes H and V.

3.5.2 Tau-leap method

There are different methods for finding the solution of the models described in section 3.5.1, ranging from symbolic procedures, numerical techniques or estimation of the symbolic solution using mathematical approximations. Mostly, the choice of method is influenced by the complexity of the reactions and size of the state-space (Érdi and Lente, 2014). Generally host-vector systems are inherently complex in structure and CCS estimation requires dealing with large population sizes. Analytic solutions to problems of these types are hard to find, and event-driven stochastic simulation is probably the most natural way to study persistence in these system. The solution of the stochastic systems shown above therefore is sought using Monte-Carlo simulations. The Gillespie Algorithm (Gillespie, 1977) is one of the widely used Stochastic Simulation Algorithms (SSA) and is used in finding accurate solutions to many ecological problems, especially related to population dynamics. The only drawback of the Gillespie Algorithm is that as population size and number of events increases, the time to next event gets smaller and event selection becomes slower; hence the algorithm requires

Event	Effect	Transition Rate
Host :		
(i) Birth of susceptible host	$S_h \uparrow$	ΛH
(ii) Exposure of susceptible host	$S_h \downarrow$ & $E_h \uparrow$	$\alpha I_v S_h / H$
(iii) Infection of exposed host	$E_h \downarrow$ & $I_h \uparrow$	σE_h
(iv) Recovery of infected host	$I_h \downarrow$ & $R_h \uparrow$	ξI_h
(v) Natural death of a host	$S_h \downarrow$, $E_h \downarrow$, $I_h \downarrow$, $R_h \downarrow$	γS_h, γE_h, γI_h, γR_h
Vector :		
(i) Birth of a susceptible vector	$S_v \uparrow$	δV
(ii) Exposure of susceptible vector	$S_v \downarrow$ & $E_v \uparrow$	$\beta S_v I_h / H$
(iii) Infection of exposed vector	$E_v \downarrow$ & $I_v \uparrow$	ρE_v
(iv) Death of a vector	$S_v \downarrow$, $E_v \downarrow$, $I_v \downarrow$	δS_v, δE_v, δI_v

Table 3.3: The events in stochastic RM_{SEIR} are shown as stochastic transition rates. All the parameters are same as in Table 3.2 except for σ and ρ which define the duration an individual spends in the exposed class for host and vectors respectively.

extensive computational resources.

A computationally attractive but approximate method is proposed by (Cao et al., 2006), which is known as the tau-leap approximation method. Tau-leap approximations make use of the cumulative rate of change in the population: if the rate of change is low, the algorithm 'leaps' through many states, hence speeding up the simulations. Two types of leap approximations, fixed tau-leap (Cao et al., 2006) and adaptive tau-leap (Cao et al., 2007) are available and most of the results were obtained by using the adaptive tau-leap due to its faster simulation times. In adaptive tau-leap, the error control parameter ϵ is used to bound the relative change in each compartment. This parameter should be set by taking into account the highest order of the reaction (loosely speaking, the order of a reaction is determined by counting the number of variables involved in all products in the differential equation for the system, and selecting the maximum of those). In both models, highest order rates were $\alpha I_v S_h / H$ and $\beta S_v I_h / H$ having order 2 therefore ϵ should be less than $\frac{1}{2}$ (see derivation in B.2). In the simulations, ϵ was set to 0.01 to avoid any computational discrepancies. All simulation experiments in the coming sections were carried out by using one of the tau-leap approximations. The computational scheme for the tau-leap method is given in Appendix B.1.

3.5.3 Estimation of CCS

The stochastic simulations were started from deterministic steady state values in each compartment, representing the deterministic analogue of the quasi-stationary distribution, and were allowed to run for twenty-five years. Two hundred repeated simulations were performed for estimating CCS. The host population was started form one hundred thousand individuals and then from 0.25 to 1.5 million individuals with intervals of 0.25 million as shown in Figures 3.1 and 3.2. The vector population V varied from one million to fifteen million individuals in the intervals of one million. The $R_0 = c \times \frac{V}{H}$ varied with respect to the vector-to-host ratio for both models. The proportion of runs which retained infection after twenty-five years, out of 200 stochastic repetitions, was used to estimate the probability of extinction $P(E)$ for the population size. CCS was selected as the population size (H) for which the $H - V$ pair yielded $P(E) = 0.5$.

3.5.4 Sensitivity analysis

Two methods were used to investigate the relationship between CCS and parameter values. These were Latin Hypercube Sampling (LHS) and estimating the Partial Rank Correlation Coefficients (PRCC) for each parameter and CCS. The next sub-sections provide an overview of both techniques.

The parameters of the model were sampled using LHS [see B.3 for details of the process]. The relationship between each parameter and CCS was plotted in monotonicity plots which were made as follows: (i) setting all the parameter values at the centre of the hypercube at their baseline values, (ii) then a parameter is chosen and the relationship between CCS and that parameter is explored by varying it from its minimum value to its maximum value while the remaining parameters remain constant (local sensitivity analysis).

Although the monotonicity plots provide information about the parameters having a strong effect on the response variable, they do not provide precise details of the strength of the relationship between the response and individual explanatory variable. Partial Rank Correlation Coefficients (PRCC) methodology is a non-parametric method used with LHS for this purpose [Blower and Dowlatabadi (1994), Blower et al. (1991) and Sanchez, M. A., & Blower (1997)] where statistics are applied on the ranks of the data rather than on data values. In the PRCC method, first we choose a parameter of the model. Then two regression models are constructed, one having that parameter as a response and other having CCS as a response variable. Remaining parameters as explanatory variables in both regression models.

Pearson's correlation coefficient is calculated on residuals from both models to obtain PRCC. Further details of this method are given in appendix B.4.

In addition to estimating the affect of individual parameters on CCS, a functional form of CCS in terms of the parameters of the models was undertaken. To identify the determinants of CCS, the data from the LHS design was used for the construction of general linear models. Two different parameter types are considered as potential explanatory variables, viz., (i) primary parameters comprises $\alpha, \beta, \xi, \delta, \Lambda$ & γ, (σ & ρ in addition for RM_{SEIR}) and (ii) secondary parameters, $R_0, R_0^{VH}, R_0^{HV}, S_h^*, I_h^*, R_h^*, S_v^*, I_v^*$, N and λ. All secondary parameters were simple algebraic combinations of primary parameters.

3.6 Results

This section reports the important findings in this chapter. First, the results of CCS for baseline values of the parameters are reported followed by sensitivity analysis where the impact of parameter variability on the CCS is explored. Through the following sub-sections, the results for RM_{SIR} and RM_{SEIR} are reported simultaneously for a better comparison between these models.

3.6.1 CCS for Ross-Macdonald models

The baseline parameters for both models described in Sections 2.2 and 2.3 are used to find CCS. CCS is defined as the minimum host population size for which half of the stochastic simulations still retain either infected hosts or vectors after twenty-five years. Figure 3.1 and 3.2 shows the variation of R_0 and CCS for the baseline parameters in RM_{SIR} and RM_{SEIR} respectively. Comparing the top heat maps in both figures shows that the range of values of R_0 are lesser in RM_{SEIR}. This is due to the fact that the formula of R_0 for RM_{SEIR} has two additional terms, $\frac{\sigma}{\sigma+\gamma}$ and $\frac{\rho}{\rho+\delta}$ as shown in equation 4.2. The first term is almost unity as the value of γ is very small, whereas $\frac{\rho}{\rho+\delta} = \frac{1}{2}$. This leads to smaller values of R_0 for RM_{SEIR} as compared to RM_{SIR} for same $H - V$ pairs and twice as many vectors are required in RM_{SEIR} to have the same value of R_0.

The bottom heat maps in both Figures 3.1 and 3.2 denotes the change in CCS with respect to the $H - V$ pair. The $H - V$ pair that result the value of $R_0 \leq 1$ leads to 100% probability of extinction as the stochastic simulations are started from the deterministic endemic equilibrium. In Figure 3.1, CCS was not found when the host population H was one hundred
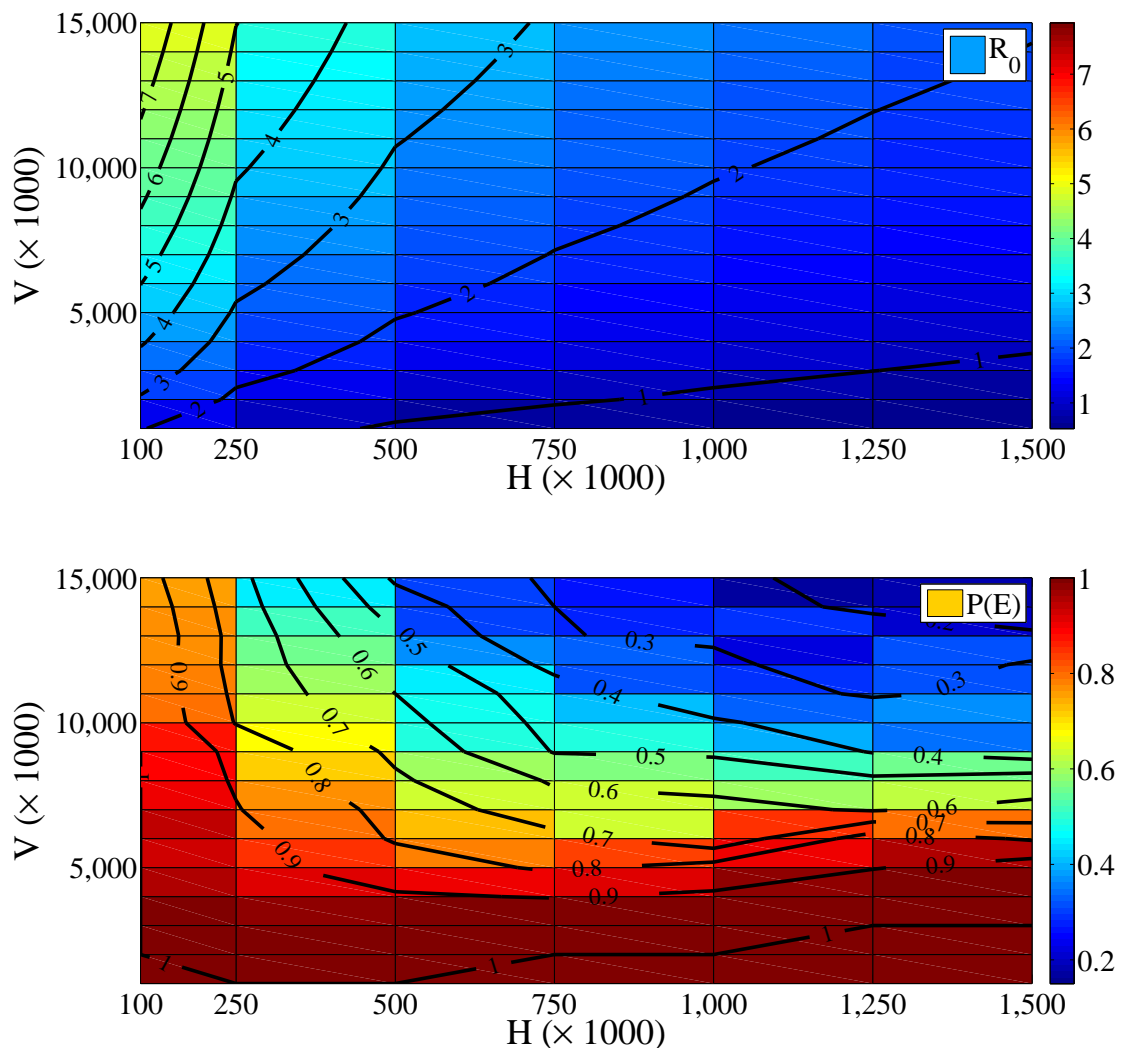
Figure 3.1: Heat maps showing the basic reproductive number R_0 (top) and probability of extinction $P(E)$ (bottom) of dengue for different host and vector population sizes in RM_{SIR} over twenty-five years. Simulations were started at deterministic endemic equilibrium and were repeated two hundred times. Host population size (H) was initially 1×10^5, then 2.5×10^5 and then it was increased in intervals of 2.5×10^5 till it reached 1.5 million individuals. The vector population V was initially one million and was increased in intervals of one million till fifteen million individuals. As $R_0 = c \times \frac{V}{H}$, the vector-to-host ratio $\frac{V}{H}$ was altered according to the $H - V$ pair. CCS was defined as the population at which half of the stochastic simulations still retained either infected hosts or vectors after twenty-five years. Constant parameters were $\alpha = 0.1$, $\beta = 0.075$, $\xi = \frac{1}{7}$, $\delta = \frac{1}{8}$, and $\Lambda = \gamma = 4.5 \times 10^{-5}$.

thousand individuals. At 250,000 hosts, fifteen million vectors are required to attain CCS. At small host populations, the number of individuals in the infected compartment I_h are very low that leads to extinction of disease in the hosts. For five hundred thousand to one million hosts, the number of vectors require to attain CCS become smaller. Finally, they remain
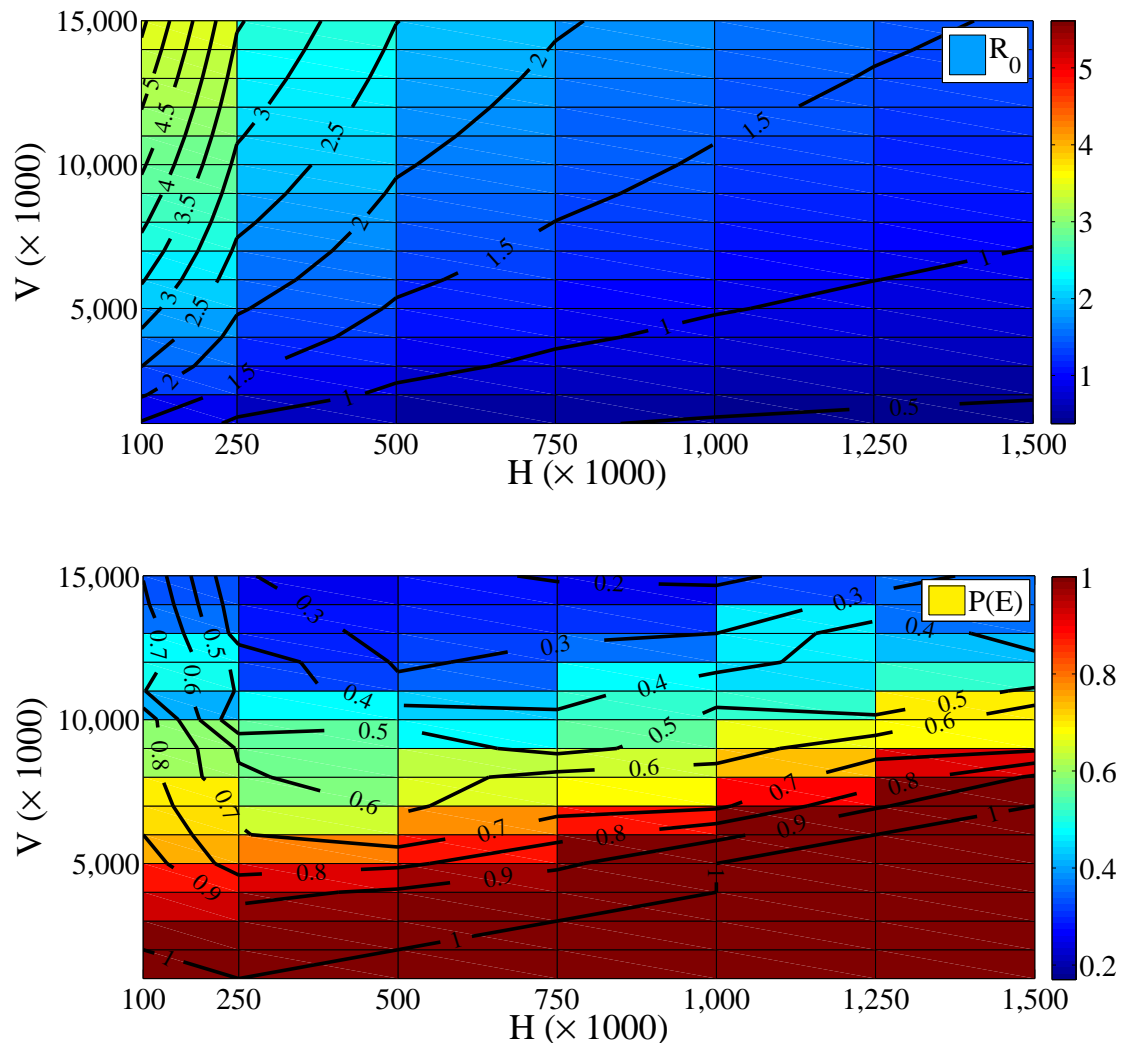
Figure 3.2: Heat maps showing the R_0 and $P(E)$ of dengue for different host population sizes in RM_{SEIR}. The detail and description are same as in Figure 3.1. The additional parameters are $\sigma = \frac{1}{5}$ and $\rho = \frac{1}{8}$.

fixed at nearly eight million vectors for one million to one and half million hosts.

A different pattern of attaining CCS is observed in Figure 3.2. The vector population required to attain CCS is initially fifteen million vectors for one hundred thousand hosts, then decreases to twelve and nine million for 250,000 and half million hosts respectively. The reason for initial decrease in number of vectors is explained in the previous paragraph. After half million hosts there are enough infectious and infected agents to sustain the disease in the population and attaining CCS needs more vectors as the host population increases. It is clear that the higher value of H decreases persistence since H is inversely related to the R_0. In comparison to RM_{SIR}, the inclusion of infectious class E_v supported the dengue persistence in RM_{SEIR}. The inclusion of latent classes markedly changes the persistence pattern of the disease as the

disease stayed for longer in both populations.

3.6.2 Results of sensitivity analysis
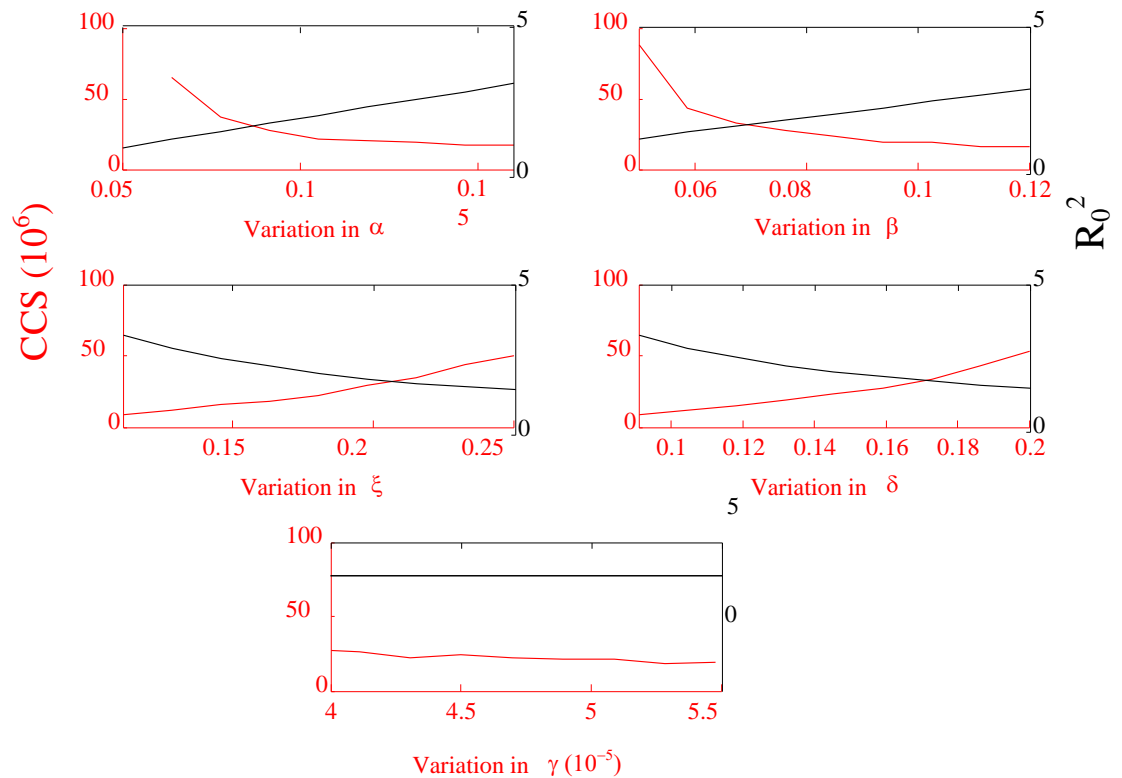
Figure 3.3: Monotonicity plots showing the relation of all parameters with CCS in RM_{SIR}. The values of CCS shown at $y - axes$ of all plots are in millions of hosts. These plots were made by setting all the parameter values at the centre of the hypercube. Each figure is made by varying a parameter from its minimum to maximum value in nine intervals, a change of 25% in the parameter value for one interval. The centre of the hypercube is near to the base-values shown in Table 3.4. The value R_0 at the right-hand y-axis is the squared value for the basic reproductive number.

The minimum, base line and the maximum values of all parameters are given in Table 3.4 and the monotonicity plots for parameters in both models are shown in Figures 3.3 and 3.4. These plots unveiled the relation between CCS and individual parameters. The average transmission rates (α and β) are inversely related to CCS, showing that increased transmission lowers the persistence threshold, provided that there is no susceptible bottleneck.

65
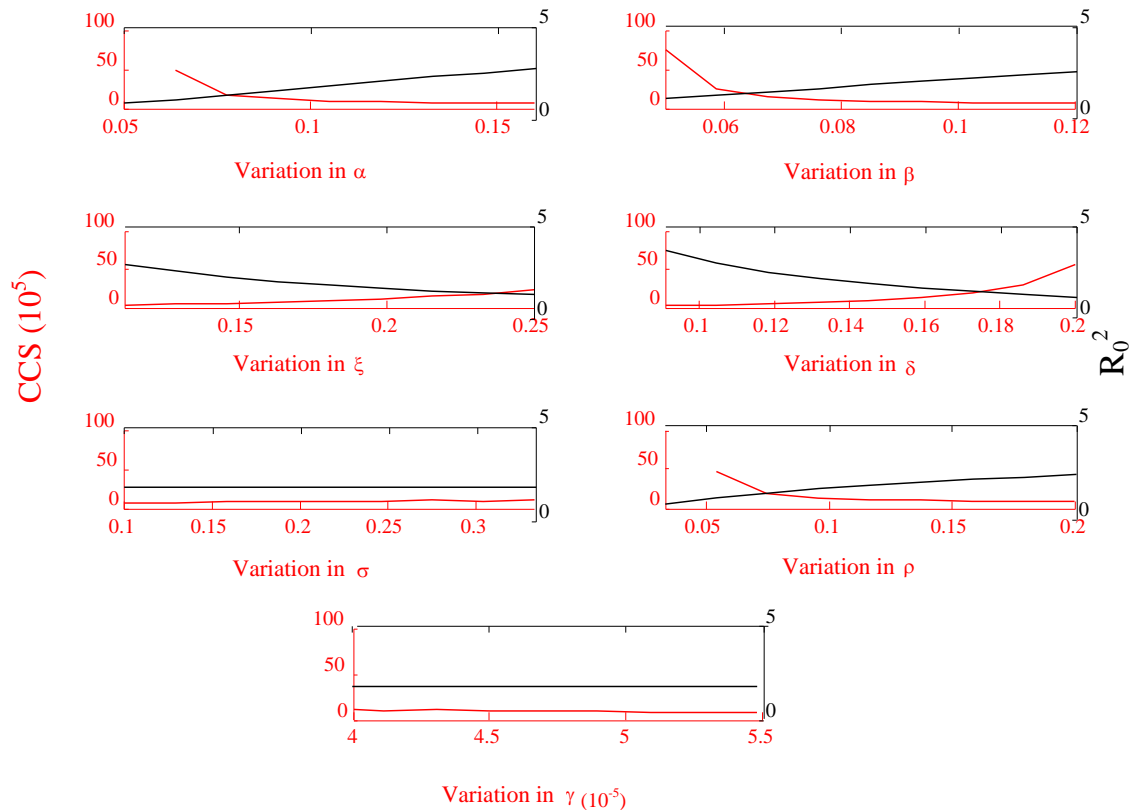
Figure 3.4: Monotonicity plots showing the relation of all parameters with CCS in RM_{SEIR}. The values of CCS shown at $y - axes$ of all plots are in hundred thousands hosts. The rate of exposure in host population IIP, σ have very little effect over CCS whereas the effect of EIP, ρ is obvious. The rest of the description is same as in Figure 3.3.

On the other hand, the average clearance rate of hosts (ξ) and average birth / death rate of vectors (δ) increased the CCS so if hosts are recovering quickly this will decrease persistence. The change in basic reproductive number (R_0) with respect to the parameters is also shown in the monotonicity plots. In both figures, the birth death rate of humans has least effect on CCS and R_0 and in Figure 3.4, the incubation rate of hosts (σ) has very little effect over the CCS and R_0. In the same figure, the incubation rate of vectors (ρ) has a significant effect on CCS. The first and last plot which shows the variation of CCS with respect to α and ρ do not have CCS at the minimum value as R_0 was less than 1 at α_{min} and ρ_{min}. The CCS either increases into unrealistically large values or R_0 falls below 1 in monotonicity plots at the lowest values of transmission rates. The same is also observed for very large duration of incubation period in vectors.

The relationship between CCS and R_0 with change in parameter values is further explored in Figure 3.5. This figure is divided into two columns. The first column shows the relationship of

CCS and R_0 with the parameters of RM_{SIR} whereas the second column is for the parameters of RM_{SEIR}. Similar to Figures 3.3 and 3.4, the values R_0 at the x-axis are the squared values for the basic reproductive number. The effect of γ and σ over R_0 is very low in both models, therefore they are plotted separately at the bottom. Therefore, the plots in the first row denotes the parameters having the most effect. It is interesting to note that CCS behaves almost asymptotically at the extreme values of R_0 in all cases. The explanation of this behaviour can be easily understood at both cases. If the variability in any parameter results value of R_0 near one, then larger population size is needed for the persistence of the disease in both models. In other words, the ratio $\frac{V}{H}$ needs to be bigger. On the other hand, values of CCS obtained by using the parameter combination that resulted large enough R_0 are relatively less changed. Therefore, except for ξ and δ in RM_{SIR}, the values of R_0 greater than $\sqrt{3}$ produces a small change in CCS value that is constant at nearly 20×10^6 hosts. Even for ξ and δ, the higher values of R_0 reveals the same pattern at lower value of CCS. This pattern is more pronounced in parameters of RM_{SEIR} where CCS is less than 1×10^6 individuals for all parameters at values of R_0 greater than $\sqrt{3.5}$.

3.6.3 Linear models of CCS using primary and secondary parameters

In the next section the results of the models based on primary and secondary parameter sets is presented. The model selection criteria in the primary model is based upon the following factors (i) the Akaike information criterion with correction for finite sample size (AICc) and (ii) the proportion of variation explained by the model, determined by the R^2 value. In the models developed using the secondary and combination of both parameters, the second criteria along with $F - ratio$ is used. Where possible, the linear models using the secondary predictors are chosen on the basis of the availability of the data related to the parameters. The main idea behind this approach is to identify models containing explanatory variables that can be easily measured or estimated in the real-world. Therefore, the best model also has the advantage that the data of its explanatory variables is easier to gather from the field.

3.6.3.1 Predicting CCS on the basis of primary parameters

Among the primary variables, it was found that birth and death rate of hosts have least effect on the dynamics of CCS, so they are not included in any of the selected models as a main effect. As the primary variables were all independently sampled from a uniform distribution, they have very low correlation values among them. In all of the models, the average clearance rate of host ξ and birth / death and infectious rate of vector, δ explain
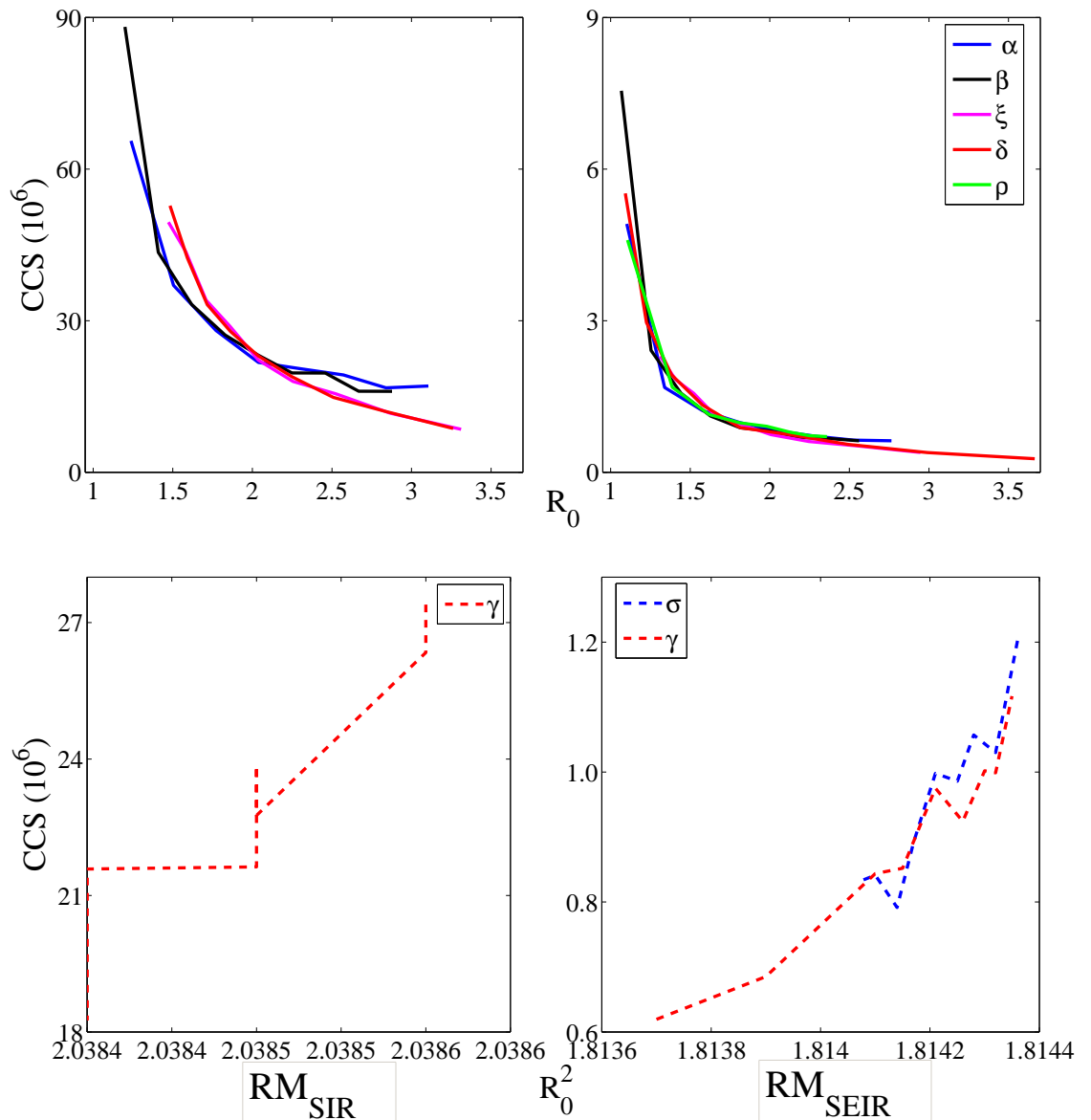
Figure 3.5: The relationship between CCS and R_0 explored using the parameters of the models. The left columns shows this relation for the parameters of RM_{SIR} whereas the right column is for RM_{SEIR}. The parameters having the most effect on the CCS and R_0 are mentioned in the top row of the figure. The bottom row is dedicated for the parameters having the least effect on CCS and R_0. The rest of the description is same as in Figures 3.3 and 3.4.

most of the variation in CCS (see Figure 3.6). The values which correspond to lower disease transmission rates α and β, faster clearance ξ and death rate of vectors δ, required larger populations to maintain the disease. The percentage variance explained was 60% with AICc value of 3304 and 57 % with AICc value of 3049 for RM_{SIR} and RM_{SEIR} respectively. As expected, both models performed badly at predicting extreme values of CCS, especially at the low values where model predicted negative values of CCS. The residuals from both models

Parameter	Minimum	Baseline	Maximum	RM_{SIR} PRCC (r)	RM_{SIR} p-value	RM_{SEIR} PRCC (r)	RM_{SEIR} p-value
α	0.05	0.1	0.16	-0.725	$< 10^{-10}$	-0.773	$< 10^{-10}$
β	0.05	0.075	0.12	-0.704	$< 10^{-10}$	-0.724	$< 10^{-10}$
ξ	0.11	0.1428	0.25	0.862	$< 10^{-10}$	0.841	$< 10^{-10}$
δ	0.09	0.125	0.20	0.845	$< 10^{-10}$	0.897	$< 10^{-10}$
γ	3.91×10^{-5}	4.215×10^{-5}	5.47×10^{-5}	-0.159	0.1241	-0.3865	0.00013
σ	0.10	0.2	0.33	.	.	0.314	0.0022
ρ	0.03	0.125	0.20	.	.	-0.699	$< 10^{-10}$

Table 3.4: Values of the parameters used in Latin Hypercube Sampling for RM_{SIR} and RM_{SEIR} models. Baseline parameter values are presented in third column. A total of a hundred parameter samples were generated. PRCC results of each parameter along with its p-value are shown. The last two parameters, σ and ρ were used only in RM_{SEIR}.

showed fairly normal behaviour by applying standard diagnostic tests. The details of the models for CCS using least squares estimation are given in Table 3.5:

Variable	RM_{SIR} Value	RM_{SEIR} Value
Intercept	-9.528×10^6	8.774×10^5
α	-3.826×10^7	-1.488×10^7
β	-6.322×10^7	-2.491×10^7
ξ	6.816×10^7	1.696×10^7
δ	7.782×10^7	2.312×10^7
σ	-	4.118×10^6
ρ	-	-7.755×10^6
γ	-	-3.898×10^{10}

Table 3.5: The values of the coefficients of the linear models for RM_{SIR} and RM_{SEIR} using primary parameters.

3.6.3.2 Predicting CCS on the basis of Secondary parameters

This section provides an overview of the determinants of CCS by using secondary predictors. As discussed previously, those parameters arose as a result of algebraic manipulation of primary predictors. The main secondary predictors which were tested as the determinants of CCS were R_0, $N = \frac{R_0}{R_0 - 1}$ and R_h^*. The next subsections presents the linear models constructed on the basis of the above secondary parameters.
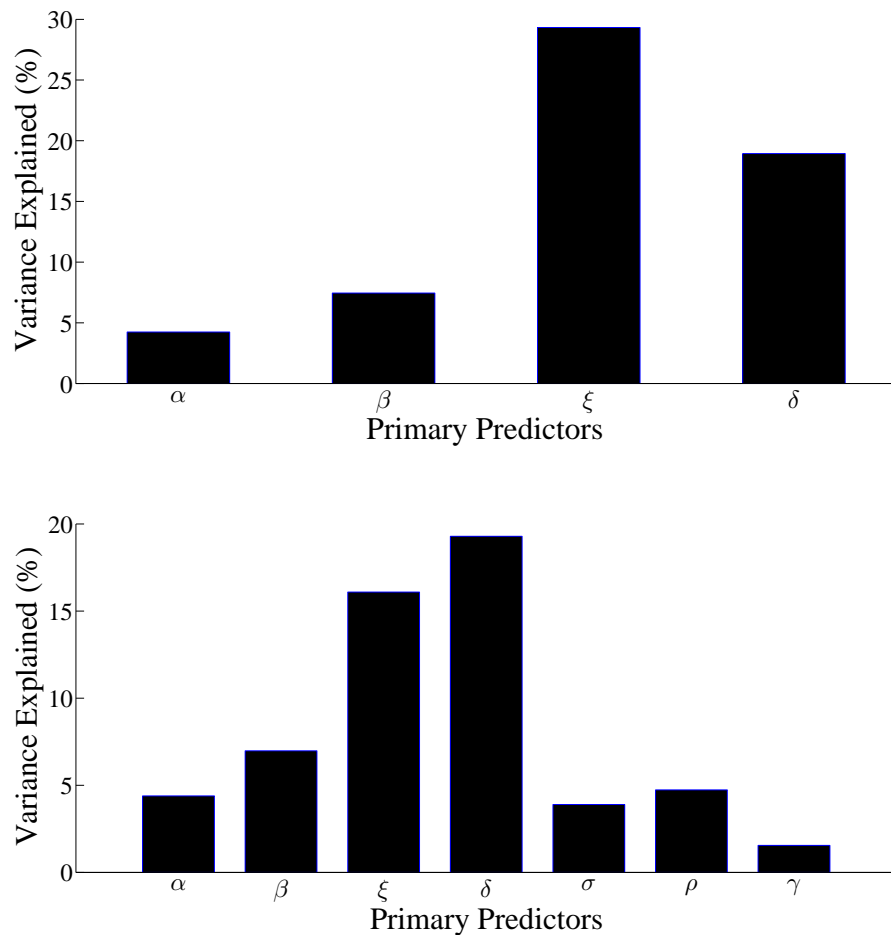
Figure 3.6: Determinants of CCS over twenty-five years using primary variables. **Top**: baseline model RM_{SIR} and bottom: Model with latent periods. The key determinants were found to be the birth / death rate (δ) of vectors and the recovery rate of hosts (ξ). As the vector remains sick for the rest of its life, (δ) accounts for the average infectious period of the insect as well. Transmission rates (α & β) and the average incubation rates (σ & ρ) have relatively lower impact on persistence. In the top figure, the model explains 60% of the variance in CCS and 57 % variance is explained in the bottom plot.

3.6.3.3 Relation between CCS and R_0

As defined in Chapter 2 Sections 2.2.1 and 2.3, the basic reproductive ratio R_0 is an important epidemiological parameter, which represents the numbers of susceptible cases that arise in a completely susceptible population caused by a single infected individual over the duration of its infectiousness. In this manner, R_0 gives insight about the next generation of infection propagation during the early stages of an epidemic. In the deterministic setting, intervention and control of a disease typically aims to reduce R_0 (or strictly speaking, effective R_0 that is called R_{eff} and $R_{eff} = R_0 \times S_h$) less than one causing the endemic equilibrium to become unstable and resulting in the eradication of the pathogen from the population. In heterogeneous populations, multiple reproductive ratios exist, one for the host and one for the vector. In Figure 3.7, the association between CCS and R_0 is shown. The sampling data is arranged from high R_0 values to low values. The plot shows an important finding that CCS behaves asymptotically with R_0 which is in agreement to the pattern seen in Figure 3.5. Provided there are enough susceptible individuals for disease to flourish, high values of R_0, above $R_0 = 5$ did not cause a considerable change in CCS. Similarly, the values slightly above $R_0 = 1$ required unrealistically large population sizes for disease to persist. The inset figures shows plots of R_0^{HV} and R_0^{VH}, where the former included the term $\frac{V}{H}$ in the numerator which scales it up as compared to the latter reproductive ratio. The inset also compares the reproductive ratios in both models used in the simulation experiments. The inset plots are made by arranging the sampling data of R_0, R_0^{HV} and R_0^{VH} from high to low values. These plots make sure that the parameter spaces generated from the LHS experiment, have a similar range of basic reproductive ratios for better comparison in both models.

3.6.3.4 Predicting CCS on the basis of $N = \frac{R_0}{R_0-1}$

Nåsell (2005) provides a foundation for deriving critical community size for childhood infections. One of the measures of CCS derived was the population size for which the mean time to extinction from quasi-stationarity is equal to $\frac{K}{\gamma}$ years, where K is a dimensionless scaling parameter and γ is the birth and death rate of hosts. Then he used the relation between expected time to extinction, starting from the quasi-stationary distribution, and population size at which the pathogen become extinct to derive the following equation: $N_{crit} = \eta^2 K \frac{R_0}{R_0-1}; R_0 > 1$ where N_{crit} is the CCS as defined above and η is the ratio between average lifespan and the average duration of infection.

This relation suggests that CCS is a function of R_0 for childhood diseases. It is worth investigating whether this formula holds for CCS in host vector systems. In the current
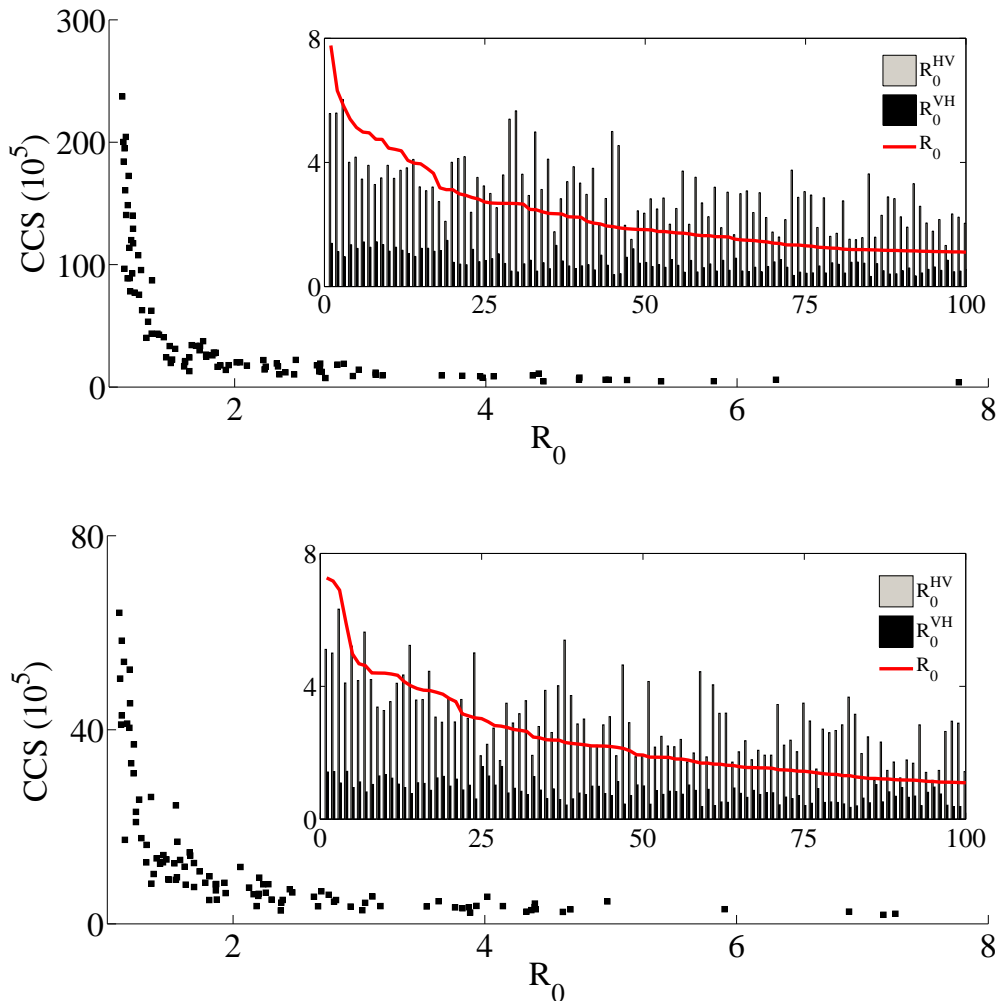
Figure 3.7: Relation between the basic reproductive ratio R_0 and CCS in both models. The top figure represents the relationship for RM_{SIR} and bottom figure denotes the relationship for RM_{SEIR}. In both models, the basic reproductive number has an asymptotic relation with CCS, showing very little change in CCS population size at high values of R_0. Insets in both figures show the relation between R_0, R_0^{HV} and R_0^{VH}, ranked from highest value of R_0 to the lowest.

work, the ratio $\frac{R_0}{R_0-1}$ is denoted by the variable N. The difference between Nasell's work and the current work is that Nasell started the simulations from the quasi-stationary distribution whereas deterministic equilibrium points are used as the starting condition in the current work. It is assumed that the quasi-stationary state is quickly achieved during the simulations.

For both RM_{SIR} and RM_{SEIR}, the variable N predicts CCS well. For RM_{SIR}, the linear model is found as CCS $= 2 \times 10^6 N - 2.25 \times 10^6$, and 90 % variance is explained by N whereas CCS $= 5.6 \times 10^5 N - 3.8 \times 10^5$ and 87% variance is explained by N for RM_{SEIR}. The intercept in both models were negative, which were not of interest as the basic reproductive number was greater than 1 in all the simulations and the limiting behaviour of N resulted in $N \to 1$.
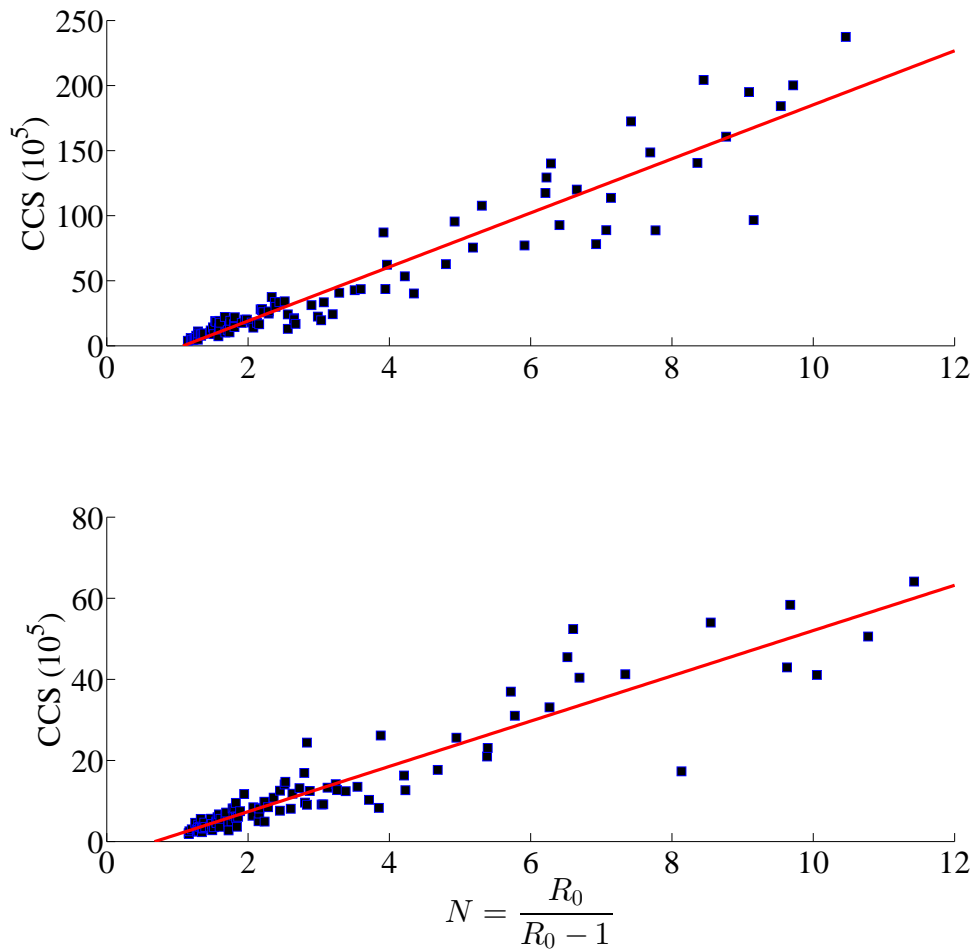
Figure 3.8: The relation between N and CCS in RM_{SIR} (top) and RM_{SEIR} (bottom). N explained 90 % and 87.6 % variation for CCS respectively. Both plots show evidence that CCS can be approximated by using the basic reproductive ratio.

The plots of N and CCS for both models are shown in Figure 3.8.

3.6.3.5   Predicting CCS as a combination of primary and secondary parameters

In this section, the prediction of CCS is attempted by using a combination of both primary and secondary predictors. In order to make prediction better, several parameter combinations were attempted on the basis of criteria defined in section 3.6.3. The goal was to obtain a set of predictors measurable in the field and which explains the maximum variance in modelling CCS. The main models considered are shown in Table 3.6.

In Table 3.6, the initial seroprevalence proportion of hosts R_h^*, at the deterministic endemic equilibrium is an important predictor for both models. For RM_{SIR}, including transmission

73

Model	Parameters	R^2	$Pr(>F)$
RM_{SIR}	(i) λ, α, β	60.00	$< 10^{-16}$, 9×10^{-5}, 4.7×10^{-4}
	(ii) R^*	63.10	$< 10^{-16}$
	(iii) R^*, R_0^{VH}, β	70.14	$< 10^{-16}$, $< 10^{-16}$, 2.6×10^{-5}
	(iv) R^*, R_0^{HV}, α	73.00	$< 10^{-16}$, $< 10^{-16}$, 1.3×10^{-5}
RM_{SEIR}	(i) R^*	62.78	$< 10^{-16}$
	(ii) $R^*, \alpha, \sigma, \rho$	72.14	$< 10^{-16}$, 2.8×10^{-3}, 6.3×10^{-4}, 2.0×10^{-3}
	(iii) N, σ, ξ	91.70	$< 10^{-16}$, 4.4×10^{-6}, 5.8×10^{-6}

Table 3.6: Models to predict CCS on the basis of secondary predictors. Models are presented in ascending order in the context of better fit to the data. The last column shows the significance of each predictor.

parameters also improved the fit of the model. The average incubation period and average infectious duration of host were important predictors for persistence.

3.7 Conclusion and discussion

The current study uses a new combination of techniques to construct modelling frameworks for (i) estimating persistence thresholds and (ii) identifying the determinants of this persistence threshold in the host-vector framework. These two goals were met by using two epidemiological models that were parametrized for dengue and investigated for long term persistence using a stochastic framework. The concept of CCS is used to relate the host population size with the extinction probability of pathogen. For achieving (i), stochastic simulations were started at deterministic endemic state of the modelling systems, hereby replacing the quasi-stationary distribution that is an analogue of the steady state in stochastic settings. It was assumed that the simulations fall quickly into the quasi-stationary state. Host population size, and corresponding vector population sizes were increased until 50% of the simulations retained infection after twenty-five years. CCS for RM_{SIR} and RM_{SEIR} was investigated for different $H - V$ pairs. To identify the determinants of CCS, i.e., goal (ii), parameter sensitivity and determinants of CCS were explored by constructing a LHS design and fitting linear models with primary and secondary parameters a independent variables. The main determinants of CCS were R_0, N, δ, ξ, α, β and σ (description of parameters is in Table 3.1).

This study focuses on long term persistence dynamics, which requires a sufficient number of susceptible hosts at the start of the simulations. Under the current parameter settings,

this is achieved by setting low values of bite rate and probability of transmission. The time scale of demographic turn-over of hosts is very slow as compared to vectors and increased transmission probability leads to a very high seroprevalence (∼90% recovered hosts). This resulted in rapid extinction of disease in the host population and consequent extinction in the vector population. Linear modelling using primary predictors shows that the low values of average transmission rates α and β have a smaller effect on long term persistence then average clearance rate of hosts, ξ and average birth / death rate in vectors, δ (see Figure 3.6). Transovarial routes of transmission in vectors are not considered here, mainly because of low evidence for the efficacy of pathogen transmission by these routes or contrast between the parameter values obtained between the field studies and laboratory measures (Adams and Boots, 2010). Moreover, the significance of vertical transmission in vectors for disease maintenance during inter-epidemic periods is not well understood (WHO, 2009).

In this study, the CCS is defined as the minimum host population size for which half of the stochastic simulations still retain either infected hosts or vectors after twenty-five years. In the literature, there is no exact value of the time required for the population to attain CCS. In our case, it is taken to be twenty-five years as the dynamics and ecology of both populations may change a lot at longer time periods. For example, the assumption of constant host population fails to hold. As it is mentioned that the simulations fall quickly into the quasi-stationary state (results not reported in this study), the time-limit of twenty-five year seems a sensible trade-off between the time required to attain the quasi-stationary state of the system and the computational time required for the experiment to finish.

The results of PRCC are shown in Table 3.4. The correlation values agrees with the behaviour seen in the monotonicity plots and gave numerical values of the relationship between CCS and parameters of the models. In both models, the birth and death rate of hosts (γ) has a very little affect on CCS as the length of the simulations time was $\sim 40\%$ of $\frac{1}{\gamma}$. In Figure 3.4, the human incubation rate σ has very low effect. The low correlation value between CCS and σ in Table 3.4 suggests that it is not efficient to include exposed class in hosts. This helps decreasing the complexity of the model representing the host-vector system. The influence of the latent period in the vector population ρ, is shown in 3.4, making it a significant addition in RM_{SEIR}. Based on the above discussion, a model having a exposed class in vectors is preferable in modelling long-term pathogen persistence than RM_{SIR} model.

The equations of infectious individuals in equation sets 2.1 and 2.2 for RM_{SIR} represent the number of new infections created per unit time. If, at any time, these rates of change become negative, then the infection will begin to decrease in both populations. In the simulations, the infectious populations shows strong positive correlation in hosts and vectors: decline in

one leads to reduced numbers in other. In order to make the rate of change of infectious hosts and vectors negative, the following inequalities must hold:

$$I_h < \frac{\alpha I_v}{\xi + \gamma} \left[\frac{S_h}{H} \right]$$
$$I_v < \frac{\beta S_v}{\delta} \left[\frac{I_h}{H} \right]$$

$$(3.1)$$

Investigating $\frac{\mathrm{d}I_h}{\mathrm{d}t}$ in 2.1 and $\frac{\mathrm{d}I_v}{\mathrm{d}t}$ in 2.2 reveals that this can be obtained as follows:

- By reducing α and β to create a transmission bottleneck.

- By reducing S_h and S_v in the populations.

- By increasing the death rate δ of infectious vectors.

- By increasing the infectious clearance rate ξ in hosts.

Reducing the rates σ and ρ in RM_{SEIR} increases the time it take for a susceptible individual to become infectious. Including the incubation period resulted major decrease in the number of hosts required for the persistence of pathogens, but has relatively low effect on CCS from the perspective of the vector population. So far there is either no vaccine developed for dengue or the vaccine efficiency is not confirmed (Mahalingam et al., 2013). In this context, minimizing the likelihood of contact between susceptible and infectious populations is one of the plausible control strategies. In infectious diseases, shortening the infection clearance rate of hosts ξ corresponds to administrating anti-viral drugs and reducing the virus load in the blood. The next three sections consider each of these options in a realistic manner in the context of vector control efforts.

As described earlier, the terms α and β consist of the bite rate of the vector times the probability of transmission. Due to the assumption of non-seasonality, the bite rate is constant through time. The easiest way to reduce α and β is to reduce the the probability of transmission from the bite of the vector. This includes vaccinating the susceptible proportion of hosts, rapid treatment and quarantining of infected hosts around 5-7 days (e.g. by encouraging them to work from home thus restricting their movement, so they cannot get bitten). As both α and β appear in the numerator of the population level basic reproductive ratios, R_0^{VH} and R_0^{HV} respectively, they can be used to reduce the values of these reproductive numbers below unity. Nonetheless, R_0^{HV} is easier to minimize by quarantining infectious hosts.

So far, eradication and control efforts of many mosquito-borne infectious diseases are centred upon the reduction in the mosquito population. Vector control programs aim to reduce the

number of older mosquitoes, as modelling results in Brownstein et al. (2003) showed that more than half of the population of thirty days older *Ae. aegypti* are capable of transmitting dengue, although the population proportion in this age class is very small. In McMeniman et al. (2009), experiments suggested the introduction of *Wolbachia* reduces the life span of *Ae. aegypti* by 50%, making it a viable vector control strategy. Targeting the older population does not only target the infectious proportion of mosquitoes, but also puts a lower selection pressure on the population. In equations for the basic reproductive ratios of RM_{SIR} and RM_{SEIR}, i.e., equations 2.3 and 4.2, killing older mosquitoes will result in the reduction of R_0^{VH}, as the vector birth / death rate appears in the denominator of the expression. This means that the vector population has more young mosquitoes but they die early, interrupting the transmission cycle.

The effect of varying δ is obvious in monotonicity plots shown in Figures 3.3 and 3.4 and it has one of the highest correlations with CCS for both models in Table 3.4. In the context of dengue, controlling *Ae. aegypti* and *Ae. albopictus* populations also reduces the outbreak risks of urban yellow fever and chikungunya disease (WHO, 2014). In RM_{SEIR} the chances for an exposed mosquito to become infectious are $\frac{\rho}{(\rho+\delta)} = 50\%$, therefore the numerical value of R_0^{VH} in RM_{SEIR} is 0.5 of its value in RM_{SIR}. In comparison, the chances of an exposed host to become infectious are $\frac{\sigma}{(\sigma+\gamma)} \approx 100\%$, meaning that the value of R_0^{HV} is roughly the same in both models. Since both of these incubation periods are exponentially distributed and the median of this distribution is less than the mean, most of the hosts and vectors have incubation periods less than five and eight days respectively. The role played by these parameters in varying CCS is evident in the monotonicity plots shown in Figure 3.4 and in Table 3.4. The potential effect and contribution of vector control is discussed in Townson et al. (2005). Lambrechts et al. (2009) highlighted the importance of diversification of research priorities in multiple disciplines for vector control strategies. The authors suggested that the focus of vector biology research to be shifted for the identification and characterization of heterogeneities related to the real-world pathogen-transmission system.

Although the current chapter has highlighted both qualitative and quantitative relationships among the drivers of persistence and critical population size, these results are valid for diseases which are not affected by seasonal variations. It would be interesting to see the changing patterns of persistence dynamics in conjunction with seasonal forcing applied to the parameters. The next chapter will investigate the determinants of CCS with seasonal time varying parameters. The aim will be to explore how current predictors, taken in a seasonal context alter the persistence of pathogen in both populations.

CHAPTER 4

Modelling persistence using Ross Macdonald dengue model with seasonality

Modelling persistence using Ross Macdonald dengue model with seasonality

4.1   Introduction

Seasonality has major effects on the spread and persistence of dengue. The number of reported cases of the dengue virus disease are clearly seasonality driven Bartley et al. (2002), mainly because of the abundance of female mosquitoes in the rainy season. In the wet season, when there is optimum temperature and rainfall, the mosquito population grows while the adult population is often negligibly small in unfavourable season (dry). Therefore, more dengue cases are reported in wet seasons as compared to dry seasons. Focks et al. (2000) have presented the effect of temperature on dengue transmission and shown that the ideal temperature for the transmission of dengue virus falls between $20\,°C$ and $35\,°C$. Temperatures $>35\,°C$ for longer periods eliminate the possibility of adult population existence as aggregate survival of eggs, larval and pupal stages of mosquito is insufficient. In particular, temperature and rainfall are the main drivers of mosquito population size in cities with a temperate climate as they usually have the optimum climatic conditions required for population maintenance.

Nowadays, dengue has spread to different parts of the world. In countries with a temperate climate and localities which are situated in the monsoon region, there is a potential risk of dengue virus transmission during the rainy season. Human mobility has been shown to be a major driver of dengue spread and re-introduction (Adams and Kapan, 2009). In areas where dengue is endemic causing periodic outbreaks, estimating the probability of disease transmission to a neighbouring naive population is an immediate need to stop the spread of the disease. Therefore, development of a mathematical framework that predicts the probability of an outbreak after the introduction of an infectious individual is important. Once dengue has been introduced to a population and has gone extinct, immediate re-introduction should be less likely because of existing seroprevalence reducing the density of susceptible humans. An estimation of the critical susceptible density required for recurrent epidemics can lead to better policy planning and disease control.

A range of models studying the dynamics of dengue in humans and vectors exist in the literature. In the early 1990s, Focks and colleagues developed detailed seasonal models to investigate the vector population dynamics and epidemiology of dengue viruses (Focks et al., 1993a,b). Their entomological model (CIMSiM) provided the necessary input parameters for their corresponding dengue simulation model (DENSiM) (Focks et al., 1995), which modelled the demographic, entomologic, serologic, and infection data for the human population in an urban environment. Both stochastic models were used to estimate transmission thresholds in terms of pupae per person (Focks et al., 2000), intended to assist in efforts to reduce sources for mosquito breeding. Both CIMSiM and DENSiM models are site-specific and very expensive because they require localized surveys for gathering human, vector and weather data. In resource poor countries, these long-term studies are even harder as localized data are usually not available.

Bartley et al. (2002) constructed a two-serotype model of dengue that included antibody dependent enhancement (ADE) following multiple infection. They introduced seasonally-varying parameters in a step-wise fashion to their two-serotype models and found a strong impact of seasonal forcing on the prevalence of dengue disease. The results were then compared to the seasonal pattern of dengue infection in Thailand. The duration of infectiousness of the host, vector mortality, and biting rate were found to be the most influential parameters. Extending their work, Wearing and Rohani (2006) developed a model considering all four serotypes of dengue. Periodic recruitment of vectors and different virulence for different serotypes were also incorporated in their model. The authors suggest that seasonal variation in vector demography combined with a short-lived period of cross-immunity is sufficient to generate infection time series that correspond with the dengue infection data in Thailand. A four-serotype model of dengue was considered by Chikaki and Ishikawa (2009) where seasonal variation is included in (i) population dynamics of *Aedes aegypti*, (ii) inapparent cases that influence disease prevalence, and (iii) the influence of antibody dependent enhancement (ADE) to model the realistic behaviour of annual cycles of dengue disease. They found that the immunity acquired by infection during a cross-immunity period, termed 'unnatural routes' of infection, changes the intensity and timing of dengue epidemics.

All of these multi-serotype models are constructed using complicated compartmental structures and require detailed information for parametrisation. The work done by Focks et al. (1995, 1993a,b) was mainly directed towards a detailed study of vector population dynamics and dengue spread in human population. They extend their work towards vector control in Focks et al. (2000) and estimate transmission thresholds in terms of pupae per person. In Bartley et al. (2002), the two-serotype model was found insensitive to the degree of cross protection with second serotype. In contrast, Chikaki and Ishikawa (2009); Wearing and Rohani

(2006) using a four serotype model found that the cross-immunity played an important role in determining the intensity and timing of the epidemics. Although dengue modelling with more than one serotype provide valuable information about the current and future trend of disease, they are difficult to use for the development of a continuous surveillance system and early detection of dengue, particularly in developing countries.

Using a fully stochastic dynamical model, Otero et al. (2006) considered the seasonal change in all life stages of an *Aedes aegypti* population in a homogeneous environment. In Otero et al. (2008), they considered the spatial population dynamics for Buenos Aires. The population dynamics of vectors is influenced by the interaction between the patches and the availability of breeding sites. Their models were updated in Otero and Solari (2010), where they showed the dependence of the epidemic size on the arrival of viremic people in Buenos Aires at different times of the year. In all of these models, seasonal variation in the vector population is controlled by introducing two non-linear regulatory processes that prevent an explosion of the vector population: (i) density dependent mortality of larvae; (ii) egg-hatching inhibition by larvae. They further improved their work by including networks to represent human mobility which is found to be the main driving force for the spread of dengue (Barmak et al., 2011). Their work can be summarized as follows: after developing a detailed map of the areas of potential distribution of *Aedes aegypti* in Buenos Aires (Otero et al., 2008, 2006), they estimated the risk of dengue outbreaks in Buenos Aires (Barmak et al., 2011; Otero and Solari, 2010).

In this chapter, the main goal is the development of a minimalist mathematical model that can be used to investigate the time evolution of a single serotype of dengue dynamics in a seasonal environment. In essence, the focus is to construct a mathematical tool for a wider range of communities that is powerful enough to answer questions like those asked in Otero and Solari (2010) but using a simpler structure. The model formation uses the classical homogeneous approach in contrast to adding individual heterogeneity in the system. The model has a seasonally dependent birth rate $\delta_b(t)$ for vectors, considered to be mediated by temperature, rainfall and humidity. In the absence of primary data, the model is parametrized using values from the available literature. The overarching research aim is to investigate the impact of seasonality on the time evolution of dengue and whether the time evolution depends on the introduction of infection at different times of the year. I also use the model to ask the following questions: how the risk of an outbreak changes if the disease is introduced in a population having pre-existing seroprevalence in humans; what are the differences in patterns of outbreak if triggered by a viremic human or by a viremic vector; how does dengue persistence relate to the month of introduction; and how does the time to extinction of dengue disease relate to the peak and timings of the epidemics. Questions like these are crucial for

the elimination and control of the disease.

The following section provides a detailed introduction to the model. The material and methods are presented in sections 4.3 and 4.4. Section 4.5 discusses the results from the deterministic version of the model while results from the stochastic version are shown in section4.6. The last section is dedicated to the discussion of the results and the limitations and caveats of the current study.

4.2 Ross-Macdonald model with seasonality, RM_{SEIR}^{s}

In this section the overview of the mathematical model RM_{SEIR}^{s} is given. Both deterministic and stochastic versions of the model are used in this chapter. The introduction and the description of symbols of the model are presented here.

Over the past seventy years, a family of different Ross-Macdonald style models have been used to simulate and predict vector-borne diseases (Smith et al., 2012). The model presented here RM_{SEIR}^{s} is another modified version of the Ross-Macdonald framework. It is constructed using a standard compartmental scheme whose description is as follows: the total number of individuals is divided into hosts (humans) and vectors (mosquitoes). The host population (H) is represented by a Susceptible-Exposed-Infected-Recovered (S E I R) compartmental model where $H = S_h + E_h + I_h + R_h$. The vector population V is represented by three compartments, a Susceptible-Exposed-Infected (S E I) model (vectors never recover from the infection) and $V = S_v + E_v + I_v$. The host-vector system is described by seven ordinary differential equations in equation set 4.1 whereas the schematic diagram is presented in Figure 4.1.

The transmission rates are defined as α and β, where α denotes the transmission rate (per day and per vector) from an infected vector to a susceptible host and β denotes the transmission rate (per day and per vector) from an infected host to a susceptible vector. The population-level rate of bites that generate new dengue infection (i.e, when the infected vector bites a susceptible member of the host population) is given by $\alpha I_v \left(\dfrac{S_h}{H} \right)$. The parameter ξ represents the host recovery rate (per day). The virus incubation rate in hosts (per day) is denoted by σ. The death rate in all host classes is γ and the birth rate Λ is equal to the death rate. There is no effect of seasonality on the parameters related to hosts. The population of female adult vectors ingest the dengue virus during a blood meal from the infected proportion of hosts at rate $\beta S_v \left(\dfrac{I_h}{H} \right)$. The per-capita birth rate of mosquitoes (per day), $\delta_b(t)$ is treated as time-dependent and varies following the wet and the dry season. The death rate of vectors δ_d and the rate of incubation ρ are kept fixed, independent of the variations in the season.

The numerical values of the above parameters chosen, with a brief explanation about the parameters related to the vector population, are provided in section 4.4.3.
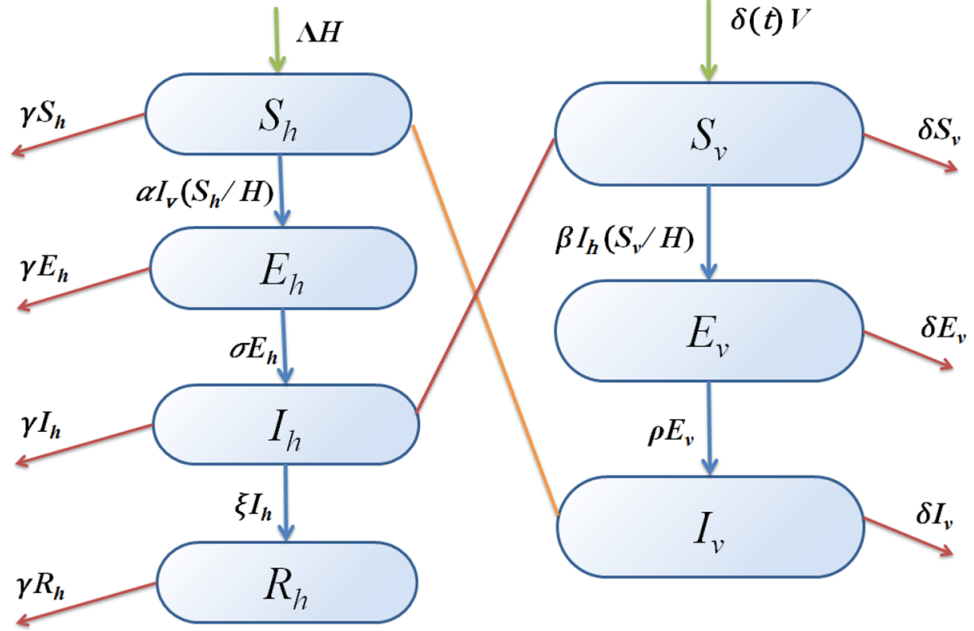
Figure 4.1: Schematic Diagram of RM^s_{SEIR}.

The modelling process is defined by the following seven non-linear time-varying state equations:

$$
\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= \Lambda H - \alpha I_v\left(\frac{S_h}{H}\right) - \gamma S_h \\
\frac{\mathrm{d}E_h}{\mathrm{d}t} &= \alpha I_v\left(\frac{S_h}{H}\right) - (\sigma + \gamma)E_h \\
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= \sigma E_h - (\xi + \gamma)I_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \xi I_h - \gamma R_h \\
\frac{\mathrm{d}S_v}{\mathrm{d}t} &= \delta_b(t)\, V - \beta S_v\left(\frac{I_h}{H}\right) - \delta_d\, S_v \\
\frac{\mathrm{d}E_v}{\mathrm{d}t} &= \beta S_v\left(\frac{I_h}{H}\right) - (\rho + \delta_d)\, E_v \\
\frac{\mathrm{d}I_v}{\mathrm{d}t} &= \rho E_v - \delta_d\, I_v
\end{aligned}
\tag{4.1}
$$

Note that the rate of change of the vector population size, which is implicit in above equations, is:

$$
\frac{\mathrm{d}V}{\mathrm{d}t} = (\delta_b(t) - \delta_d)V
$$

If $\delta_b(t) > \delta_d$, then the vector population will grow at an exponential rate and if $\delta_b(t) < \delta_d$, the population will decrease. If $\delta_b(t) = \delta_d$, then the population remains constant. The non-seasonal version of the model, RM_{SEIR} is proven locally asymptotically stable in Chapter 2, section 2.3.1 at disease-free (when $R_0 \leq 1$) and endemic (when $R_0 > 1$) equilibrium. The stability of the seasonal model RM_{SEIR}^s at endemic equilibrium is discussed in section 4.5.1.

4.3 Methods for analytic derivations

For investigating the effect of introducing disease at different times of the year in RM_{SEIR}^s, the analytic expression of several measures, including the basic reproductive numbers and probabilities of invasion is presented in the next section.

4.3.1 Seasonal reproductive numbers and invasion probabilities

For simple models, the basic reproduction number R_0 is the average number of secondary infections generated from a single infected individual introduced into a susceptible population during its lifetime of infection. In host-vector systems there is more than one basic reproductive number. The non-seasonal model that was discussed in last chapters, RM_{SEIR} comprises of two reproductive numbers denoted by R_0^{VH} and R_0^{HV}. R_0^{VH} is the average number of secondary cases in the susceptible host population, resulting from the introduction of one infected vector. R_0^{HV} is defined by reversing the roles of hosts and vectors in this definition. The value of the basic reproductive ratio R_0 used in previous chapters (Appendix A.1.2) for the non-seasonal model RM_{SEIR} using the next generation method described in Diekmann et al. (2010) is:

$$R_0 = \sqrt{R_0^{VH} \times R_0^{HV}} = \frac{\alpha\rho}{\delta_d(\rho + \delta_d)} \times \frac{\beta\sigma V}{(\xi + \gamma)(\sigma + \gamma)H} = \sqrt{\frac{\alpha\beta\sigma\rho V}{\delta_d(\xi + \gamma)(\sigma + \gamma)(\rho + \delta_d)}} \quad (4.2)$$

Note that the per capita birth rate of vectors, δ_b is not included in the expression for R_0. R_0^{VH} has an inverse relationship with the death rate δ_d of the vectors. The term $\frac{\rho}{\rho + \delta_d}$ is the probability that a vector will survive the exposed state (E_v) and move to the infectious compartment (I_v). In R_0^{VH}, the rate of transmission from an infectious human is β, the quantity $\frac{1}{\xi + \gamma}$ denotes the average amount of time an individual who enters state I_h spends in state I_h and $\frac{\sigma}{\sigma + \gamma}$ is the probability that a host will survive the exposed state (E_h) to enter the infectious state (I_h). Mathematically, the term R_0 is the geometric mean of the number of infected vectors per infected host and the number of infected hosts per infected vector.

The product R_0 acts as a general threshold and there is disease extinction when this number is less than 1 and either invasion or persistence (in deterministic settings) of disease, if this number is greater than 1.

The expression for R_0 in equation 4.2 is suitable for a non-seasonal model. However, for the seasonal models R_0 does not provide much valuable information because the number of vectors, and thus other values, vary over time. Intuitively, the reproductive number of the seasonal model differs as a function of the time of year of introduction; we therefore write $R_{0|t_0}$, where $_{|t_0}$ is read "*given t-zero*", and t_0 denotes the time of year of introduction (i.e. the season, measured in months from April).

Following (Diekmann et al., 1990) and (Heffernan et al., 2005) in Appendix A.1.2, the matrix F reflecting the rate at which new infections arise in the seasonal model is written as

$$
F = \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 \\ 0 & \beta\frac{V}{H} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\tag{4.3}
$$

The derivation of $R_{0|t0}$ assumes the disease-free condition, therefore all individuals are initially susceptible as shown in matrix 4.3.

In the transmission matrix F in 4.4, the terms $\alpha\frac{H}{H}$ and $\beta\frac{V}{H}$ are replaced by $\alpha\frac{S_h}{H}$ and $\beta\frac{S_v}{H}$ respectively since the total population is not all susceptible at the beginning.

$$
F = \begin{pmatrix} 0 & 0 & 0 & \alpha\frac{S_h}{H} \\ 0 & 0 & 0 & 0 \\ 0 & \beta\frac{S_v}{H} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\tag{4.4}
$$

Since the size of the vector population V is changing with respect to the seasons, the fraction $\frac{V}{H}$ is replaced by $\frac{V(t_0)}{H}$. The expression for $R_{0|t_0}$ can be written as follows:

$$
\begin{aligned}
R_{0|t_0} &= \sqrt{R_{0|t_0}^{VH} \times R_{0|t_0}^{HV}} = \frac{\alpha\rho}{\delta_d(\rho + \delta_d)} \times \frac{\beta\sigma V(t_0)}{(\xi + \gamma)(\sigma + \gamma)H}. \\
&= \sqrt{\frac{\alpha\beta\sigma\rho V(t_0)}{\delta_d(\xi + \gamma)(\sigma + \gamma)(\rho + \delta_d)H}}
\end{aligned}
\tag{4.5}
$$

The term $R_{0|t_0}$ assumes a fully susceptible population, at the initial introduction, given a specified month. In the non-seasonal model, we recall that the effective reproductive number, R_{eff} depends on susceptible depletion. In the seasonal model, the effective reproductive number depends both on susceptible depletion and current transmission which depends on the time of year (i.e. season). Susceptible depletion, in turn, depends on the season of introduction and time elapsed since. The term $R_{t|t_0}$ underscores that the effective reproductive number is time-varying, and also depends on the month of introduction. This reproductive number can be written by substituting S_h by $S_h(t, t_0)$ and S_v by $S_v(t, t_0)$ in the transmission matrix given in equation 4.4. The expression of $R_{t|t_0}$ becomes:

$$R_{t|t_0} = \sqrt{R_{t|t_0}^{VH} \times R_{t|t_0}^{HV}} = \frac{\alpha \rho S_h(t, t_0)}{\delta_d(\rho + \delta_d)H} \times \frac{\beta \sigma S_v(t, t_0)}{(\xi + \gamma)(\sigma + \gamma)H}. \tag{4.6}$$
$$= \sqrt{\frac{\alpha \beta \sigma \rho S_h(t, t_0)S_v(t, t_0)}{\delta_d(\xi + \gamma)(\sigma + \gamma)(\rho + \delta_d)H^2}}$$

In $R_{t|t_0}$, the number of susceptible individuals in both populations is dependent upon the variation over time and the month of introduction. In a further development, we are interested in the characteristics of the system after re-introduction (i.e. when there is existing seroprevalence in the human population). In the case of a re-introduction of the disease, the number of susceptible hosts S_h are further influenced by the existing seroprevalence level of the host population. The effective reproductive number $R_{t|t_0,S_h}$ is defined as the effective reproductive number that is time-varying and depends on the month of introduction, but given an initial seroprevalence level in the host population. The expression for this reproductive number is derived in the same manner as previously, but taking into account that the initial population of the susceptible humans at a particular month is not only dependent on t_0, but on the seroprevalence level of the host population as well. The resulting expression for $R_{t|t_0,S_h}$ becomes:

$$R_{t|t_0,S_h} = \sqrt{R_{t|t_0,S_h}^{VH} \times R_{t|t_0,S_h}^{HV}} = \frac{\alpha \rho S_h(t, t_0, R_h)}{\delta_d(\rho + \delta_d)H} \times \frac{\beta \sigma S_v(t, t_0)}{(\xi + \gamma)(\sigma + \gamma)H}. \tag{4.7}$$
$$= \sqrt{\frac{\alpha \beta \sigma \rho S_h(t, t_0, R_h)S_v(t, t_0)}{\delta_d(\xi + \gamma)(\sigma + \gamma)(\rho + \delta_d)H^2}}$$

The term R_h account for the number of recovered hosts present in the population. The above expression is used to investigate the effect of seroprevalence levels on dengue outbreaks in section 4.5.3. It is important to note that only the initial number of the susceptible humans

are of interest in the analysis performed in that section and the time-varying behaviour is not required. Therefore $R_{t|t_0,S_h}$ is simplified to $R_{0|t_0,S_h}$, that is just a single number dependent on a given value of t_0 and the number of recovered individuals present in the host population by setting $t = t_0$ from the above expression in the terms.

For the non-seasonal model RM_{SEIR} discussed in chapter 2, the derivation for the probabilities of a major outbreak after the introduction of one infectious vector and one infectious human, conditioned upon $R_0 > 1$ are (derivation in Appendix C.2):

$$P_{Inv|I_v=1} = 1 - \frac{R_0^{HV} + 1}{R_0^{HV}(1 + R_0^{VH})} \tag{4.8}$$

$$P_{Inv|I_h=1} = 1 - \frac{R_0^{VH} + 1}{R_0^{VH}(1 + R_0^{HV})}. \tag{4.9}$$

For the seasonal model, the roles of R_0, R_0^{HV}, and R_0^{VH} that previously gave an existing seroprevalence level, are replaced by $R_{0|t_0,S_h}$, $R_{0|t_0,S_h}^{HV}$, and $R_{0|t_0,S_h}^{VH}$ in measuring the invasion probabilities. The probabilities of a outbreak after the introduction of one infectious vector and one infectious host potentially with pre-existing seroprevelence level are:

$$P_{Inv|I_v=1,t_0} = 1 - \frac{R_{0|t_0,S_h}^{HV} + 1}{R_{0|t_0,S_h}^{HV}(1 + R_{0|t_0,S_h}^{VH})} \tag{4.10}$$

$$P_{Inv|I_h=1,t_0} = 1 - \frac{R_{0|t_0,S_h}^{VH} + 1}{R_{0|t_0,S_h}^{VH}(1 + R_{0|t_0,S_h}^{HV})}. \tag{4.11}$$

In Appendix C.2, it is mentioned that the infection in the system goes extinct if the basic reproductive number R_0 is below one. This condition is used to derive the expression for the probabilities of invasion; $P_{Inv|I_h=1}$ and $P_{Inv|I_v=1}$ respectively. For the non-seasonal system this is true as R_0 remains fixed so the estimation of the probability of major outbreak as $P_{Inv|I=1} = 1 - P(\text{ext})$, where $P(\text{ext})$ is the probability of extinction. Here, it is important to note that this invasion probability is obtained by estimating the probability of extinction of the disease. In the case of the seasonal model RM_{SEIR}^s, there are two major differences compared to the non-seasonal set-up for P_{Inv}: (i) R_0 is replaced by $R_{0|t_0,S_h}$ which is influenced by the point of introduction of the disease; (ii) The pre-existing seroprevalence level in the host population alters the probability of an outbreak. At higher seroprevalence levels, an invasion is less likely, even if a viremic individual is introduced in the favourable season.

4.4 Methods of quantitative analysis

The behaviour of the RM^s_{SEIR} was investigated using deterministic and stochastic models. As the system is non-autonomous (the birth rate of vectors $\delta_b(t)$ varies with time), it is converted into autonomous system by using the technique described in Appendix C.1. Converting into an autonomous system of ODEs ensures that both numerical and stochastic solutions remain 'well-behaved'.

The birth rate of the vector population varied over the course of the year whereas other rates affecting the vector and host population were set constant. The formula for the birth rate was

$$\delta_b(t) = \delta_b \left(1 + a\,sin\left(\frac{360t}{365} + shift\right) ^\circ \right) \tag{4.12}$$

Here, δ_b is the baseline value for the average birth rate, a is the amplitude of the fluctuations, $shift$ is the phase change according to the season. The baseline value of $\delta_b = \frac{1}{11}$, $a = 0.15$ and seasonal shift started from $0\,^\circ$ in April and increased by equal intervals of $30\,^\circ$ up to $330\,^\circ$ for the month of March. The vector-to-host ratio $\frac{V}{H}$ is kept fixed at one at 1st April. This month is at the peak of the dry season and it has the lowest number of vectors. For all other months this ratio is calculated from the endemic equilibrium by integrating the system of ODEs under the the disease-free condition. In general, the presence or absence of disease has no impact on the amplitude and frequency of the vector population cycle. During the sinusoidal fluctuation of the vector population, the term $\frac{V}{H}$ increases to a maximum value of 4.875 on 1st October. October is the peak of the wet season and as a result, demonstrates the highest number of mosquitoes. An initial condition matrix ICM was constructed for both deterministic and stochastic models. It consisted of three columns containing the information about the (i) month of the year (1-12), (ii) seasonal shift in the sine curve relative to that month ($0\,^\circ$ - $330\,^\circ$) and (iii) the vector-to-host ratio for that month.

One important decision in the seasonal model, where seasonal forcing is implemented using a sinusoidal function, is to choose the number of days in one calendar year. In most of the studies, one year consists of 360 days. This is usually done to synchronize the number of years and the period of the sine wave. In this study, instead of considering three hundred and sixty days, three hundred and sixty five days are considered in one year and while the shift is $30\,^\circ$ for each month, the vector-to host ratio is estimated by assuming 365 days in one calendar year. This is done to prevent the timing of 'simulation' month varying from the 'actual' month of the year.

4.4.1 Deterministic model

The deterministic version of RM^s_{SEIR} is introduced in Section 4.2. The system of ODEs described in equation set 4.1 was solved numerically from twelve different starting points, one for each month, and disease invasion was investigated by either introducing one viremic host or one viremic vector. The first twelve initial conditions were dedicated to looking at the dynamics of invasion after the introduction of one infectious host at different months; the second twelve studying invasion dynamics after the introduction of one infectious vector. At $t = 0$, the host population was either totally susceptible or has different seroprevalence levels. The vector population was $V - 1$ when the disease is introduced by a viremic vector. In other cases, all vectors were susceptible at $t = 0$. The time period of numerical integration varied for different initial conditions and is referred to in the plots.

4.4.2 Stochastic model

The event-based, stochastic version of RM^s_{SEIR} has seven state variables, four (S_h, E_h, I_h, R_h) for host populations and three (S_v, E_v, I_v) for vector populations that can take only positive integer values. The time evolution of these state variables is affected by nine different possible events shown in Table 4.1. The seasonal fluctuation in the birth rate δ_b introduces an additional time dependence in the event rates. These events occur in continuous time with rates that depend on the population values, mediated by seasonality. The evolution of the populations is modelled by using adaptive tau-leaping algorithm (Cao et al., 2007). The tau-leaping algorithm is an approximation of the Gillespie algorithm (an exact algorithm for simulating individuals in a population) and assumes that the change in reaction rates during one time step are negligible.

This method chooses the simulation time increment in an adaptive manner depending on the current state of the populations. At each time increment, each reaction is considered as an independent event, so the number of events of each type (one of nine types in this case) are drawn from a Poisson distribution with mean $a_j dt$, where a_j is the rate of event type j and dt is the time step. The error control parameter ϵ is set to 0.01, to avoid occurrence of more than one reaction during leaping in time. Further details of the method are provided in Chapter 3, section 3.5.2 whereas the algorithm is presented in Appendix B.1, and the choice of the highest order rate required to select ϵ is shown in Appendix B.2 of the same chapter.

Event	Effect	Transition Rate
Host :		
(i) Birth of susceptible host	$S_h \uparrow$	ΛH
(ii) Exposure of susceptible host	$S_h \downarrow$ & $E_h \uparrow$	$\alpha I_v S_h / H$
(iii) Infection of exposed host	$E_h \downarrow$ & $I_h \uparrow$	σE_h
(iv) Recovery of infected host	$I_h \downarrow$ & $R_h \uparrow$	ξI_h
(v) Natural death of a host	$S_h \downarrow,\ E_h \downarrow,\ I_h \downarrow,\ R_h \downarrow$	$\gamma S_h,\ \gamma E_h,\ \gamma I_h,\ \gamma R_h$
Vector :		
(i) Birth of a susceptible vector	$S_v \uparrow$	$\delta_b V$
(ii) Exposure of susceptible vector	$S_v \downarrow$ & $E_v \uparrow$	$\beta S_v I_h / H$
(iii) Infection of exposed vector	$E_v \downarrow$ & $I_v \uparrow$	ρE_v
(iv) Death of a vector	$S_v \downarrow,\ E_v \downarrow,\ I_v \downarrow$	$\delta_d S_v,\ \delta_d E_v,\ \delta_d I_v$

Table 4.1: The events in the stochastic RM^s_{SEIR} model are shown as stochastic transition rates. Here the subscript h denotes the host population and v denotes vector population. The direction of the arrow in second column *'Effect'* denotes either addition or subtraction of an individual to and from a compartment. Here the birth rate of susceptible vectors δ_b is time dependent, varying with seasons.

4.4.3 Value of the entomological parameters

In order to obtain quantitative predictions, the model was parametrised as described below.

Most of the Ross-Macdonald style host-vector models contain either the same or very similar quantities which are represented by different parameters (Smith et al., 2012). The way these parameters are defined and their quantitative values have a significant effect on the behaviour of the model. This suggests the importance of finding a suitable parameter space for modelling the dynamics of dengue. Many studies for modelling dengue were performed in different field and laboratory conditions. As a result, parameter values in the literature are based upon the data obtained from a specific locality and reflect the meteorological conditions, vector-to-host ratio, spatial distribution of hosts, lifestyle of people, daily commuting patterns and water storage practices of humans living in that area (Adams and Kapan, 2009; Harrington et al., 2005; Jansen and Beebe, 2010; Otero et al., 2011; Otero and Solari, 2010; Reiter et al., 2003).

After emerging as adults A. *aegypti* live in close proximity to humans, mainly in houses, containers and sheds. As noted in a longitudinal study performed by Harrington et al. (2005), most adult *Aedes aegypti* travel relatively short distances. Their strong anthropophilic nature and weak flying ability is documented in different studies (Harrington et al., 2005; Maciel-De-Freitas et al., 2007; Muir and Kay, 1998). In this study, it is assumed that a large proportion of the vector population moves near to the human habitat shortly after emerging as adults.

As they spend most of their lives in the artificial habitat, there is a very low impact of seasonality on their rates of incubation and mortality. The average death rate of adult vectors δ_d is thus taken to be independent of temperature and $\delta_d = \frac{1}{11}$ days^{-1} in the temperature range of $5\,°C$ to $30\,°C$ (Christophers, 1960). Similarly the average rate of extrinsic incubation in vectors $\rho = \frac{1}{10}$ days^{-1} is modelled to be independent of seasonal conditions (Otero and Solari, 2010). The transmission rates are α and β are taken as 0.55 days^{-1}. These rates ranged from $0.2 - 0.67$ days^{-1} in Adams and Boots (2010) and were 0.5025 days^{-1} in Bartley et al. (2002). For dengue modelling, usually $\alpha = \beta$ and they are represented as a product of constant bite rate (per vector) and the transmission probability from an infectious individual of one population to other. In this study, higher transmission probabilities are chosen to lead to five yearly cycles of dengue after invasion. The dengue outbreak data from different studies reveal fluctuations with a period of between 3 and 4 years (Nishiura, 2006; Wearing and Rohani, 2006), whereas individual serotypes have longer periods and cycle in and out of phase (Wearing and Rohani, 2006). The impact of high transmission probabilities is shown implicitly through high transmission rates. The list of parameters for dengue used in the modelling process are given in Table 4.2.

Symbol	Explanation	Value used	Reference
α	Vector-to-host transmission rate, in days^{-1}	0.55	Adams and Boots (2010)
β	Host-to-vector transmission rate, in days^{-1}	0.55	Adams and Boots (2010)
$\frac{1}{\xi}$	Average infectious period of hosts, in days	7	Adams and Boots (2010)
γ	Birth / death rate of hosts, in days^{-1}	4.215×10^{-5}	-
$\frac{1}{\sigma}$	Average latent period in hosts, in days	5	Adams and Boots (2010) Newton and Reiter (1992)
δ_d	Mortality rate of vectors, in days^{-1}	0.0909	Otero and Solari (2010) Otero et al. (2008)
$\frac{1}{\rho}$	Average latent period in vectors, in days	10	Otero and Solari (2010) Otero et al. (2008)

Table 4.2: List of parameters used in RM_{SEIR}^s. Here the average birth rate of hosts is equal to the death rate, i.e, $\Lambda = \gamma$. The life expectancy of a single host is set to 65 years, so birth / death rate of hosts becomes $\frac{1}{65 \times 365}$ days$^{-1} = 4.215 \times 10^{-5}$ days^{-1}.

4.5 Results from deterministic model of RM_{SEIR}^s

This section reports the main findings of the chapter obtained by using the deterministic framework. The deterministic compartmental model which was presented in Figure 4.1 is used to explore the research questions of interest described at the end of section 4.1. In particular,

the research questions addressed are: (i) how does the time evolution of $R_{t|t_0}$ depend upon the month of dengue introduction?, (ii) how do these patterns relate to underlying $R^{HV}_{t|t_0}$ and $R^{VH}_{t|t_0}$? and (iii) how does the probability of (re-) invasion with a different serotype change with the pre-existing seroprevalence? All of the investigations are conducted by separately considering the case of (re-) introduction by a viremic human or a viremic mosquito. The results in this section are reported after a formal analysis of the model. The model is first examined for the time evolution of the following quantities: (i) per-capita birth rate $\delta_b(t)$, (ii) population growth rate $M(t)$ and (iii) the total vector population $V(t)$. Then the dynamics of infectious individuals in both populations, after disease introduction by a single viremic individual is presented. The behaviour of the model at the endemic equilibrium concludes the preliminary analysis of RM^s_{SEIR}. After the analysis of the model, the presentation of the text follows the order of the questions.
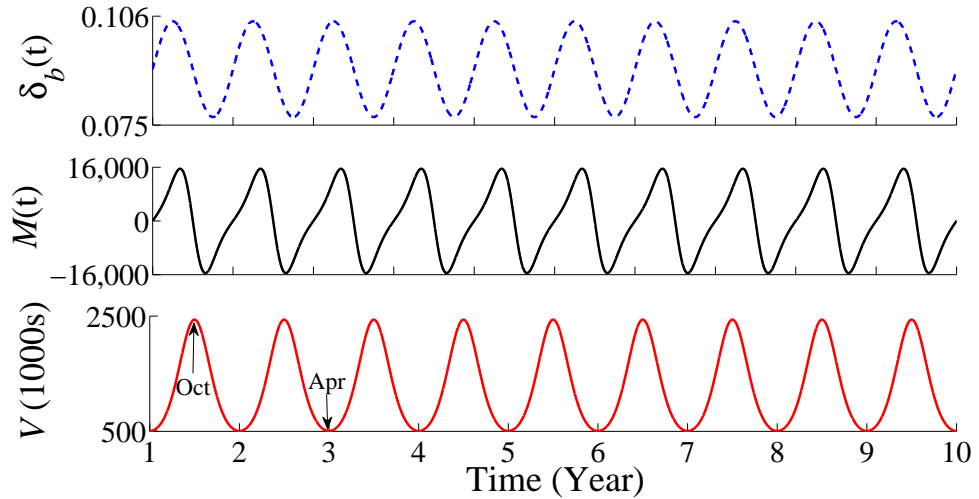
Figure 4.2: Time series plots of variation in per capita birth rate $\delta_b(t)$, population growth rate $M(t)$ and the vector population size V. All quantities are shown by integrating the ODE system defined in equation set 4.1 from April. In this study, it is assumed that there are two seasons per year: a favourable (wet) season that peaks at the start of October and an unfavourable (dry) season that peaks in April. These terminologies of season are defined on the basis of change in $M(t)$, which affects the total number of vectors V. Both seasonal extremes are pointed out by arrows in the last panel. The ratio between hosts and vectors $\frac{V}{H}$ is one on 1st April. Host population H at the start is five hundred thousand individuals.

Population parameters related to the demography of the vectors are presented in the three panels of Figure 4.2. Here $\delta_b(t)$ is the per capita birth rate, $M(t) = (\delta_b(t) - \delta_d)\,V(t)$ is the population growth rate and $V(t)$ is the population fluctuation of the vectors. The relationship between these panels is explained as: the change in $M(t)$ (middle panel) is a result of change in $\delta_b(t)$ (top) and the change in V (bottom) is a result of change in $M(t)$. The figure was

plotted by integrating from 1st April, where $shift = 0\,°$. It was solved numerically for ten years under disease-free initial conditions where the host-to-vector ratio was set to one, so both populations had 500,000 individuals at the start. The month of April corresponded to the peak of the dry season in the modelling settings and the trough of the vector population occurred in mid-April (indicated by an arrow in the last panel). The per-capita birth rate of vectors follows a sinusoidal cycle as described in Equation 4.12. The cycle starts from the base-line value, which is is $\frac{1}{11}$. In the middle panel, the population growth $M(t)$ at any instant t is found as the product of vectors present at that time and the difference between birth and death rates. The impact of growth rate $M(t)$ on the change in vector population is explained as follows: As mentioned in Equation 4.2, the vector population grows if the quantity $M(t)$ is positive. When the population growth rate remains positive (i.e, birth rate is higher than the death rate) the plot shows the growth of the vector population up to nearly two and a half million individuals in the month of October (indicated by an arrow). For the negative values of $M(t)$, the population starts falling until it reached to five hundred thousand vectors, thus completing the cycle.

The time evolution of both infectious populations is shown in Figure 4.3 after the introduction of a single viremic host (first and third row) or a single viremic mosquito (second and last row). The description of the first two rows is as follows: They clearly showed the similar temporal behaviour for the first nine months after the disease is introduced. A susceptible bottleneck occurred in S_h after the epidemics and the numerical solver is run for only one year since the infectious populations reached fractional values ($< 10^{-5}$) after six to eight months, except then these fractional values are allowed to persist for fifteen years in the lower two panels of Figure 4.3. The initial host population consisted of five hundred thousand individuals and the vector population is adjusted according to the $\frac{V}{H}$ ratio for that month. It is interesting to note that irrespective of the starting month, the epidemic peak occurs in between two and four months after the introduction of an infectious human (top) or mosquito (bottom). The explanation of this phenomenon lies in Figure 4.4, in which a very small proportion of S_h is present at the endemic equilibrium. After an introduction, initially the impact of seasonal dynamics on the system is overtaken by the number of susceptible humans present in the population. This leads to an outbreak shortly after the disease is introduced into the community. The last two rows showed the medium-term transient dynamics of the infectious populations with the peak of the first year not shown. In these figures, the dengue epidemic showed a cyclic behaviour of roughly five years.

The effect of seasonal variation in Figure 4.3 for the first nine months can be described as follows: The impact is stronger on the vector population as it has the maximum number of infectious individuals. $I_v > 500,000$, if the numerical solver is initiated at favourable seasons
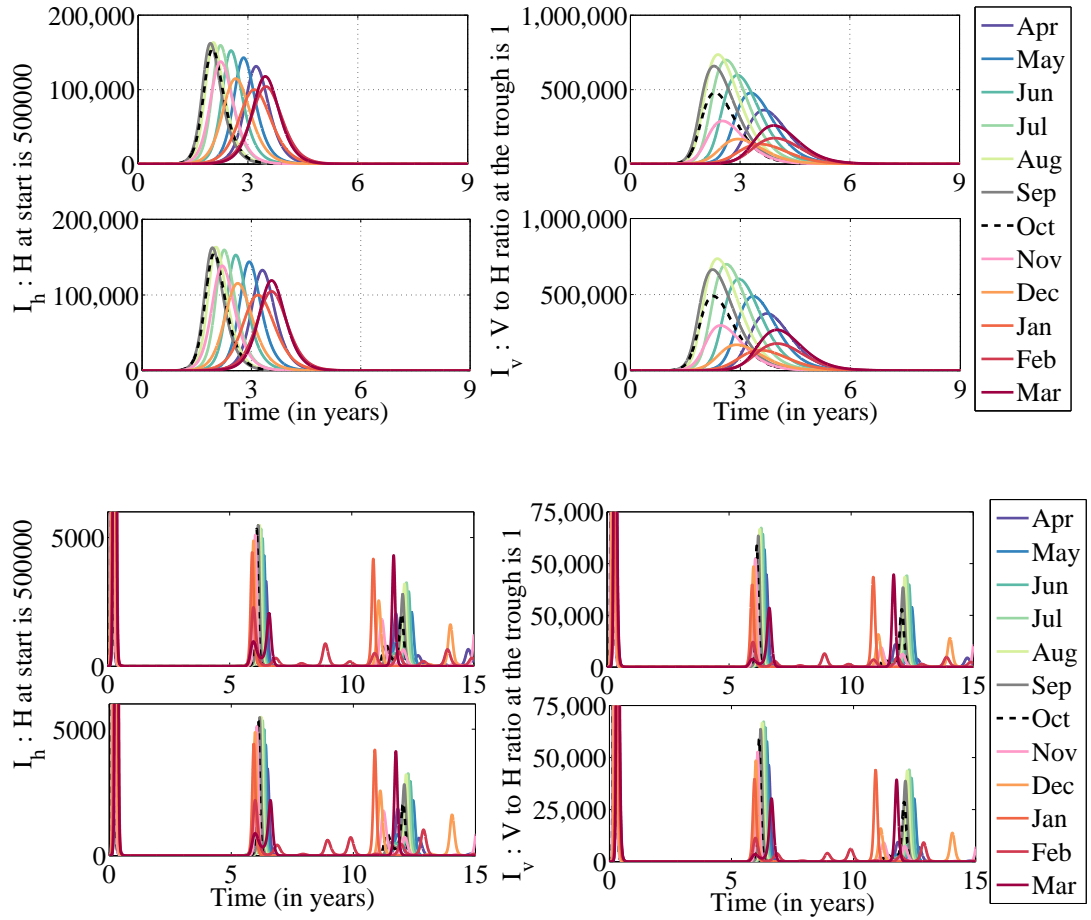
Figure 4.3: Seasonal time series plot for the infectious compartment of both populations. In the first and the third row, the epidemics are caused by introducing a single infectious human in a totally susceptible vector population. In the second and the last row, a viremic vector is introduced in a totally susceptible human population. The last two rows shows the long term behaviour of the infectious population with the peak of the first year not shown. Epidemics occur shortly after the introduction.

and $I_v < 250,000$ if it is started from the dry seasons (right column). If the invasion occurs in January, the peak of the epidemic is in April which is the driest month of the year. That why the peak has the lowest number of both infectious vectors $\approx 150,000$ and infectious humans, $\approx 100,000$. The peaks in both infectious populations occur later following introductions in February as compared to April; however fewer infectious mosquitoes are present at the peak. This shows that the total number of I_h at the peak is less affected by the change in seasons than the total number of I_v. In the human population, the peaks of the infectious population showed less variation, the maximum number of infectious humans are $\approx 165,000$ and minimum number is $\approx 100,000$. If the invasion starts from either in January or February, the smallest peaks are observed just before or after the month of April. The relationship

between the highest peaks with faster invasion and the lowest peaks with delayed invasions can be explained in terms of population growth rate of vectors $M(t)$. Higher peaks and faster invasion is observed when the $M(t)$ is at the highest values (August-October) at the tome of introduction. For introductions in the month of November to January, the peaks of I_h starts shifting later and a considerable decline is observed in the peak of I_h since $M(t)$ has the lowest value in January.

4.5.1 Behaviour of the model at equilibrium
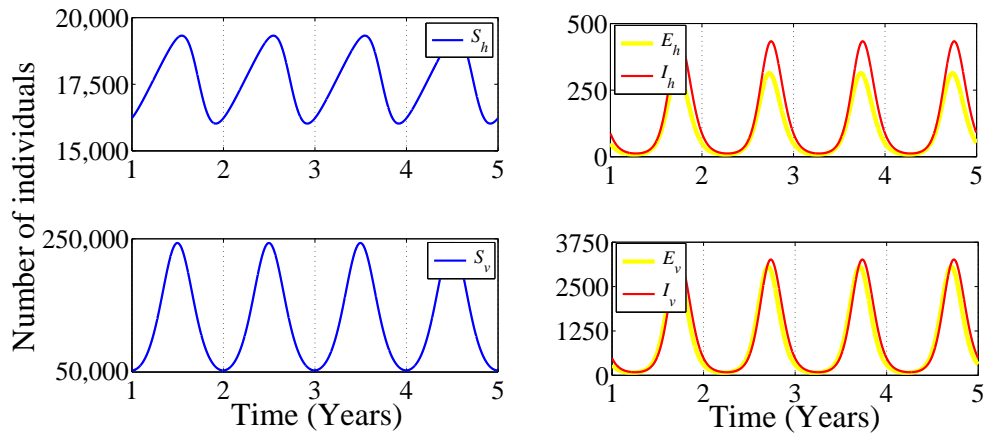
Figure 4.4: Compartments of RM_{SEIR}^{s} at endemic equilibrium when started from 1st April. The initial value of H and V is set to five hundred thousand individuals. The plots show the endemic equilibrium for six out of seven compartments of RM_{SEIR}^{s}. The instantaneous number of recovered humans can be obtained from the relation $H = S_h + E_h + I_h + R_h$. The results were obtained by running the numerical solver from April and the years in x-axis are the number of years after the oscillating endemic equilibrium was attained.

This section presents the model's behaviour at the deterministic equilibrium. The oscillating behaviour of the compartments of the model at the endemic equilibrium is presented in Figure 4.4. The differential equation model for RM_{SEIR}^{s} given in 4.1 is integrated numerically starting from the month of April with 500,000 hosts and vectors and allowed to run until the transient period has passed and an oscillatory endemic equilibrium is attained. At equilibrium, the population of hosts consists of almost all immune individuals. At this state, small annual outbreaks in the human and mosquito population are observed and the number of infected individuals E_h and E_v are closely tracked by the infectious population I_h and I_v respectively. Here the number of years on the x-axis represents the output from the last five years of the solution.
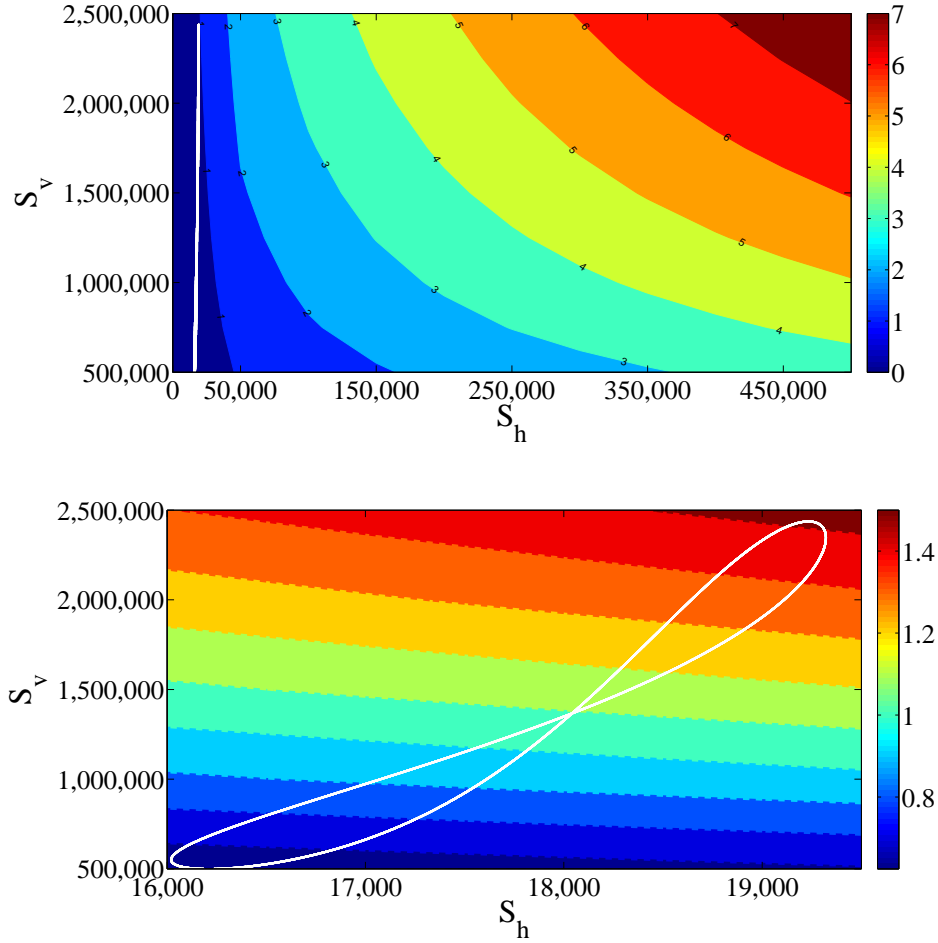
Figure 4.5: Phase diagram showing $R_{t=0|t_0}$ values at equilibrium for different months. In both figures, the solid curves are contours of $R_{t=0|t_0}$ in the S_h-S_v plane. Top: Plot showing the critical susceptible numbers S_h^{crit} and S_v^{crit} that are required for disease endemicity. Here S_h denotes the susceptible humans and S_v represents the susceptible vectors. The plot is obtained by solving the equations with one infectious host and allowing the model to run until the oscillating endemic equilibrium state is attained. At equilibrium, the blue vertical line indicates the equilibrium trajectory. Bottom: Zoomed view of the region of interest. The white bow-tie structure corresponds to the white line of the left figure. The contour lines of $R_{t=0|t_0}$ are varied from 0.7 to 1.5 so when $R_{t=0|t_0} > 1$, susceptible depletion occurs reducing it below 1 and halting the dengue transmission.

The model is used to estimate the critical number of susceptible individuals required for recurrent dengue epidemics in both host and vector populations. In Figure 4.5, a phase-space plot between S_h and S_v at endemic equilibrium is drawn by numerically integrating over twelve starting points, one for each month. The solid curves are contours of $R_{0|t_0}$ in the S_h-S_v plane. Irrespective of the month of the year chosen to start the numerical solver, the same bow-tie structure is obtained ensuring that the oscillating equilibrium state is locally stable. The plot on the top of Figure 4.5 is the zoomed view of the structure (blue line) on the left plot. At the start, both host and vector populations consist of 500,000 individuals. The

critical susceptible density of hosts S_h^{crit} at equilibrium was found to be between 16,000 and 19,500 individuals. The critical susceptible vector density S_v^{crit} ranges from 50,000 to 250,000 individuals. The bow-tie structure on the bottom of Figure 4.5 is explained as follows: As the susceptible host population reach levels that yields $R_{0|t_0} > 1$ (values along the contour lines), dengue transmission starts and the number of susceptible hosts is reduced. The number of S_h starts falling causing $R_{0|t_0}$ to fall below one, thus halting the transmission. These results shows that re-invasion by the same serotype of dengue requires a very small fraction of susceptible humans.

The endemic behaviour of the model highlighted in Figures 4.4 and 4.5 also gives information about the disease-free condition. For dengue disease, it is known that infection caused by one serotype provides life-long immunity to that serotype whereas infected patients have limited cross-immunity against remaining serotypes (WHO, 2014). In the case of dengue re-invasion from a different serotype, a very large pool of susceptible hosts is present, i.e, $S_h \approx H$. This causes bigger and quicker epidemics. The worst case scenario can be seen from Figure 4.3 which shows huge peaks of epidemics, irrespective of the month of introduction.

4.5.2 Time evolution of reproductive numbers: $R_{t|t_0}$, $R_{t|t_0}^{HV}$ and $R_{t|t_0}^{VH}$
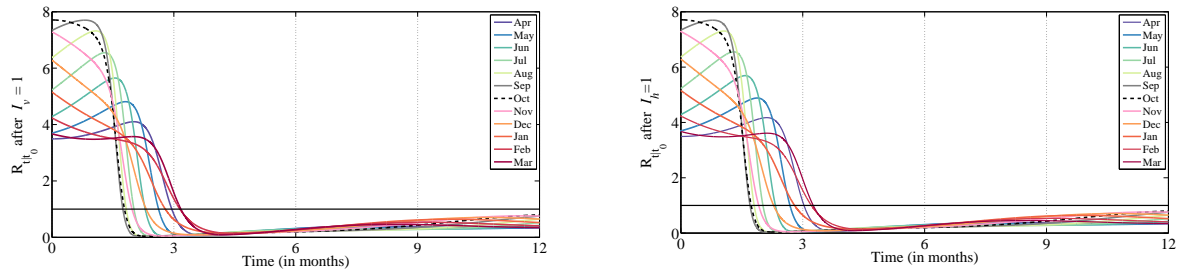
Figure 4.6: Time evolution of the reproductive number. Seasonal variation in $R_{0|t_0}$ with respect to different starting conditions during one year post introduction with one infectious human (left) and one infectious vector (right). Each line corresponds to the time evolution for a given month of introduction. Irrespective of the initial conditions, $R_{0|t_0}$ quickly falls below one. The host population H comprises five hundred thousand individuals.

The question addressed in this section highlights the importance of the month of disease introduction upon the time-varying basic reproductive number $R_{t|t_0}$. We further want to compare the dynamics when dengue disease is introduced by an infectious vector or an infectious human. The first situation can be thought as a community having a lot of infectious mosquitoes at the boundary. Disease can then be transmitted to the neighbouring community having no history of dengue. The second case corresponds to humans travelling from one

place to another. In both cases, the effect of the month of dengue introduction are discussed in the next paragraphs.

Figure 4.6 presents the time-evolution in $R_{t|t_0}$ during a year. Temporal patterns in $R_{t|t_0}$ on the left hand side is when the disease is introduced by one infectious vector (situation one). The right hand side denotes the second case of dengue, i.e., introduced by one viremic host. The point where $R_{t|t_0} = 1$ on the y-axis is indicated by a black horizontal line. Three points are worth mentioning in these plots. (i) The time evolution of the reproductive ratios are almost identical, irrespective of whether disease is introduced by host or vector. (ii) The time of disease introduction into a naive population has a strong impact on the reproductive number. Higher values of $R_{t|t_0}$ are obtained when the viremic individual is introduced in the wet season. (iii) Starting at a range of values that are roughly between 3.5 to 8, $R_{t|t_0}$ quickly falls below one. From May to September the quantity initially grows with increasing slope for each month before sharply falling to lower values. The maximum time delay before the fall occurs in the month of May whereas an immediate decline is observed during November to February. By comparing these observations to Figure 4.2, the vector population V increases as it moves away from April (peak of dry season) with maximum number of individuals in October (peak of wet season), therefore the time delay required for $R_{t|t_0}$ to peak gets smaller as an infected individual is introduced in later months till September. From the months of October to February, the rate of fall in $R_{t|t_0}$ gets less as we move away from the peak of the wet season. All these points are in agreement with the temporal patterns that reflects the changes in the number of the infectious individuals in Figure 4.3.

The decomposition of the temporal patterns in terms of population based reproductive ratios are presented in Figures 4.6 and 4.7. $R_{t|t_0}^{HV}$ denotes the instantaneous number of secondary cases in the vector population and $R_{t|t_0}^{VH}$ for the human population. Similar patterns for $R_{t|t_0}^{HV}$ and $R_{t|t_0}^{HV}$ are found for both cases. $R_{t|t_0}^{HV}$ quickly goes up after an epidemic as it depends upon the ratio of the number of susceptible vectors S_v to the host population H. The dip that occurs at around three months on the left column gets deeper if the disease is introduced in the favourable season. $R_{t|t_0}^{HV}$ never goes below one during the first year. In contrast, $R_{t|t_0}^{VH}$ which depends upon S_h quickly falls to very low values (< 1) as all the hosts become infected soon after the introduction of an infectious individual causing a susceptible bottleneck in hosts halting the transmission. The left column of the above figure indicates that the vector population has a lot of susceptible individuals present to invade the system during the first year. The right column shows that the secondary cases in the host population caused by a single infectious vector, $R_{t|t_0}^{VH}$ are, on average, fewer than one shortly after the introduction. In summary, the introduction of an infectious individual under current parameter settings creates a huge epidemic at varying time conditioned upon the month of the introduction. This
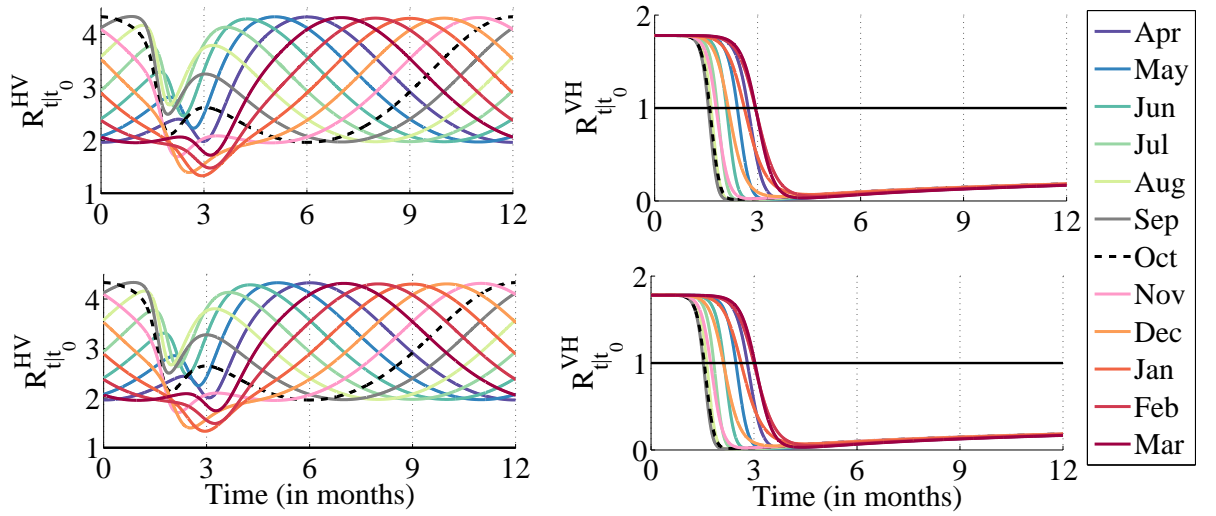
Figure 4.7: Population based reproductive ratios; $R_{t|t_0}^{HV}$ and $R_{t|t_0}^{VH}$. The plot shows the population based reproductive numbers for all months, for the case of: (i) introduction of dengue virus by an infectious vector (top row) and (ii) introduction by an infectious host (bottom row). Initial conditions are same as in Figure 4.6. Black horizontal lines mark the value of $R_{t|t_0} = 1$ in all sub-figures.

leads to an epidemic burnout of susceptible hosts during the first year (see Figure 4.3). In addition, the vector versus human introduction does not create a difference in the dynamics of an epidemic.

4.5.3   Effect of seroprevalence on outbreaks

The final investigation using the deterministic model was to examine whether the probabilities of (re-)invasion changes with pre-existing seroprevalence in humans and if this depend on whether re-introduction of dengue is caused by a vector or a human. This situation is important for two reasons: (i) the design of health policy in communities where a different serotype of dengue invades after an initial outbreak from one serotype in the past, and (ii) re-invasion from the same serotype but after a long time so that the susceptible pool has grown large enough during that period. The Figure 4.8 highlights this situation. This figure is obtained by integrating the deterministic model for one year and considering different seroprevalence levels (0% to 100%) in hosts. For every one percent increase in the seroprevalence level, the values at $t = 0$ for the invasion probabilities $P_{Inv|I_h=1,t_0}$, $P_{Inv|I_v=1,t_0}$ (which are defined in section 4.3.1) and $R_{0|t_0,S_h}$ were plotted. This procedure is repeated for all months and for both (re-)introduction conditions.

The probability of invasion in the first column shows a gradual reduction until the sero-prevalence level of 65%. The decrease in the invasion probabilities becomes rapid when the seroprevalence levels are greater than 65%. This change of behaviour is more obvious if one infectious host is re-introduced at high seroprevalence levels. In the case of introduction of one infectious vector, the transition is smoother. The change in $P_{Inv|I_h=1,t_0}$ with respect to seroprevalence is more spread-out and has a larger effect than $P_{Inv|I_v=1,t_0}$. At the start when $t = 0$, $P_{Inv|I_h=1,t_0}$ is always higher than $P_{Inv|I_v=1,t_0}$. At $t = 0$, the difference between the life span of a host and the vector is the reason for this disparity. Moreover, the highest probabilities of invasion are found if the introduction / re-introduction is from September to November as the peak of vector population occurs in October as shown in Figure 4.2. More-over, this difference decreases when the season of (re-)introduction approaches the trough of the vector population. During March, April and May the population of vectors V is very close to the number of humans available and hence difference in the invasion probabilities is the lowest.

Comparison of the plots in the left column of Figure 4.8 also shows the asymmetric behaviour of both invasion probabilities, whereas the product of the individual reproductive numbers, $R_{t|t_0,S_h}$ remains the same in both cases. This is due to the fact that they were obtained by taking the composition of two functions, which does not follow the commutative law (Lloyd et al., 2007). The right-hand side column shows the change in $R_{t|t_0,S_h}$ with increasing seroprevalence. $R_{t|t_0,S_h}$ remains the same for the two (re-) introduction conditions. The horizontal line at $R_{t|t_0,S_h} = 1$ corresponds to the deterministic threshold condition for the invasion probability as it become zero below this line. This is the reason why during some months, the invasion probabilities fall to zero even before 100% seroprevalance level.

4.6 Results from Stochastic model of RM^s_{SEIR}

This section is devoted to the results obtained from the stochastic version of RM^s_{SEIR} de-scribed in section 4.4.2. The stochastic version of the model is used to ask the following research questions: (i) how does the probability of persistence of dengue infection depend upon the month of introduction? (ii) How is the time to extinction of dengue t_e and the starting conditions in the seasonal model related and how they are related to the type of individual that introduced the disease? (iii) how the patterns related to the peak of the infectious humans I_h and time taken to attain peak change with the month of dengue intro-duction. These investigations are conducted in similar fashion to section 4.5 by separately considering the case of introduction by an infectious human or an infectious mosquito. In the next sections these questions are addressed in a sequential order.
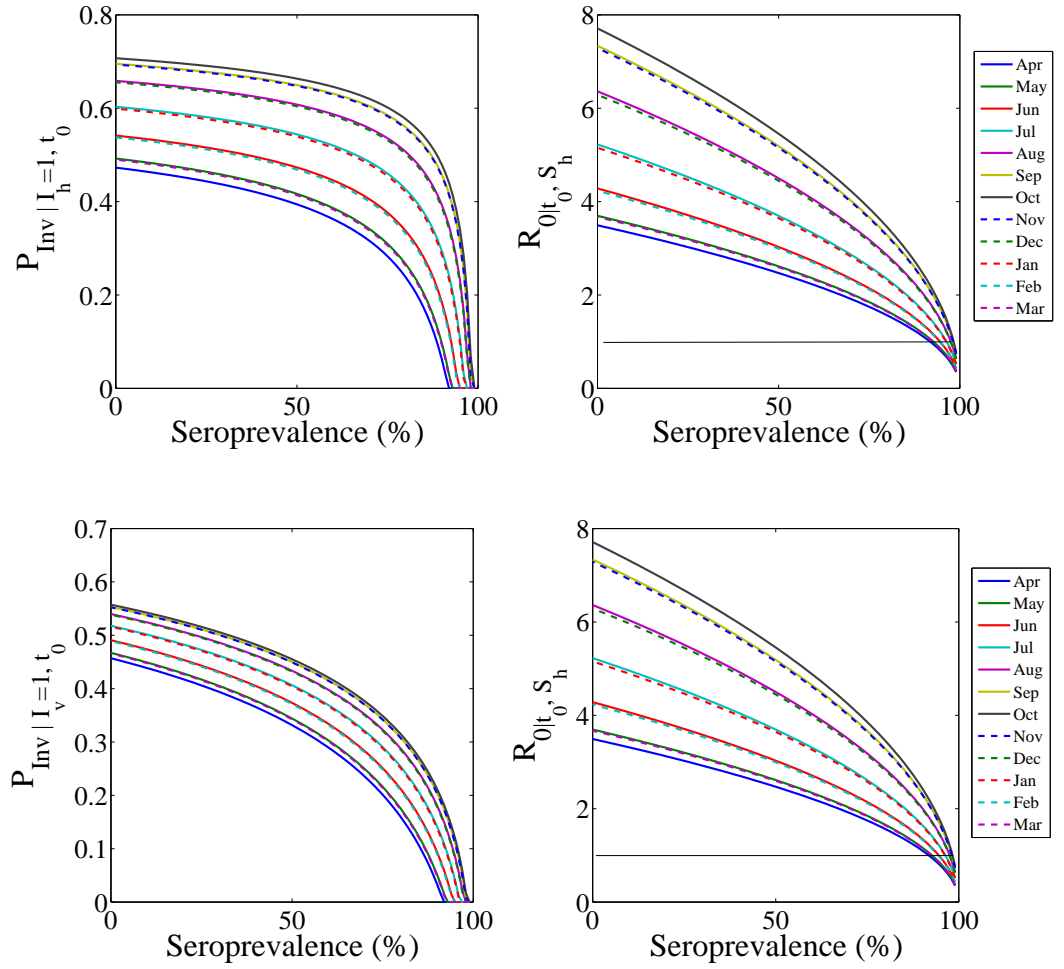
Figure 4.8: Probability of (re-) invasion as a function of pre-existing seroprevalence $P_{Inv|I_h=1,t_0}$ and $P_{Inv|I_h=1,t_0}$. Initial values of both probabilities and $R_{0|t_0,S_h}$ is displayed at different seroprevalence levels for all months.

4.6.1 Dependence of the probability of persistent runs on seasons

The baseline parameters for RM_{SEIR}^s described in Table 4.2 are used to investigate whether the infection is present in both human and mosquito populations if the disease is introduced at different times of the year. For this purpose, the probability of persistent runs is estimated. The time period of stochastic simulations is set to one year and dengue virus is introduced in naive host and vector populations by an infectious individual. This procedure is repeated one thousand times for 12 different dates of arrival of one infectious human and one infectious mosquito in the susceptible human and vector populations. The host population size (H) is initially 5×10^5 and the size of the vector population is estimated as defined in section 4.4.

Figure 4.9 shows the proportion of runs persisting to one year out of one thousand stochastic

repetitions, after introduction by either a viremic host (blue star) or a viremic vector (red circle). In general, the proportion of persistent runs in both cases shows similar behaviour except for the months of January to March where the Probability of persistence is nearly 4% greater if the disease in introduced by a viremic vector. The impact of seasonal changes can be explained as follows: if the disease is introduced in the favourable season, very few of the stochastic runs retain infection after one year whereas the chances of retaining infection if the disease is introduced in unfavourable season are higher. If an infectious host or a vector arrives in the population three months before the peak of the unfavourable season, it is more likely to spread infection in the long term as compared to arrival before or immediately after the peak of the favourable season. The relationship between dengue persistence and peak number of I_h with change in seasons is further explored in the next section.
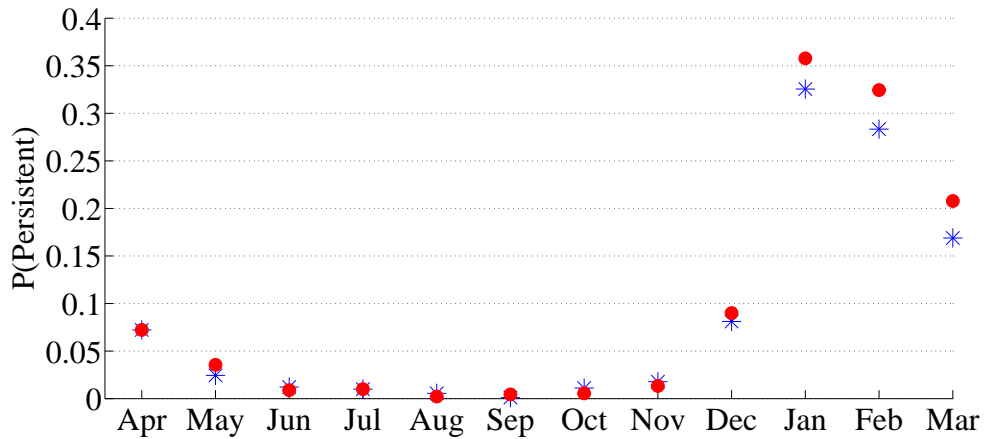
Figure 4.9: Probability of runs in which dengue infection is persistent in the population. One thousand stochastic repetitions are performed for each month. The blue stars shown the proportion of persistent runs when disease is introduced by an infectious human whereas the introduction by a single viremic vector is shown by red circles. The parameters for the stochastic model are taken from Table 4.2. The host population size (H) is initially 5×10^5 and the size of the vector population is estimated as defined in section 4.4.

4.6.2 Distribution of time to extinction of the disease

Figure 4.9 can be used to find the proportion of stochastic runs in which the infection is extinct after one year but it does not provide information about the distribution of the stochastic repetitions around extinction of the disease. To further look at the patterns of dengue extinction, Figure 4.10 is plotted. Here t_e is defined as the time to extinction of disease in both of the populations in one year. There are two cases in this figure: (i) the distribution of t_e after disease invasion caused by an infectious vector and (ii) the distribution

of t_e after disease invasion caused by an infectious host. Red circles denote situation one whereas situation two is denoted by blue stars. The size of host and vector population is the same as in previous figure. The numbers written at the top of every month show the number of stochastic repetitions that went extinct during one year. The first row denotes the number of runs for situation one and the second row for situation two.

The impact of seasonal differences is evident from Figure 4.10. The Figure shows a clear bimodal pattern of t_e. There is either an earlier extinction of the disease after the introduction (less than three weeks with most of the extinction during the first week) or persistence until seven months to a year. There are less early extinctions of disease in the favourable months and the distribution of persistent runs (i.e., infection is more than seven months) shows early extinction of dengue in these months as compared to the simulations from the unfavourable season. The unusually low frequency of disease extinction during January to March can be matched to the pattern shown in Figure 4.9.
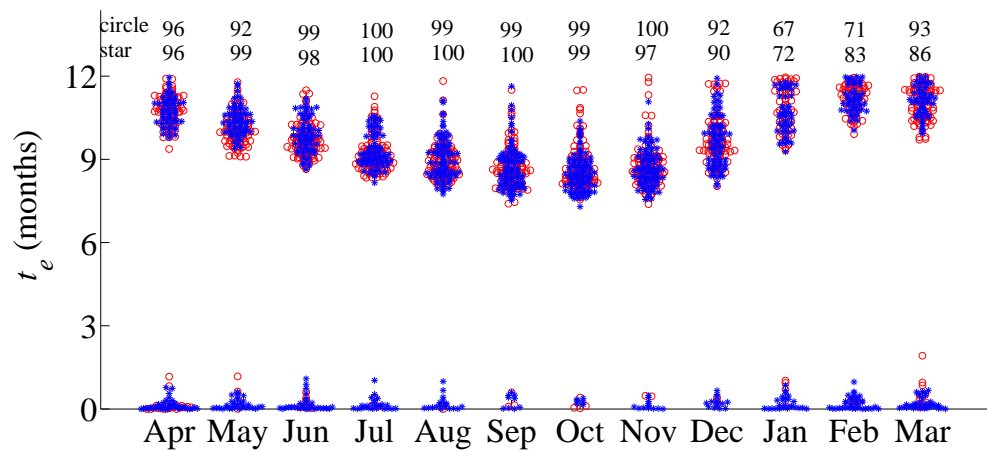
Figure 4.10: Distribution of t_e. Here red circles shows the probability of time to extinction in less than one year after the introduction of dengue with one infectious vector and blue stars shows the same quantity after the introduction with one infectious host.

4.6.3 Seasonal variation in timing and peak of infectious humans

Estimating the number of infectious people during an epidemic is crucial for public health planning and infrastructure. The time evolution of seasons during an epidemic can affect the patterns of I_h. As mentioned as the last question in Section 4.6, Figure 4.11 is constructed to investigate the maximum number of I_h during an epidemic at different starting points. In this figure, only invasive runs are plotted. Here dengue invasion is defined if there is more than

one secondary case in the human population. The patterns related to the peak of infectious humans I_h are shown at the top of Figure 4.11. The numbers written at the top of every month shows the number of invasive runs after the introduction of an infectious vector (top row) or an infectious host (bottom row). The population of hosts and vectors are the same as in previous sections. The bottom half of the figure denotes the time taken to attain peak.

In Figure 4.11, an important point in the top panel is the difference in the number of invasive simulations. If disease is introduced by an infectious vector in April, there is an 80% chance that the disease persists. I_h is going a peak around 130,000-145,000 individuals whereas in the case of disease introduction by an infectious human, this probability reduces to 64%. Although the peak number of I_h is nearly the same in the case of both introductions, the probability of attaining these peak values is different. In general, the chances of invasion are greater if one infectious vector is introduced in the population.

The top panel of Figure 4.11 shows the distribution of the patterns of peak I_h to the seasons. Higher number of infectious humans are present at the peak if the simulations are run from July to September. Similarly, the lowest number is when the introduction of dengue is in January. The pattern shown here closely matches with the cyclic pattern of the time evolution of the vector population V in Figure 4.2 with a time-lag of nearly three to four months. This is exactly what is present in the bottom panel of Figure 4.11 that shows the time in months for the I_h population to attain its maximum value.

4.7 Conclusion and discussion

This chapter further extends the host-vector model, RM_{SEIR} developed in Chapter 2 for modelling transmission dynamics of dengue and investigating the persistence of the virus in the host and vector population. The model constructed here is for a single serotype of dengue with no vertical transmission. The mosquito birth rate $\delta_b(t)$ is seasonally dependent, considered to be mediated by temperature, rainfall and humidity. The new model is termed RM_{SEIR}^s, the superscript s denoting seasonality. The birth rate of the vectors varies throughout the year during the wet (favourable) and dry (unfavourable) seasons. The deterministic version of RM_{SEIR}^s was presented first and and the terms of the model are explained. Methods were divided into qualitative and quantitative sections, where analytic derivations of different measures are presented in the first part followed by the numerical methods used for the deterministic and stochastic versions of the models. The time evolution of the vector population, V and the behaviour of the model at equilibria is presented. Different scenarios are presented for the investigation of (i) the influence of varying seasonal conditions on
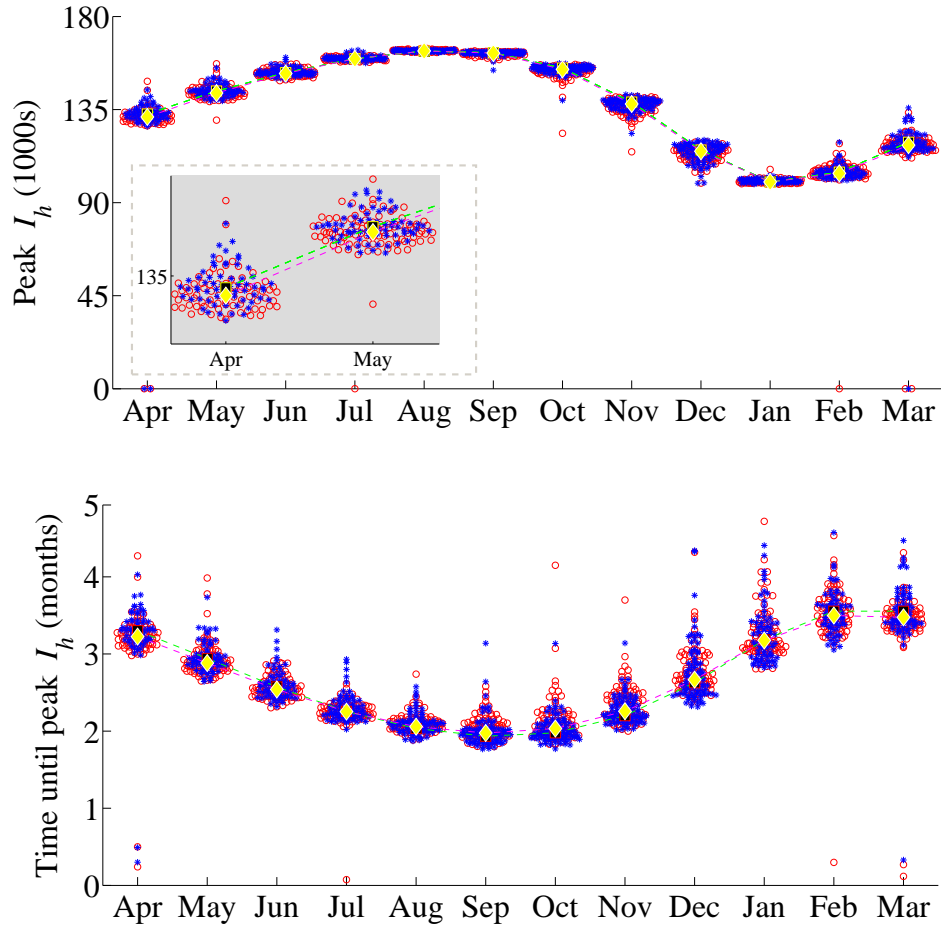
Figure 4.11: Beehive plot shows the time taken by I_h to reach the peak value, after the introduction of dengue disease from a viremic human (blue stars) and viremic mosquito (red circles) at different times of the year. The month at the y-axis of the bottom plot is the month in which median peak time occurs. In both figures, the black squares connected by green coloured lines are the peak I_h (top) and time to peak (bottom) values from the deterministic model for every month, as a result of introducing a single viremic human. Similarly, the yellow diamonds connected by magenta coloured lines are peak I_h (top) and time to peak (bottom) values from the deterministic model for every month, as a result of introducing a single viremic vector. These peak values are mentioned in the time series plots in Figure 4.3. The numbers written at the top of every month shows the number of invasive runs after the introduction of an infectious vector (top row) or an infectious host (bottom row), for both figures. The host population is taken as 5×10^5 individuals. Inset of the top figure: The zoomed view of the peak I_h for the month of April and May.

dengue persistence, and (ii) the effect on the persistence of disease if introduced by a viremic host or a viremic mosquito.

From the literature cited in Section 4.1, it can be concluded that the temperature, rainfall and relative humidity are the most influential parameters, both for modelling the dynamics of the vector population and the incidence of dengue disease. The effect of the above seasonal factors is simulated implicitly in the current study through the change in total vector population size. The birth rate of vectors, $\delta_b(t)$ fluctuates with seasonal changes. Thus vector recruitment over time is driven by season, and its value alters as a function of temperature, rainfall and humidity. For simplicity, death δ_d and incubation rates ρ of mosquitoes (per day) are assumed constant throughout the time period of simulations (Otero et al., 2008; Otero and Solari, 2010). The only rate which varies with the change in season is the birth rate of vectors δ_b. The population of male mosquitoes is not considered. Estimates of sex ratios are dependent upon the method of collection (traps, aspirators etc), time of the year when sampling takes place and geographical location of the study. Chen and Hsieh (2012); de Castro Medeiros et al. (2011) considered this ratio to be 2-fold the number of humans. For simplicity, I considered this ratio to be 1:1 in the dry season.

The time evolution of both infectious populations showed the same temporal behaviour for after the disease is introduced. Smaller peaks of infectious individuals are observed if the disease is introduced prior to the peak of the dry season (April) and higher peaks of infectious individuals are seen if the infection is introduced prior to the peak of the wet season (October). The high seroprevalence level in the human population leads to the extinction of the disease (I_h and $I_v < 10^{-6}$) within a year post introduction. The high seroprevalence level is also found by Bartley et al. (2002) from the dengue data of Bangkok. There is a time-delay of two to four months between the introduction of the disease and the peak of the infectious population which is caused by the number of susceptible vectors present at that time of the year. The results from RM_{SEIR}^s suggests a delay of one and half to three months depending upon the number of susceptible individuals present in that season. Andraud et al. (2013) fitted a seasonal model to the dengue incidence data of Singapore. The found a strong effect of climate on the vector density and mentions a delay of 16.8 weeks between the peaks of vector density and dengue incidence. The effect of season on the peak by the infectious vector population is more than that of the infectious human population.

The number of secondary cases during an epidemic is estimated using the seasonal reproductive numbers in Section 4.3.1. Polwiang (2015) derived the seasonal reproductive number for dengue and concluded that the variation in the amount of rainfall and temperature plays an important role in the dengue incidence. In this chapter, many variants of basic repro-

ductive number are derived that are not cited anywhere else. These reproductive numbers are useful for the estimation of disease spread in different seasonal conditions. The time evolution of the reproductive ratios is presented in Figures 4.6 and 4.7. As noticed in Figure 4.3, the temporal patterns of $R_{t|t_0}$, $R_{t|t_0}^{HV}$, and $R_{t|t_0}^{VH}$ are not affected by whether the disease is introduced by an infectious human or an infectious mosquito. The fraction of susceptible human population becomes very small soon after the outbreak, resulting in the values of $R_{t|t_0}$ and $R_{t|t_0}^{VH}$ becoming less than one. In equation 4.7, the proportions $\frac{S_h}{H}$ and $\frac{S_v}{H}$ are the only varying quantities with time. In the absence of vaccination, one possible measure is to apply insecticides after two to three months of an outbreak, as shown in Figure 4.7, to bring the value of the basic reproductive number of the vector population $R_{t|t_0}^{HV}$ below one. The results obtained form these sections are in line with the conclusions drawn by (Bartley et al., 2002; Otero and Solari, 2010; Polwiang, 2015).

The studies conducted using dengue models show that the seroprevalence level of the human population is inversely proportional to the size of the vector population to prevent or interrupt the dengue transmission (Focks et al., 2000, 1995; Newton and Reiter, 1992). Figure 4.8 shows the probability of invasion for one infectious human and mosquito for different seroprevalence levels. As seroprevalence level increases the probability of invasion becomes less which is in agreement to the work of Otero and Solari (2010). It is mentioned in Institute of Medicine (2008) that the saturation in the host population due to immune individuals makes it difficult to sustain dengue transmission. The results in Figure 4.8 also highlights this fact as there are cases where the change in seroprevalence after a certain level ($\sim 65\%$) resulted a quick decline in the probability of invasion. This level can be thought as the saturation point of the recovered individuals after which the transmission in the susceptible population gets harder. However, re-invasion in the case of dengue gives rise to complications arising in the immune system if re-infected with a different serotype (WHO, 2014).

Although the deterministic version of the model is used in this Chapter to answer different research question, some research questions are answered using the stochastic version of the model. The stochastic version of the model first is used to investigate the questions related to the probability of dengue persistence, distribution of the time to extinction of the disease and estimation of the peak of the infectious humans I_h and time taken to attain the peak. The date of arrival of an infectious individual effects the distribution of the infectious humans (Otero and Solari, 2010). Outbreaks are likely to produce larger epidemics that occur quickly if dengue introduction occurs in favourable season. If introduced in the unfavourable season, the outbreak takes longer to attain the peak value of I_h. This is because the infectious population gets a longer time to evolve before the next unfavourable season (Otero and Solari, 2010). This is the reason why there are more persistent runs for January to March,

with lower number of infectious humans that takes longer to attain the peak. If an epidemic occurs during or after October, the peak of I_h starts to decrease as the epidemic is modulated by the decrease in the vector population.

The model used in this chapter is a simplified version of seasonal dengue transmission models. This model can be adapted to other regions and parametrised for different epidemic and endemic scenario in urban environment. The model can be used to represent the host-vector system of different vector-borne diseases where seasonal conditions affect the transmission. By using the modelling framework presented, the immigration of an infectious individual and its impact of the epidemic can be discussed in a well-mixed population. The concept of disease introduction by an infectious vector is particularly important for the diseases where vectors are dispersed to or transmitted long distances. Results presented here can be used for cost-effectiveness analysis, including costs of hospitalization and planning for dengue outbreaks.

4.7.1 Caveats in the current dengue modelling setting

The parameters used in RM^s_{SEIR} are chosen from the literature, so that they provide similar patterns of dengue incidence as observed in nature. There are different biological and meteorological quantities related to the vector population and several factors related to the human population which are not included in this study. There are a number of reasons for this, including their level of impact on the population and difficulties in obtaining information on these factors from the field or the laboratory. Some of the parameters are either difficult to incorporate or not the main focus of the current study. In addition, there is a trade-off between decreasing the complexity of the model and the amount of data required to parametrize the model. Some of the factors which are not included in the current study are listed below:

- The population of vectors is affected by a change in three meteorological parameters namely temperature, rainfall and relative humidity. It is assumed that the annual change in these is synchronous. In addition, the effects of extreme values of temperature, humidity and rainfall are not taken into account.

- Temperature variation during the day is ignored. At huge population levels and the rates that have a time span of days, this simplification is not expected to introduce important distortions into the results (Focks et al., 1993b).

- Optimum humidity levels are assumed in the model. Higher humidity levels increase the chances of survival of the eggs and the adult population of mosquitoes. This can

lead to variable total population size of vectors across different years.

- Certain biological factors are ignored including: larval predation, effect of wind speed on bite rate, host preference, multiple feeding behaviour of *Aedes aegypti*, competition between *Aedes aegypti* and *Aedes albopictus* for resources, and availability of breeding sites and spatial distribution of vectors. While other factors such as photo-period and wind speed may affect the survival and developmental rates of dengue vectors, their effects are irrelevant in the current settings (Otero et al., 2008; Service, 1980). Since *Aedes aegypti* and *Aedes albopictus* have adapted to live in the close proximity of humans, they mostly use sheltered artificial and natural sites for their activities. As the effect of air turbulence and flow direction are usually more important in the day time they can effect *Aedes* mosquitoes. In general, the effect of wind is highlighted by Service (1980) where it is concluded that the day time biting mosquitoes are carried away long distances due to wind currents. This can have a potential impact on the transmission of the disease.

- In the host population the effect of urbanization, socio-economic factors, spatial distribution of hosts, migration and commuting patterns of humans and time spent outdoors are not considered.

- A single serotype model for dengue transmission is used, therefore factors related to the immunity of hosts like waning cross-immunity after infection from one serotype and antibody dependent enhancement (ADE) are irrelevant in the current study.

CHAPTER 5

Direct transmission models to represent host-vector systems

Direct transmission models to represent host-vector systems

5.1 Introduction

Understanding the mechanisms of pathogen spread and persistence in a population requires good quality data and sophisticated mathematical models. These mechanisms are even harder to uncover in vector-borne diseases mainly due to the different levels of complexity involved in disease transmission. For instance, the differences between the life span and pathogen incubation period in the host(s) and the vector(s), pathogen clearance rate in both populations and seasonal fluctuation in the size of the vector and the host population(s) varies the transmission dynamics. The magnitude of this variation is sometimes so large that the disease transmission follows inter- and intra-annual cycles in both populations. In addition, data required for parameters related to vector populations is typically scarce, influenced by the micro-environment and require a long-term investment for conducting longitudinal studies.

Vector-borne diseases create a significant annual health burden and contribute significantly to emerging diseases (Institute of Medicine, 2008). In particular, mosquito-borne diseases infect a substantial proportion of the human population, especially in tropical and sub-tropical regions of the world (WHO, 2014). A review of the mechanistic models used for mosquito-borne pathogen transmission is provided by Reiner et al. (2013) whereas the theory of mosquito-borne pathogens is recently reviewed by Smith et al. (2014). In both of these articles the evolution of mathematical modelling and theory related to mosquito-borne pathogens over four decades i.e., from 1970 to 2010 is presented. Different aspects of the mosquito life cycle and transmission are considered; however, it is interesting to note that most of the models are extensions of the Ross Macdonald mathematical framework.

The time evolution of all biological processes is influenced by stochastic events. These stochastic events are the parts of the dynamics that are not included or predicted by using the deterministic models (Ditlevsen and Samson, 2013). In contrast to a deterministic model, a

stochastic model can generate more realistic output trajectories and encompasses a broader spectrum of possible scenarios of disease spread. This is why mathematical models that are able to investigate persistence are mainly stochastic in nature since the real patterns of endemic fade-out and extinction of the pathogen are not uncovered using the deterministic modelling framework. However, one of the drawbacks of using stochastic models is that they are computationally expensive, although different approximations have been developed to increase their computational efficiency (e.g. see Cao et al., 2006, 2007, for details).

In this chapter, a simple SIR model RM_{SIR}, here termed the 'Full model' is considered with two different approximations to represent the vector disease dynamics. These models are (i) the Vector Proxy model (VP), and (ii) the Reservoir model. These models capture the effect of the vector population using two different approaches. In the first model, the latent class L_h acts as a proxy for the effect of infectious vectors in the host population: the number of humans in L_h class can be regarded as "pre-bitten" hosts. The Reservoir model contains a pool P of infection from which the infection is transmitted to the host population. The pool P can be thought of as a place for the survival of the infectious agents and a source for the transmission of the infectious particles. Because, in earlier chapters, infection in hosts and vectors followed a paired distribution with a strong and positive relationship, the number of infectious individuals in P is proportional to the number of infectious individuals in the host population. Also for simplicity, the birth and death rates of the host population are kept the same in both models so that the host population size stays the same over time. The introduction of a latent class and Reservoir to represent the dynamics of the vectors, make these models computationally efficient and analytically more tractable. In this chapter, the emphasis is put upon the validation of the approximations used in constructing simple models and the comparison of the stochastic versions of these systems with the Full model.

Modelling using these approximations is not an entirely new idea. The concept of having a population class in hosts that represent the dynamics of infectious vectors is first attempted by Dye and Williams (1995). In addition to tracking infection in the mosquito population, they investigated whether there is a need to include the population dynamics of the vector in modelling indirectly transmitted diseases. As a very small proportion of the mosquito population takes part in the process of infection transmission, they suggested that the depletion of susceptible vectors can be ignored and the vector population can be considered at steady state throughout the time-scale of simulations. Pandey et al. (2013) implicitly modelled the effect of infectious vectors in the transmission term by comparing a simple SIR model for hosts to a vector-host model. They found that the SIR model was substantially better than the vectorhost model for the DHF data from Thailand. In both of these studies, simpler models are proven to be either better or similar to the models representing the full dynamics

of the system.

The population dynamics of hosts can be highly influenced by the pathogens and mathematical models are developed to investigate their effects. Similar to the Reservoir model, Anderson and May (1981) alter a basic SIR model to incorporate the 'free-living' state for directly transmitted parasites. This is achieved by adding an equation for the population of free- living infective stages, W, assuming that an infected individual produces these stages at some rate λ , and a susceptible individual becomes infected at rate vW. As an application, they modelled the dynamics of larch budmoth, Zeiraphera diniana, and its infection with a granulosis virus. The model, although simple, was sufficient to account at least for most long-term population cycles in forest insects. Boots (1999) extended the above model by reducing the uptake of pathogen free-living infective particles by infected individuals after the pathogen infection. They found that the reduced equilibrium population density and the cyclic dynamics in the host population are as a result of regulation by the pathogen. In most of these studies, the main concern was the effect of pathogens on the host population dynamics therefore the assertion was put on the main sub-classes of the host-vector system. These approximations worked well in above studies.

Although more complex models allow us to see the structure more clearly, adding complexity in the model has some trade-offs. For instance, the complexity of model dynamics increases with multiple populations and hence it becomes difficult to understand the mechanism of persistence in these systems. This suggests that simplified models could be useful. In addition to this benefit, the approximations taken in this chapter are interesting in many aspects. First, model structures like that of the VP model are only developed for SI and SIR systems (Dye and Williams, 1995; Pandey et al., 2013). It would be interesting to develop an approximate model by incorporating the effect of infectious vectors implicitly in the latent class L_h of hosts than in infectious hosts class I_h. In the reservoir model, the free-living modelling structure is altered to account for the indirectly transmitted diseases, where rate of change in the individuals of the infectious host population I_h is connected to the free-living agents in the pool P. In the context of epidemiology and population ecology, these two approximate models are entirely different and comparing these models on the basis of a host-vector model is interesting in its own right.

The chapter is divided into two parts: (i) the first part deals with the deterministic analysis and the analytical derivation of key terms for the models. It starts with the ordinary differential equation (ODE) representation of the models including analytical derivation for the unknown quantities of the approximate models. The derivation of these terms is done to compare both models to the Full model. The basic reproductive ratios and the stability

analysis of the models at the disease free and endemic equilibrium are also presented in this section. Important mathematical derivations are provided in the appendix of the chapter. (ii) The analysis undertaken in the second part is concentrated upon the stochastic treatment of the models. A section is devoted to the identification of the regions where the model approximation fails to hold. The results obtained from the approximate models are compared with the Full model. In particular, both VP and Reservoir models are compared with the Full model by using three different criteria: (a) By comparing the stochastic trajectories; (b) by comparing the persistence threshold (CCS); and (c) by comparing the Quasi-Stationary Distribution (QSD).

5.2 Description and deterministic analysis of the models

This section comprises of the first part of the analysis performed in this chapter. A brief description of the Full model followed by the stability analysis of both the VP and Reservoir model is presented to complete the formal deterministic analysis. The derivation of main terms that arise by the approximation of parameters is also provided with the description of the models. In the end, the endemic equilibrium state of all models is compared and expressions are presented for the host compartments in all models.

5.2.1 Full model

The Ross Macdonald compartmental framework with immunity, RM_{SIR} is referred in this study as the 'Full model'. This model is made up of a Susceptible-Infected-Recovered ($S\ I\ R$) system of equations for host population dynamics and a Susceptible-Infected ($S\ I$) system for the mosquito population. This model has been extensively studied in this thesis and further details of the model can be found in section 2.2. Here, the equations of the deterministic version of the model are provided.

$$\frac{\mathrm{d}S_h}{\mathrm{d}t} = \Lambda H - \alpha I_v \left(\frac{S_h}{H}\right) - \gamma S_h$$

$$\frac{\mathrm{d}I_h}{\mathrm{d}t} = \alpha I_v \left(\frac{S_h}{H}\right) - (\xi + \gamma)I_h$$

$$\frac{\mathrm{d}R_h}{\mathrm{d}t} = \xi I_h - \gamma R_h \tag{5.1}$$

$$\frac{\mathrm{d}S_v}{\mathrm{d}t} = \delta V - \beta S_v \left(\frac{I_h}{H}\right) - \delta S_v$$

$$\frac{\mathrm{d}I_v}{\mathrm{d}t} = \beta S_v \left(\frac{I_h}{H}\right) - \delta I_v$$

5.2.2 Vector proxy (VP) model

The first of the two approximate models, the Vector proxy (VP) model uses the idea that a sub-population of hosts can be identified that represents the impact of the vectors on the host population. This sub-population acts as a proxy or an indirect representation of the vector population and transforms the indirectly transmitted host-vector system into a directly transmitted system. This is achieved by setting up a class of individuals L_h in the host population which consists of the hosts that are infected as a result of incurring an infectious bite from a vector. The hosts in this compartment are regarded as "pre-bitten" individuals. Assuming a latent class L_h introduces a delay for susceptible hosts to become infectious. This delay can be thought of as the time required for the infection to transmit to the hosts as every bite from an infectious mosquito is not assumed to be an infectious bite.

Mathematically, in addition to the SIR compartments for the host population, the VP model involves a latent class. The differential equations for this model are written as follows:

$$\frac{\mathrm{d}S_h}{\mathrm{d}t} = \Lambda H - b_{vp} I_h \left(\frac{S_h}{H}\right) - \gamma S_h$$

$$\frac{\mathrm{d}L_h}{\mathrm{d}t} = b_{vp} I_h \left(\frac{S_h}{H}\right) - (\overline{\sigma} + \gamma)L_h$$

$$\frac{\mathrm{d}I_h}{\mathrm{d}t} = \overline{\sigma} L_h - (\xi + \gamma)I_h \tag{5.2}$$

$$\frac{\mathrm{d}R_h}{\mathrm{d}t} = \xi I_h - \gamma R_h.$$

Here $\Lambda = \gamma$ denotes the birth and death rate of hosts, $\overline{\sigma}$ denotes the duration which hosts spend in the latent class and ξ denotes the recovery rates of hosts. The rate b_{vp} is the transition coefficient from susceptible compartment S_h to the latent compartment L_h. The latent class, when compared with the Full model, can be thought of a compartment that consists of 'pre-bitten' humans. The flow of infected individuals to the infectious compartment

takes place after on average $\overline{\sigma}^{-1}$ days. Due to the low mortality rate of hosts, most of the individuals in the L_h class will become infectious. The term b_{vp} is multiplied by the term $\frac{S_h}{H}$ showing frequency dependent transmission.

In order to parametrise the model for the stochastic simulations, values of the unknown parameters b_p and $\overline{\sigma}$ are required. The Full model has no latent class in hosts, therefore it cannot be used for finding the analytical form of $\overline{\sigma}$. For this purpose a model called $RM_{SEIR-SI}$ is constructed that includes a latent class in the host population. The term $\overline{\sigma}$ is derived by setting the endemic equilibrium of the VP model to that of the $RM_{SEIR-SI}$. In a similar fashion, the value of b_{vp} is obtained when comparing the equilibrium point of the VP model with $RM_{SEIR-SI}$. The same value of b_{vp} is obtained by comparing the equilibrium point of the VP model with the Full model.

The value of the unknown parameters b_vp and $\overline{\sigma}$ in terms of the parameters of the host-vector models are (see Appenidx D.3.1).

$$b_{vp} \approx \frac{\alpha\beta}{\delta}\left(\frac{V}{H}\right) \qquad\qquad \overline{\sigma} \approx \sigma. \tag{5.3}$$

In comparison with RM_{SIR}, the terms b_{vp} and $\overline{\sigma}$ are the composite host-to-host transmission and latency rate. The term b_vp is enhanced by the product of transmission rates in the $RM_{SEIR-SI}$ model ($\alpha\beta$) times the number of vectors per host ($\frac{V}{H}$) present in the system and transmission in the VP model is affected by the longevity of the vector (δ). It implicitly represents the effect of the vector population rather than adding a separate set of equations for vector population. The second unknown parameter $\overline{\sigma}$ contains one additional term $\frac{\sigma}{b_{vp}}$, as the delay in VP model is proportional to the average latent period in hosts in $RM_{SEIR-SI}$ and is hindered by the composite host-to-host transmission. By replacing the value of b_{vp} in the approximation of $\overline{\sigma}$ in equation 5.3, the average latent period remains unaffected in VP model and $\overline{\sigma} = \sigma$.

The basic reproductive number R_0 is found in Appendix D.1.1:

$$R_0 = \frac{b_{vp}\overline{\sigma}}{(\xi + \gamma)(\overline{\sigma} + \gamma)}. \tag{5.4}$$

There are two equilibrium points: the Disease Free Equilibrium (DFE) and the Endemic Equilibrium (EE). The details of stability analysis of this model are provided in Appendix

D.2.1. Standard mathematical analysis yields the following endemic equilibrium points.

$$S_h^* = \frac{H(\xi + \gamma)(\overline{\sigma} + \gamma)}{b_{vp}\overline{\sigma}} = \frac{H}{R_0}$$

$$I_h^* = \frac{H\gamma\overline{\sigma}}{(\xi + \gamma)(\overline{\sigma} + \gamma)} - \frac{H\gamma}{b_{vp}} = \frac{\gamma H}{b_{vp}}(R_0 - 1)$$

$$R_h^* = \frac{H\xi\gamma\overline{\sigma}}{\gamma(\xi + \gamma)(\overline{\sigma} + \gamma)} - \frac{H\xi}{b_{vp}} = \frac{H\xi}{b_{vp}}(R_0 - 1) \qquad (5.5)$$

$$L_h^* = \frac{-H\gamma(\xi + \gamma)}{b_{vp}\overline{\sigma}(\overline{\sigma} + \gamma)}\left(\gamma + H\left(\overline{\sigma} - R_0\left(\overline{\sigma} + \gamma\right)\right)\right).$$

These equilibrium points are comparable with the equilibrium points of the Full model. In chapter 2, the equilibrium points of hosts in RM_{SIR} are derived as:

$$S_h^* = \frac{H(\gamma H R_0 + \alpha V)}{R_0(\alpha V + \gamma H)}$$

$$I_h^* = \frac{\delta\gamma H^2(R_0 - 1)}{\beta(\alpha V + \gamma H)} \qquad (5.6)$$

$$R_h^* = \frac{\xi\delta H^2(R_0 - 1)}{\beta(\alpha V + \gamma H)}.$$

The equilibrium points for the vector population are:

$$S_v^* = \frac{V(\alpha V + \gamma H)}{(\alpha V + \gamma H R_0)}$$

$$I_v^* = \frac{\delta\gamma H(\xi + \gamma)(R_0 - 1)}{\alpha\left(\delta(\xi + \gamma) + \beta\gamma\right)}. \qquad (5.7)$$

The term γH has very little contribution (0.007%) in the expression $\alpha V + \gamma H$ for the parameter values employed. As a result, this term can be dropped in the expression $(\gamma H R_0 + \alpha V)$ and $(\gamma H + \alpha V)$ in the above equation set 5.6 to obtain the equilibrium points of the VP model.

5.2.3 Reservoir model

The second alternative model uses the concept of a Reservoir by considering a pool P of infection. The 'individuals' in the pool give birth with a rate $c = kq\ (V/H) = \beta V/H$ multiplied by the number of infectious hosts. Here c is the per human rate of generating infectious vectors. The bite rate per vector is denoted by k and q denotes the probability of transmission of infection from a bite. The vector to host ratio is V/H. The term cI_h is defined as the mean number of bites that lead to an infectious vector for the entire human population and $\overline{\delta}$ denotes the vector mortality rate. The vector population is assumed constant and the depletion of susceptible individuals is ignored and so $S_v \approx V$.

The equations of the Reservoir model are as follows:

$$
\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= \Lambda H - b_r P\left(\frac{S_h}{H}\right) - \gamma S_h \\
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= b_r P\left(\frac{S_h}{H}\right) - (\xi + \gamma)I_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \xi I_h - \gamma R_h \\
\frac{\mathrm{d}P}{\mathrm{d}t} &= cI_h - \bar{\delta}P
\end{aligned}
\tag{5.8}
$$

By comparing the endemic equilibrium of RM_{SIR} and Reservoir model (derivation is in Appendix D.3.2), we obtain:

$$
b_r \approx \alpha, \qquad \bar{\delta} \approx \delta.
\tag{5.9}
$$

The basic reproductive ratio is found as

$$
R_0 = \frac{b_r\, c}{\bar{\delta}(\xi + \gamma)}.
\tag{5.10}
$$

The endemic equilibrium points are:

$$
\begin{aligned}
S_h^* &= \frac{\bar{\delta}H(\xi + \gamma)}{b_r c} = \frac{H}{R_0} \\
I_h^* &= \frac{H\gamma}{\xi + \gamma} - \frac{\bar{\delta}\gamma H}{c\, b_r} = \frac{\bar{\delta}\gamma H(R_0 - 1)}{c\, b_r} \\
R_h^* &= \frac{\xi H}{\xi + \gamma} - \frac{\xi\bar{\delta}H}{c\, b_r} = \frac{\xi\bar{\delta}H(R_0 - 1)}{c\, b_r} \\
P^* &= \frac{H\gamma c}{\bar{\delta}(\xi + \gamma)} - \frac{\gamma H}{b_r} = \frac{\gamma H(R_0 - 1)}{b_r}.
\end{aligned}
\tag{5.11}
$$

The mathematical analysis of the models is given in Appendices D.1.2 and D.2.2. By putting the values of c, $\bar{\delta}$ and b_r, and since $\gamma H \ll 1$ when it is added to a large quantity, the equilibrium points of the Reservoir model are deductible from the equilibrium points of the Full model. The comparison of the equilibrium points of all models for the susceptible, infectious and recovered hosts is shown in equation 5.12, whereas the baseline parameters are shown in Table 5.1.

$$
\begin{aligned}
S_h^* &= \frac{H}{R_0} \\
I_h^* &= \frac{\delta\gamma H^2(R_0 - 1)}{\alpha\beta V} \\
R_h^* &= \frac{\xi\delta H^2(R_0 - 1)}{\alpha\beta V}.
\end{aligned}
\tag{5.12}
$$

Symbol	Explanation	Value used
α	Per-bite vector-to-host transmission rate, in days^{-1}	0.1
β	Per-bite host-to-vector transmission rate, in days^{-1}	0.075
ξ	Average infectious period of hosts, in days^{-1}	0.1428
δ	Average birth / death rate of vectors, in days^{-1}	0.125
γ	Average mortality rate of hosts, in days^{-1}	4.215×10^{-5}
b_{vp}	Composite host-to-host transmission rate, in days^{-1}	$\frac{\alpha\beta}{\delta}\left(\frac{V}{H}\right)$
$\overline{\sigma}$	Average incubation rate of hosts, in days^{-1}	$\sigma = 0.2$
c	Per human rate of generating infectious vectors, in days^{-1}	$\beta\dfrac{V}{H}$
$\overline{\delta}$	Average birth / death rate of individuals in the Reservoir, in days^{-1}	$\delta = 0.125$

Table 5.1: List of parameters used in all the models. Here, the average birth rate of hosts is equal to the death rate, i.e., $\Lambda = \gamma$. The life expectancy of a single host is set to 65 years. The first five parameters are same as used in Chapter 2 and 3. The last four entries of the table denote the derived parameters obtained by comparing the approximate models to the Full model at the endemic equilibrium.

5.3 Methods for the evaluation of the stochastic version of the models

The second part is related to the study of the stochastic version of the models. In the part, the VP and Reservoir models are compared with the Full model. The approximate models are initially compared with the Full model to identify the possible parameter values corresponding to epidemiological scenarios where the model approximation fails to hold. After that, the models are compared by plotting the stochastic trajectories of different compartments. Next, the CCS is obtained from the models. At the end, the QSD of different compartments is compared. The models are parametrized by using the baseline parameter values mentioned in Table 5.1.

5.3.1 Breakdown of approximation in both models

The models presented in Sections 5.2.2 and 5.2.3 are two distinct approximations of the Full model, RM_{SIR}. The introduction of a latent class and Reservoir to represent the dynamics of the vectors, make these models computationally efficient and analytically more tractable. However, there are cases where these approximations lead to incorrect results. This section highlights the parameter space under which the model assumptions fail to hold. After putting the values of b_{vp}, $\overline{\sigma}$, b_r and $\overline{\delta}$, the expression for R_0 derived through the equivalence of the

models at equilibrium is similar.

In the VP model, it is assumed that the vectors in the Full model can be approximated as a proxy class of latent hosts, where the effect of the vector is implicit in the transmission term b_{vp}. In other words, the class L_h contains the hosts which are almost certain of getting infection after the average period of $\frac{1}{\sigma}$ days (death rate is very low in hosts). Comparing the equations of infectious humans in VP model and RM_{SIR} at equilibrium yields the following value of L_h^*, in terms of the entities of RM_{SIR}:

$$L_h^* = \frac{\alpha I_v^*}{\sigma}\left(\frac{S_h^*}{H}\right). \tag{5.13}$$

By substituting in the baseline values of $\alpha = 0.1$, $\sigma = 0.2$ from Table 5.1 results $\frac{S_h^*}{H} \approx 0.4$. That leads to $L_h^* \approx 0.2\,I_v^*$. Equivalently the number of hosts in the latent class of the VP model is one fifth of the number of infectious vectors in the Full model. In general, the approximation at equilibrium follows $L_h^* \approx \sigma I_v^*$. This assumption is valid when both models are at the deterministic endemic equilibria. As mentioned in the above paragraph, hosts in L_h are almost guaranteed to become infected. The low number of hosts in this class as compared to I_v^* makes sense as not all vectors in I_v^* are successful in passing infection on to hosts, mainly because of their high death rate.

In order to gain a better understanding of the conditions under which the simplified models provide an acceptable approximation of the Full model, the parameter values for which the assumptions may fail are considered. The models are then solved for these parameter values showing the effects of this breakdown. For the VP model, let there be five infectious hosts and vectors present at the start of simulations in RM_{SIR}. The simulations are now started to represent an epidemic in a totally susceptible population, so in this case $S_h = H$. Inserting the values of the parameters into equation 5.13 produces $L_h \approx 0.5 I_v$ meaning that the number of hosts in the latent class in the VP model is half the number of vectors in the infectious class of RM_{SIR} when simulations are started away from the steady state. Although, this relation converges to $L_h^* \approx 0.2\,I_v^*$ when the endemic equilibrium state is attained, this difference leads to different stochastic trajectories in I_h and L_h : I_v classes of both models, implying that these assumptions are valid near the deterministic steady state (see top row of Figure 5.1).

Similarly, in the model with infectious Reservoir, it was assumed that $S_v \approx V$ rather than $V = S_v + I_v$. The infectious proportion of vectors is neglected as a very small proportion takes part in the transmission of the infection (results from previous chapters), as is shown in the *Aedes aegypti* survival curve in Brownstein et al. (2003). It is straightforward to note that
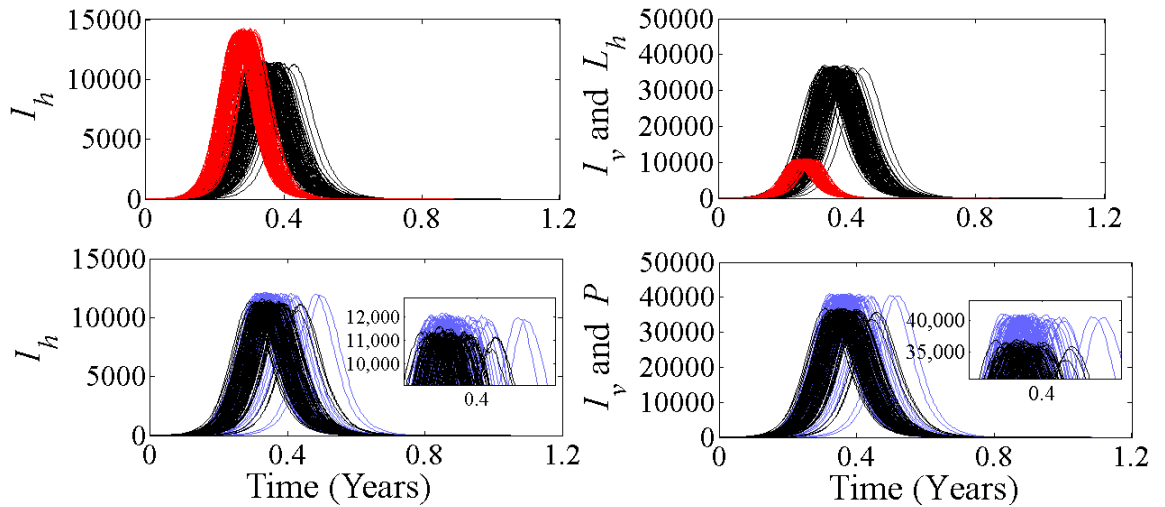
Figure 5.1: Graph showing the comparison of VP (in red) and Reservoir (in blue) models with RM_{SIR} (in black) for an epidemic scenario. The top row represents the comparison of VP model and RM_{SIR} and bottom row shows the comparison of Reservoir model and RM_{SIR}. The inset in the bottom row highlights the peak of the populations. One hundred stochastic trajectories of both models are shown. Population of hosts is 100000 individuals and vector-to-host ratio is six. The initial conditions are five infectious hosts (I_h) for all three models, whereas five infectious vectors I_v are introduced in RM_{SIR} and in Reservoir model P respectively. The value of pre-bitten hosts at equilibrium is taken from $L_h^* \approx 0.5\, I_v^*$. The value of parameters is listed in Table 5.1.

the above assumption becomes invalid if there are many infectious vectors in the population. One way to violate this assumption is to increase the rates α and β, which accelerates the transmission process and introduces a lot of infectious mosquitoes. The second way is to introduce only a few infective hosts and vectors so that an epidemic outbreak is observed. The difference between the models becomes evident by observing the number of individuals at the peaks of the infectious compartments, as shown in the bottom row of Figure 5.1.

The equation for I_v in the Full model is

$$\frac{\mathrm{d}I_v}{\mathrm{d}t} = \beta S_v \left(\frac{I_h}{H} \right) - \delta I_v. \tag{5.14}$$

By comparing this equation with the equation of pool P in section 5.8, the rate at which a new infectious vector is generated in equation 5.14 is a product of the proportion of infectious humans $\frac{I_h}{H}$ and the susceptible vectors S_v. The rate at which the population of vectors changes in P is obtained by multiplying $\frac{I_h}{H}$ by the total vector population V. As the vector population is set constant in both models, $V = S_v + I_v$. In the situation where plenty of infectious vectors are present; i.e, disease invasion, the difference in the '*generation of new*

infections' becomes evident between the two models since $V \not\approx S_v$. The bottom row of Figure 5.1 presents this scenario where the number of infectious hosts and vectors at the peak of the epidemics are different. The difference is more prominent at the peak of infectious vector and population of the pool (Bottom right of Figure 5.1) which, in turn, create more infections in host population of the Reservoir model.

5.3.2 Estimation of the CCS

The estimation of the CCS is performed in similar fashion as undertaken in Section 3.5.3. A summary of the method is as follows: the stochastic simulations are started from the deterministic equilibrium point and allowed to run for twenty-five years. Adaptive tau-leap simulations for every population size are repeated one thousand times and population size is increased until half of the stochastic simulations retain infection at the end of twenty-five years. This population size is termed as CCS and it is computed for the Full, VP and Reservoir models by initializing them using the baseline parameter values.

5.3.3 Estimation of the Quasi-Sationary Distribution (QSD)

The QSD distribution is defined by Nåsell (2005) as:

> *The so-called quasi-stationary distribution, which is a stationary distribution, conditional on non-extinction, is supported on the transient, non-absorbing states. It is a useful approximation of the state of the process when it has been going on for a long time without extinction. It is a counterpart to the endemic infection level in the deterministic model.*

Here transient states are states where the number of infected and infectious individuals is non-zero.

The quasi-stationary distribution is the distribution of states of the system modelled as Markov chains and it is conditioned on non-extinction of the disease. The Markov chain should have at least one absorbing state, for example, the extinction state of the organisms. In the context of population dynamics, a disease is considered endemic in the stochastic setting where the QSD is reached. Intuitively, QSD describes the long-run distribution of the states of the system.

For estimating the QSD, the simulations were allowed to run for 100 years. A host population of 1.5 million individuals was selected and a vector-to-host ratio of six was assumed for all models. After parametrization and finding the unknown parameters, the stochastic simulations were initialized with same initial conditions in all of the three models. The result of the simulations using the simulations incorporating Gillespie and adaptive tau-leap algorithms were then compared for estimating the accuracy of the tau-leap algorithm. The estimation of the QSD can be described as follows: assume that I is the Markov chain for a compartment of the model, and i is a particular count of the number of individuals of that state. The formula for estimating the probability of each state i is then:

$$P(I = i) = \frac{\sum \text{ length of time}(I = i)}{\sum t_{tot}} \tag{5.15}$$

Here, numerator sums the time spent in the state i in the simulation. t_{tot} denotes the running time of a simulation until the infectious populations falls to zero or the maximum time limit is reached. The summation runs to hundred years for hundred stochastic repetitions using tau-leap algorithm and thirty using the Gillespie algorithm. It was assumed that the simulations fell quickly into the QSD after starting from the deterministic endemic steady state. Each simulation was stopped if there was no infected individual in the populations.

5.4 Results from the stochastic version of the models

The second part of this Chapter deals with the analysis undertaken with the stochastic version of the models. The main idea behind this analysis is to investigate the similarity between all the models in the stochastic settings. The populations in a stochastic model follow different trajectories even if started using the same initial conditions. Therefore, comparing the dynamics of individuals in different classes in all three models is a valid starting point. The time-varying trajectories provides an initial idea of the distribution of the profiles of individuals in different compartments. Another interesting area is to estimate the Critical Community Size (CCS) for these models. This measure highlights the similarity of the models in relation to the host population sizes and disease persistence. Moreover, in the first part of the chapter the models have almost similar values of individuals at the deterministic endemic equilibrium, therefore the quasi-stationary distribution (QSD) of the different compartments of the models is compared. As this distribution is the stochastic counterpart of the endemic equilibrium points in the deterministic setting, this distribution helps investigating the behaviour of the model at endemicity in the stochastic setting. The results of these three analyses are presented in a sequential manner. At the end, a discussion about important outcomes that can be established by using these model approximations are presented.

5.4.1 Stochastic trajectory comparison
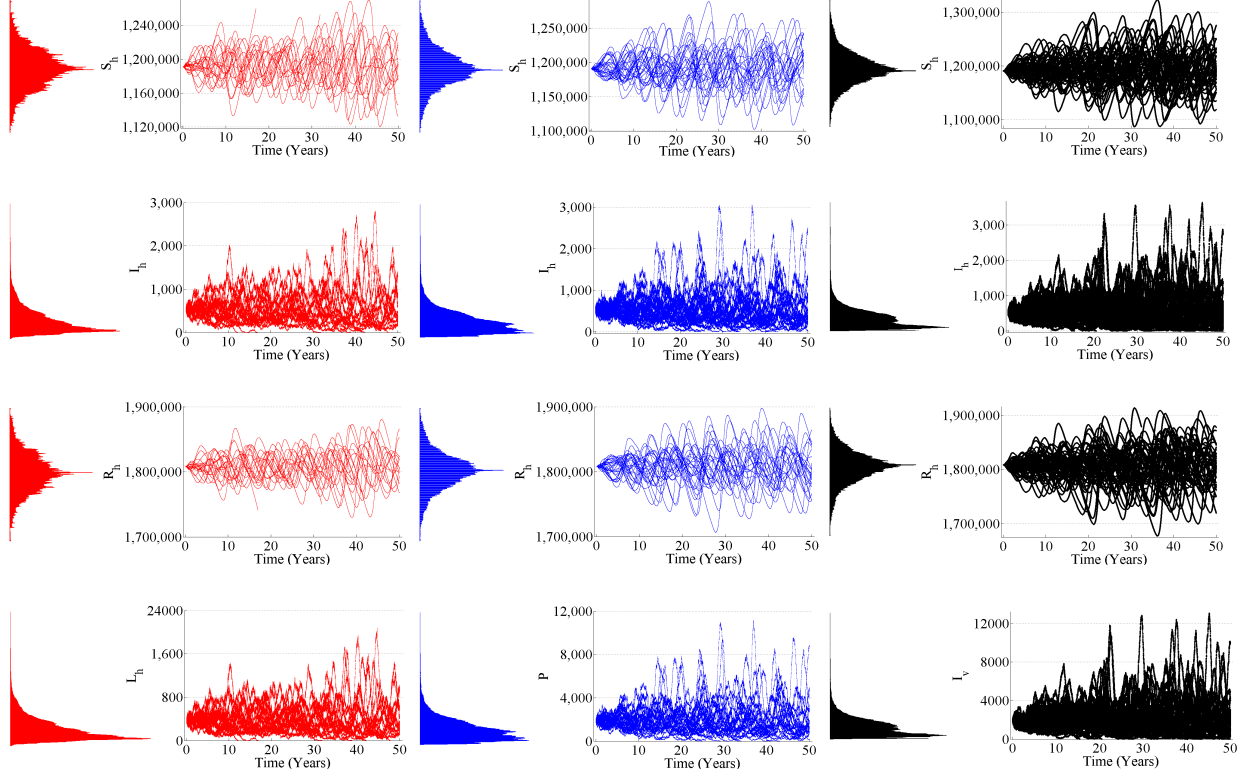
Figure 5.2: Time evolution of different population compartments in the approximated models. The VP model is plotted in red and the Reservoir model is plotted in blue colour. The number of susceptible (S_h), infectious (I_h) and Recovered (R_h) hosts are directly comparable whereas the number of individuals in P are nearly five times as in L_h. The graph showing infectious vectors I_v in the Full model has nearly the same number of individuals as the pool P in the Reservoir model. In the VP model, $L_h^* \approx 0.2\, I_v^*$, as shown in section 5.3.1. All stochastic simulations are started form the deterministic endemic equilibrium. Parameter values used in these models are listed in Table 5.1.

The first measure for comparing the models is by comparing the stochastic trajectories of the same or similar compartments. For this purpose, the stochastic simulations in all three models are run for twenty-five years starting form the deterministic steady state in each compartment. One hundred stochastic repetitions are performed for all models. Incorporating the parameter values of b_{vp} and $\overline{\sigma}$ in the VP model leads to roughly the same number of individuals in S_h^*, I_h^* and R_h^* and the value of R_0 as in RM_{SIR} at the start of the simulations. This is also true for the Reservoir model when the values for b_r and $\overline{\delta}$ are incorporated. The host population size at the start is 3 million individuals and the vector-to-host ratio is set to six. Simulations are run for twenty five years and their trajectories are plotted in Figure 5.2. The compartments of the Full model are plotted in black, whereas the compartments of VP and Reservoir models are

shown in red and blue respectively. The population of susceptible and recovered individuals in the VP model showed less dispersion around the deterministic endemic equilibrium as compared to the same classes in RM_{SIR}.

5.4.2 CCS comparison

The CCS is regarded as the population size above which a pathogen can persist in the population without introduction from external source. It is a measure that is similar to a persistence threshold of a stochastic population model. As the Full model is being approximated by VP and Reservoir models, it is worthwhile to examine the similarity of their CCS with the Full model. In this section the CCS of both approximation models is found and compared with RM_{SIR}.
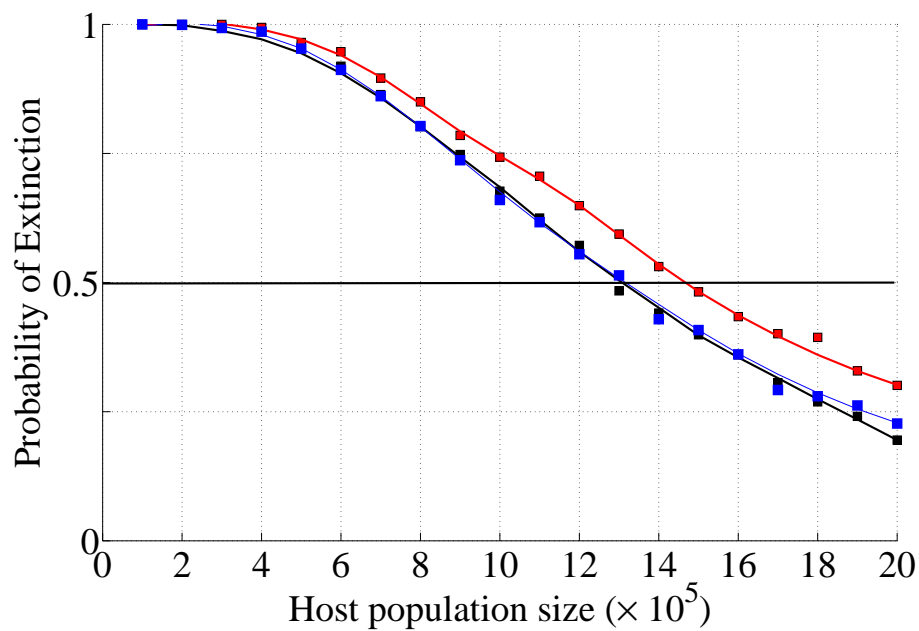
Figure 5.3: Comparison of CCS for RM_{SIR} (black), VP (red) and Reservoir model (blue). The x-axis represent the host population size in 10,000 and y-axis represents the probability of extinction. CCS is attained when half of stochastic repetitions contained infection at the end of twenty five years.

The CCS for RM_{SIR} was found to be ~ 1.3 million hosts in Chapter 3. Model parametrization resulted in approximately the same initial conditions for these models as the Full model. The comparison of the CCS for these models is shown in Figure 5.3. The Reservoir model followed the same path as the Full model (CCS ~ 1.3 million hosts) but a marked difference is observed

in probability values in the VP model especially when $H > 0.3$ million, where CCS for the VP model is found to be ~ 1.5 million hosts. This difference mainly arises due to the low number of individuals in the L_h compartment, as shown in Figure 5.2. At the endemic equilibrium $L_h^* \approx 0.2\, I_v^*$; the infection in the VP model is more prone to the demographic stochasticity at low numbers of individuals in I_h and L_h classes.

5.4.3   Comparison on the basis of QSD of the models

This section provides a final equivalence measure by comparing the Quasi-stationary distribution (QSD) of the models. The QSD is compared for the number of individuals in susceptible, infectious and recovered compartments of hosts in all models. The QSD of the infectious hosts I_h is estimated by using both the tau-leap and Gillespie algorithms, whereas the QSD for susceptible and recovered hosts is found by using the tau-leap algorithm. When Gillespie algorithm is employed for the Full model, the birth and death of a vector is represented as a single event, i.e, the death of an infectious vector results in an immediate birth of a susceptible individual. With the exception of a few runs (<3 in all models), infection in all of the simulations went extinct within the time frame of one hundred years.

The comparison of adaptive tau-leap (top) and Gillespie algorithms (bottom) for infectious humans (I_h) is shown in Figure 5.4. The x-axis represents the states and y-axis represent the probability in each state. The QSD for infectious humans (I_h) was obtained by using equation 5.15. The QSD shows high probability for states corresponding to low numbers of infectious humans. This high probability of having a small number of infectious individuals is an artefact of using the tau-leap method for the VP model. In the bottom figure, the comparison of the QSD for I_h in the VP model obtained by employing the Gillespie algorithm does not show this behaviour. This suggests that the tau-leap method is not a right choice for investigating the behaviour of the system near the absorbing state, for models that are similar to the VP model.

In the bottom figure where the Gillespie algorithm was used, the shape of the QSD of infectious individuals in the VP model is better. There is a bit of variation at the mode of the distribution in the bottom figure because of fewer stochastic repetitions; thirty as compared to a hundred for the tau-leap method. An important finding is that the tau-leap method shows roughly the same level of accuracy as the Gillespie algorithm for the Full and the Reservoir model, in addition to being much faster in implementation to the models. The QSD for the susceptible and recovered individuals is plotted in Figure 5.5. The distributions of both compartments for the RM_{SIR} model has slightly lower values near the peak of the
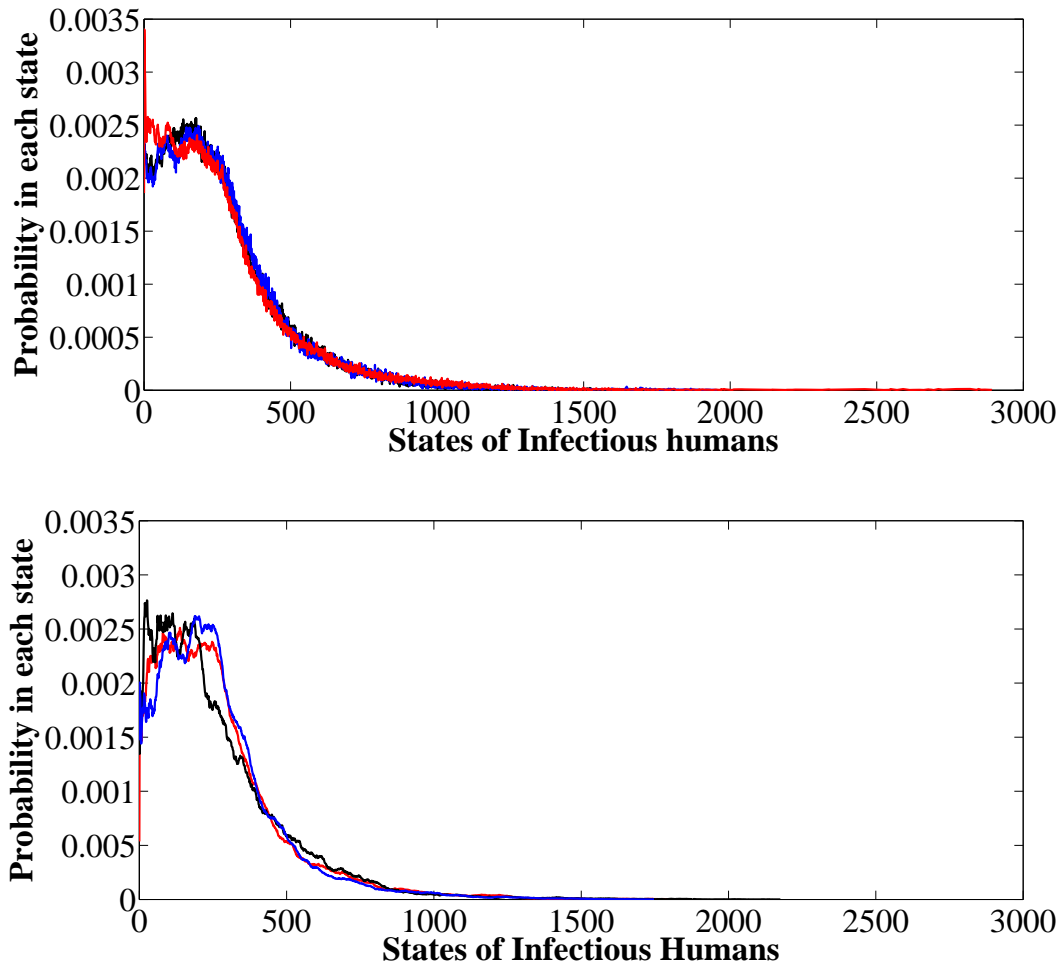
Figure 5.4: Comparison of the QSD for infectious humans (I_h) in all three models. Top: Comparison of the QSD estimated using adaptive tau-leap. RM_{SIR} in black, VP model in red and Reservoir model in blue. Bottom: Estimation of QSD using Gillespie algorithm. The approximations show more variability in the bottom figure due to fewer stochastic repetitions of the models.

distribution. There is a cyclic pattern at the tail of the distribution of both classes. This pattern is may be due to the way the critical events are handled using the tau-leap algorithm as in this study they are not done as described in Cao et al. (2007). This also explain why CCS is higher for the VP model as shown in Figure 5.3– it is not that VP predicts CCS incorrectly, but that the tau-leaping approximation with the VP model does not work.
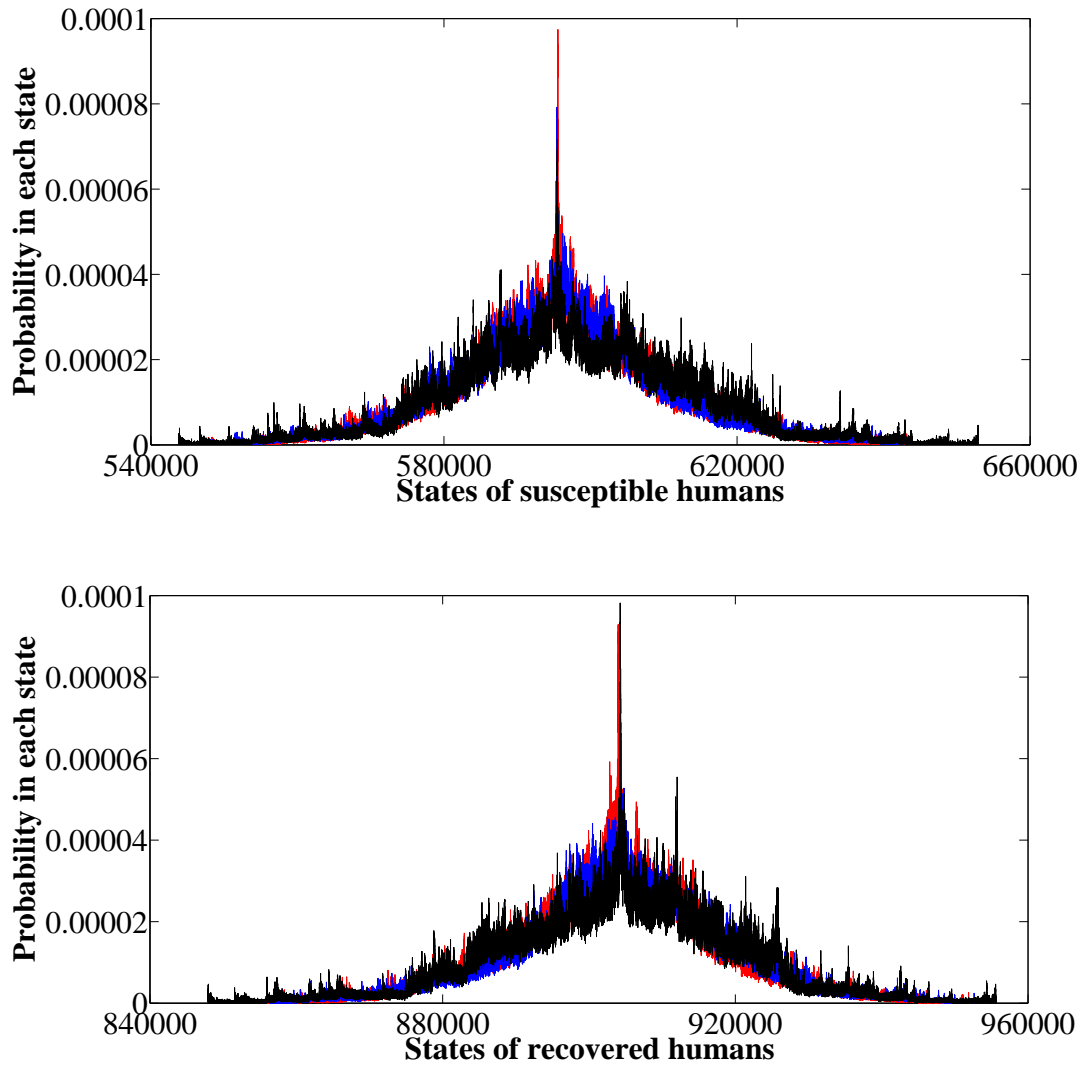
Figure 5.5: Comparison of the QSD for the susceptible and recovered hosts in all three models using the tau-leap method. The Host population comprises of one million individuals. All three plots in susceptible and recovered compartments followed a similar profile, except for some small deviations between the Full and approximate models. Here the colour scheme is same as in Figure 5.4.

5.5   Conclusion and discussion

The mathematical formulation of any biological process can rapidly become complicated with the inclusion of factors like the number of species and reactions involved in the modelling framework. This additional structural information can impair the benefits and advantages of using the mathematical tools to explore the behaviour of the system. The attractiveness of simplifying the host-vector system comes from the potential gains in computation time required for the understanding of the biological process. It allows for the fact that the data for the complex vector dynamics are also limited. A simple but powerful approximation may provide insights into the persistence dynamics of more complex systems (*e.g.*, multi-host and multi-pathogen) which are currently not easy to investigate using mathematical models that represents the complete detail of the process. Moreover, the lesser the uncertainty related to the data and the number of parameters, the better the results from the mathematical model. The model with fewer parameters is strongly selected by the AIC and other model selection criteria (Pandey et al., 2013). The authors further found that the simple SIR model fits the dengue incidence data of Thailand as well as the host-vector model and conclude that the inclusion of vector population may not be necessary to model prevalence and incidence in human or other primary host. Furthermore, at the very least, a thoughtful simple mathematical framework can provide an idea of the behaviour of the system without going into every detail of the process.

It is anticipated that the analysis done in this chapter is not a well-studied area in host-vector systems and very few studies (*e.g.* Dye and Williams (1995); Pandey et al. (2013)) address the problems of this type. The current chapter is an attempt to represent the dynamics of a non-seasonal host-vector model in two ways. First a simple host population based model, termed as the VP model, where the effect of the vector population is represented as a latent class of humans. The hosts enter this compartment at a composite host-to-host transmission rate b_{vp} and leave at the rate $\bar{\sigma}$. These terms are found by comparing the models at the deterministic equilibrium. Second way is by reducing the dimension of the host-vector system, i.e, the Reservoir model. The dimensional reduction is undertaken by assuming that the susceptible population of vectors is fixed at the deterministic endemic equilibrium. The composite terms b_r and $\bar{\delta}$ are found by using the same procedure as undertaken for the VP model. These assumptions create a considerable change in the time required for the processing of the stochastic version of the models ($\approx$ five to ten times) although the model contains a large number of individuals. In this study, both of the approximate models are investigated for the parameter regions where the approximation fails to hold and then compared with the Full model by using three metrics.

By comparing the results from the models, it can be seen that the approximations considered here work well if the behaviour of the system is studied around or near the endemic state (see Figure 5.3.1). The VP model is sensitive to the low population numbers, as there are few infected individuals in L_h at the start of the simulations when the system is at the deterministic endemic equilibrium (see Figure 5.2 for the number of individuals in L_h) that resulted in a bigger host population to attain the CCS. The Reservoir model seems better in handling this issue as seen from Figure 5.3. With large enough and well-mixed host and vector populations that overcome the effects of demographic stochasticity these model approximations are good enough to be incorporated in a range of host-vector systems. The exclusion of seasonal effects in these approximate models means that they are suitable for short-term analysis of the system so that the population size of vectors is not considerably changed due to seasonal fluctuations. However, both models are well-suited to investigate the temporal trends and basic epidemiological features of diseases like dengue, yellow fever and malaria.

In Section 5.4.3, the QSD is approximated by using two different algorithms and results are compared for the infectious hosts in each model. In Figures 5.4 and 5.5, the comparison is shown. The tau-leap method resulted a close approximation of the QSD of the Full model and Reservoir model. For the VP model, it does not provide a good approximation at lower states of I_h. However, it does better at the peak and the later states and at the tail of the distribution. The similarity of the mode and tail of the distributions in all three models indicate that the "stationary state" of the models is comparable in the stochastic setting. Although the mean time to extinction can be estimated when the system is at Quasi-stationary state, the emphasis here is to compare the models at the "stochastic endemic equilibrium state" since the number of individuals in every compartment is roughly the same for all three models at the deterministic endemic state.

In addition to the computational benefits, these approximations are used to investigate to what degree the inclusion of the vector population in the dynamics of the system is important. This question can be of a huge importance as the data related to vectors is usually not very easily available. It requires time and costs to conduct longitudinal field studies for obtaining data. In contrast, the human cases are very well monitored. As quoted in Dye and Williams (1995) for flea-borne diseases:

> *Interestingly, all species in general exhibited stronger correlations with host dynamics than those of their vectors, supporting the assertion that flea-borne microparasites can often be incorporated effectively into epidemiological models as directly-transmitted pathogens.*

In general, the inclusion of the vector population having a very high birth and death rate as compared to hosts results two very different time scales of demographic turn-over. As a very small fraction of the vector population takes part in the process of disease transmission, it can be assumed without the loss of generality that there is no depletion in the number of susceptible vectors present in the population. The structure of both VP and Reservoir model excludes the population dynamics of susceptible vectors and assumes that they are at a stable equilibrium value. This has two major benefits: (i) as the dimension of the system is reduced so that the analytical treatment of the approximate models is easier (Dye and Williams, 1995), (ii) the models are compared using stochastic individual-based algorithms, which can be extremely slow if the birth and death of the vectors are included as individual events. As a result, the estimation of QSD and CCS becomes very time consuming in that case.

While all the models of mosquito-borne pathogens address some biological question of interest, there is no clear justification of which is the most appropriate model. In epidemiology, approximate models are intended to obtain expressions for epidemiological parameters that can be used to determine the observed patterns of susceptibility and infection of diseases in different countries. The analysis done in this chapter surround both deterministic and stochastic modelling frameworks and the conditions of epidemics and endemicity of the disease. The main advantage of these approximations is computational ease, as the VP and Reservoir model were faster to simulate, even using the Gillespie algorithm. As a result, those stochastic simulation experiments can be used to gain insight to the mechanisms at work which are complex or time and memory intensive by using the Full model. Another main contribution of chapter is the rigorous treatment of the approximate models since they are compared by using very important metrics, the estimation of CCS and QSD. Depending on the nature of the research question, many interesting questions related to the mechanism of disease transmission can still be answered by using the approximations used, mainly because they are proven mathematically sound for further analysis. Possible extensions of the approach taken in this chapter can be incorporating the meta-population structure, extension to a spatial model or including age structure and using method of stages for the infectious period distribution for the host-vector systems.

It is important to note that the model approximations here are valid for modelling the dynamics of a vector-borne disease without seasonal fluctuations. Further, critical events are not handled as defined by Cao et al. (2007), that suggests switching to the Gillespie algorithm at low number of individuals in any compartment. By definition, an event is critical if happening fewer than 'n' times would make a population negative. Depending on the model, the value of n is usually set to 5 or 10. It is interesting to note that this does not

produce a huge difference in the results from Reservoir and Full models as shown in Figure 5.4. Another important point from this figure is that the tau-leap method without handling the critical events is not suitable for the VP model. So if the mathematical structure of the model is similar to that of the VP model, care should be taken if the stochastic system is modelled using an approximate method to investigate the behaviour of the population compartments that are far from the equilibrium. With small enough population sizes, it may will be that tau-leap method would fail for all three models without properly handling the critical events.

CHAPTER 6

Thesis overview and conclusion

Thesis overview and conclusion

The main aim of this thesis was to develop our understanding of pathogen persistence in host-vector systems by using efficient modelling paradigms. This was achieved by conducting three studies that were used to investigate the following: (i) The relationship between the long-term pathogen persistence in host-vector models and the parameter values if the model; (ii) the effects of seasonal variation on the introduction and persistence of disease in a host-vector system; (iii) the use of approximate models to investigate the dynamics of the full host-vector system. An overview is now provided of each of the chapters.

6.1   Overview

The objective of Chapter 3 was to establish to what extent persistence of a pathogen in a host population, measured by the Critical Community Size (CCS), depends on the parameters of host-vector system. For this purpose, two different models, Ross Macdonald model with host immunity (RM_{SIR}) and Ross Macdonald model with latent periods and immunity (RM_{SEIR}) are used. The deterministic behaviour and the stability of these models was investigated in Chapter 2 and the stochastic version of these models is used to understand the mechanism of pathogen persistence in this chapter.

To measure the probability of disease extinction in human and vector populations, individual-based models were simulated using tau-leaping approximations of the Gillespie Algorithm. These were used to find the CCS for both models, parametrised for dengue by keeping same basic reproductive ration R_0 in both models. CCS for the host population was found to be 1.3 million hosts for the baseline model RM_{SIR} and was reduced to less than half (0.6 million hosts) in RM_{SEIR} which shows that the inclusion of latent periods have a dramatic impact on the persistence of dengue virus in hosts. For the vector population, the CCS was found to be 7.8 and 6.6 million for the two models respectively.

To further investigate the association between parameters of the models and CCS, sensitivity analysis was performed and general linear models were used to quantify the relationship

between CCS and its determinants. The parameters of the models were divided into primary and secondary (the algebraic combination of primary parameters). The average infection clearance rate of hosts (denoted by ξ) and the average birth / death rate of vectors (this rate is denoted by same parameter δ; once infectious vectors are assumed to remain infectious for the rest of their lives) are found to be the most significant primary predictors for CCS in both models, followed by infection transmission rate in hosts (α) and vectors (β). Moreover, in modelling CCS using RM_{SEIR}, the average rate of latency in vectors (ρ) is found to be an important predictor for CCS whereas host latent period (σ) has very low effect. In both models, the average birth and death rate of hosts (γ) has least effect on CCS, as the length of the simulations time was $\sim 40\%$ of $\frac{1}{\gamma}$. Modelling using the secondary parameters yielded three main determinants for CCS in dengue, viz., (i) the basic reproductive number R_0, (ii) $N = \frac{R_0}{R_0-1}$, and (iii) the number of recovered individuals R^* at the deterministic endemic equilibrium, which is taken as initial value at the start of the stochastic simulations.

The temporal dynamics of dengue show remarkable variation due to change in seasonal patterns and real-world dengue data shows a strong seasonal impact. Chapter 4 is devoted to exploring the effect of seasons on the persistence of a single serotype of dengue, taken as a motivational example of a host-vector system. The seasonal dynamics are viewed using RM^s_{SEIR} model, which is identical to RM_{SEIR} used in the previous chapters, except for the fact that vector birth is affected by the change of seasons. The parameter space used in this work corresponds to that of dengue as measured in empirical studies. The seasonal impact affects the vector birth rate as a sinusoidal function whereas the parameters related to human population remain invariant to seasonal fluctuations. The sinusoidal function introduces two seasonal extremes for the vector population, termed as favourable and unfavourable. Twelve different seasonal points (months) during a year were chosen as starting points for the deterministic and stochastic versions of the model. The derivation of different analytic measures including seasonal basic reproductive ratio $R_{t=0|t_0}$ and probabilities of invasion $P_{Inv|I_v=1,t_0}$, and $P_{Inv|I_h=1,t_0}$ were presented. The model was initially investigated at the disease free and endemic equilibrium. The time evolution of the system follows five yearly epidemic cycles at the beginning that were reduced to small annual outbreaks at the endemic equilibrium. The introduction of disease in the naive population at different times of the year resulted in outbreaks of varying intensity within two-and-a-half to four months. The effective reproductive ratio for hosts after the introduction of one viremic mosquito, denoted by $R^{VH}_{t|t_0}$ drops quickly below one during this time period due to the depletion of susceptible host population whereas $R^{HV}_{t|t_0}$ stays above one. It was also shown that the increase in the seroprevalence levels of hosts reduced the probability of invasion. The probability of persistence of dengue infection during one year was greater when introduced during unfavourable seasons, as the change in mosquito population alters the transmission mechanism. This is the reason why there are more persis-

tent runs for January to March (Figure 4.9), with lower number of infectious humans that takes longer to attain the peak (Figure 4.11). From Figure 4.11 it can be seen that if an epidemic occurs during or after October, the peak of I_h starts to decrease as the epidemics is modulated by the decrease in the vector population. The distribution of time to extinction t_e was affected by the seasons. Following introduction during the favourable season, t_e was less after higher peaks of infectious individuals. The peak of infectious humans and the time to attain the peak value was dominated by the seasonal fluctuation in the vector population.

In Chapter 5, two alternative models of indirect transmission for RM_{SIR} are presented. (i) An SIR model with a latent class L_h that acts as a proxy effect of vectors in the host population by delaying transmission compared with direct transmission. (ii) An SIR model that contains a 'Pool' or reservoir of infection P which infects the host population. The concept of reservoir is used to mimic the population of the infectious vectors and the infection is transferred to the hosts by the reservoir. The models are termed as Vector Proxy (VP) model and Reservoir model respectively.

In the VP model, the compartment L_h comprises pre-bitten infected hosts, and the number of individuals entering are dependent upon the rate b_p, the number of infectious hosts and the proportion of susceptible hosts. They leave this compartment to become infectious with rate $\bar{\sigma}$. In the Reservoir model, the number of infectious humans depends upon the rate b_r, the population of individuals in the pool P, and the proportion of susceptible individuals. The individuals in the pool die at a rate $\bar{\delta}$. Mathematically, the VP model can be thought as a $SEIR$ type host model and the Reservoir model as a SIR model in hosts and the dynamics of the infectious compartment of the pool P is represented by a separate equation. The unknown rates in both models are derived analytically and their biological explanation is provided. In the process of approximating a biological process, the first and foremost step is to check the validity of the approximation, i.e., to identify the appropriate parameter regimes where the approximation is valid and where it is not. The regions where these model approximations fail to hold are identified and discussed.

As the main theme of the thesis surrounds the estimation of the pathogen persistence, the CCS calculated by using the VP and reservoir models is compared to the full model. In addition, the deterministic and endemic equilibrium state and the quasi-stationary state of all the models is compared. Computationally, these model approximations are quicker to obtain important results. The approximations of the host-vector model RM_{SIR} presented in the current chapter can help understanding the biology and mechanisms behind the persistence of dengue disease. The results of the above mentioned models can be compared and cross-validated with the full model to visualize the impact of parameters in different endemic

systems. Moreover, these models are easier to treat analytically, and they can be used to answer a wider range of questions as compared to RM_{SIR}.

6.2 Concluding remarks

In summary, in this thesis, a range of theoretical and practical issues pertaining to pathogen persistence are explored. All of the studies conducted in this thesis use different modelling structures with efficient computational approaches to investigate different questions related to the persistence of the pathogen in the host and the vector populations. Therefore, the main areas in which the work undertaken in this thesis fall are investigating pathogen persistence in host-vector systems using minimalist models and efficient algorithms. From the perspective of understanding pathogen persistence, CCS as a measure of long-term pathogen persistence is investigated in two non-seasonal models, RM_{SIR} and RM_{SEIR}. Persistence in the seasonal model is investigated in RM_{SEIR}^s where the relationship between pathogen persistence and (i) the month of disease introduction and (ii) whether the introduction is caused by a viremic host or a viremic vector is investigated. In relation to investigating the efficient modelling structure, two simple models are used to approximate the dynamics of a host-vector system. Under appropriate parameter values both models are able to represent the dynamics of the full model, but their ability to do so depends on the simulation algorithm.

The study conducted in this thesis can be expanded in different directions. One such direction includes incorporating the spatial dimension into the host-vector system. A meta-population structure can be used to investigate the persistence dynamics in the patches and helps understanding the persistence of vector-borne diseases (Adams and Kapan, 2009; Stoddard et al., 2009). The adaptive tau-leap algorithm used in this thesis can help investigate the persistence dynamics in the patches where in case of multiple patches, the approximate models developed in Chapter 5 can be considered to accelerate the simulation process. A very important point of using the tau-leap method is that it can give a good approximation of the full Gillespie algorithm as shown in Figure 5.4. It can even be better if the critical events are handled as described in Cao et al. (2007), as low number of individuals in any compartment are are dealt by using the Gillespie algorithm. In disease ecology, this technique is particularly useful for investigating the extinction process in the patches. The expansion of the models studied in this thesis into the meta-population structure is an interesting area to explore. The extension of the stochastic model to include seasonal dynamics can also be used in the context of the meta-population, if the disease has seasonal dynamics.

The estimation of CCS for the non-seasonal host-vector system is an attempt to quantify

pathogen persistence for the host and the vector population. As the determinants of CCS are the parameters of the model, the diseases which are less affected with the seasonal changes can be incorporated into the models. It is interesting to note that in the current study, the number of infectious hosts and infectious vectors are strongly correlated. A similar parameter regime for mosquito-borne diseases may lead to similar estimates of the CCS as undertaken in Chapter 3. It is anticipated that the inclusion of the latent class especially for the vector population is very important for the realistic measure of CCS. The population size should not be too large to have sub-population structures having different dynamics that affect the persistence of the disease. By using the linear equations constructed in Chapter 3, the seroprevalence level of hosts can be used to estimate the CCS for the host population. The dependency of the CCS can be further explored by using different modelling structures and the linear models can be used to have an idea of the CCS, depending on the availability of the parameters.

Another interesting direction is to investigate the spread of the disease using graphs or network theory. A network is a collection of a set of nodes denoting the individuals and edges that denotes the connection between the nodes. One of the advantages of using networks is that they give a realistic representation of the distribution of the population. The dynamics of interacting epidemics using networks is discussed in Funk et al. (2010) and an overview of networks used in epidemiology are described in Danon et al. (2011). The strength of the relationship between the nodes can affect the disease extinction risk in a population, the degree of spread of the disease and the lifetime of an epidemic. The network structure can be easily incorporated to the models studied in this thesis. The contact heterogeneity can be studied for finding the optimal route required to maintain disease transmission.

In summary, a foundation for estimating the disease persistence and the affect of seasonal forcing on the disease introduction using simple models has been laid by using efficient algorithms in the current thesis. The extensions proposed in this section will allow the researchers to have a deeper understanding of the processes that govern the mechanisms of pathogen persistence. It will be an interesting research problem if the models proposed in this thesis are validated by data acquired from a particular region and examined for the patterns of the 'real' disease. In this regard, the concepts of the meta-population or the network theory will be a valuable addition to the current modelling framework.

APPENDIX A

Appendix for Chapter 2

Appendix for Chapter 2

A.1   Derivation of the basic reproductive number, R_0

In this section, the derivation of basic reproductive number R_0 for both models is presented. The Next Generation Matrix (NGM) method as given in (Diekmann et al., 2010) is employed to obtain R_0 from the ODE system. The equations representing the transmission and transition of infection are referred as the infected subsystem. This system will be decomposed into transmission and transition matrices and these two matrices will be linearised at the stable infection-free or disease free state (DFE). Epidemiologically, as R_0 defines the number of secondary cases that arise with the introduction of an infected individual in an entirely susceptible population, change in number of susceptible individuals is negligible during the initial spread so at DFE, $S_h \approx H$ and $S_v \approx V$.

A.1.1 R_0 for RM_{SIR}

The equation for infected hosts and infected vectors from 2.1 and 2.2 are given as:

$$
\begin{aligned}
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= \alpha I_v \left(\frac{S_h}{H} \right) - (\xi + \gamma)I_h \\
\frac{\mathrm{d}I_v}{\mathrm{d}t} &= \beta S_v \left(\frac{I_h}{H} \right) - \delta I_v
\end{aligned}
\tag{A.1}
$$

The matrix form of transmission and transition for above system can be written as:

$$
A = \begin{pmatrix} -(\xi + \gamma) & \alpha \frac{S_h}{H} \\ \beta \frac{S_v}{H} & -\delta \end{pmatrix}
$$

the above matrix is decomposed in to two matrices, one denoting the transmission of infection F and other representing the transition or change of state, V. At DFE, the matrix F, containing the entry points of infection and the matrix V, containing the transition / exit points of infection (including death) are:

$$
F = \begin{pmatrix} 0 & \alpha \\ \beta \frac{V}{H} & 0 \end{pmatrix}
$$

$$
V = \begin{pmatrix} -(\xi + \gamma) & 0 \\ 0 & -\delta \end{pmatrix}.
$$

The matrix K is obtained by taking the negative of the inverse of V and multiplying by the matrix F.

$$K = F \times -V^{-1} = \begin{pmatrix} 0 & \dfrac{\alpha}{\delta} \\ \dfrac{\beta V}{(\xi + \gamma)H} & 0 \end{pmatrix}$$

In terms of the population based basic reproductive numbers R_0^{HV} and R_0^{VH},

$$K = \begin{pmatrix} 0 & R_0^{VH} \\ R_0^{HV} & 0 \end{pmatrix}$$

The dominant eigenvalue or the spectral radius of the matrix K gives the basic reproductive number R_0, which is explained below:

$$|\lambda I_2 - K| = 0$$

$$\lambda^2 - \left(R_0^{VH} \times R_0^{VH} \right) = \lambda^2 - R_0 = 0$$

Here R_0 is a general threshold of the host-vector system. It is a smultiple of R_0^{HV} and R_0^{VH} which are the threshold conditions on vector and host populations respectively.

$$R_0 = \sqrt{\dfrac{\alpha \beta V}{(\xi + \gamma)\delta H}}.$$

$\blacksquare$

A.1.2 $\quad R_0$ for RM_{SEIR}

Using the same analogy as in A.1.1, the matrices A, F and V were formed by taking the infected subsystem from equation sets 2.14 and 2.15.

$$A = \begin{pmatrix} -(\sigma + \gamma) & 0 & 0 & \alpha \\ \sigma & -(\xi + \gamma) & 0 & 0 \\ 0 & \beta\frac{V}{H} & -(\rho + \delta) & 0 \\ 0 & 0 & \rho & -\delta \end{pmatrix}$$

$$F = \begin{pmatrix} 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 \\ 0 & \beta\frac{V}{H} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad ; \quad V = \begin{pmatrix} -(\sigma + \gamma) & 0 & 0 & 0 \\ \sigma & -(\xi + \gamma) & 0 & 0 \\ 0 & 0 & -(\rho + \delta) & 0 \\ 0 & 0 & \rho & -\delta \end{pmatrix}$$

The inverse of the matrix V is following;

$$
V^{-1} = \begin{pmatrix}
\dfrac{1}{\sigma + \gamma} & 0 & 0 & 0 \\[2ex]
\dfrac{\sigma}{(\sigma + \gamma)(\xi + \gamma)} & \dfrac{1}{\xi + \gamma} & 0 & 0 \\[2ex]
0 & 0 & \dfrac{1}{\rho + \delta} & 0 \\[2ex]
0 & 0 & \dfrac{\rho}{\delta(\rho + \delta)} & \dfrac{1}{\delta}
\end{pmatrix}
$$

The matrix K has the following 4×4 form:

$$
K = \begin{pmatrix}
0 & 0 & \dfrac{\alpha\,\rho}{\delta(\rho + \delta)} & \dfrac{\alpha}{\delta} \\[2ex]
0 & 0 & 0 & 0 \\[2ex]
\dfrac{\beta\,\sigma\,V}{H(\sigma + \gamma)(\xi + \gamma)} & \dfrac{\beta\,V}{H(\xi + \gamma)} & 0 & 0 \\[2ex]
0 & 0 & 0 & 0
\end{pmatrix}
$$

In terms of the population based basic reproductive numbers for RM_{SEIR},

$$
K = \begin{pmatrix}
0 & 0 & R_0^{VH} & \dfrac{\alpha}{\delta} \\[2ex]
0 & 0 & 0 & 0 \\[2ex]
R_0^{HV} & \dfrac{\beta\,V}{H(\xi + \gamma)} & 0 & 0 \\[2ex]
0 & 0 & 0 & 0
\end{pmatrix}
$$

By evaluating the determinant $|\lambda I_4 - K| = 0$, the following characteristic equation is obtained:

$$
\lambda^2 \left(\lambda^2 - \left(R_0^{VH} \times R_0^{VH}\right)\right) = \lambda^2 \left(\lambda^2 - R_0\right) = 0
$$

Two of the four eigenvalues are zero. The dominant eigenvalue obtained is the basic reproductive number of RM_{SEIR}.

$$
R_0 = \sqrt{\dfrac{\alpha\beta\sigma\rho V}{H\delta(\xi + \gamma)(\sigma + \gamma)(\rho + \delta)}}
$$

■

APPENDIX B

Appendix for Chapter 3

Appendix for Chapter 3

B.1 Tau-Leaping algorithm

- Calculate the event rates R_j for all events E_j.

- Calculate the auxiliary quantities μ_i and σ_i^2 for every population P_i.

$$\mu_i = \sum_j v_{ij} R_j \qquad\qquad \sigma_i^2 = \sum_j v_{ij}^2 R_j$$

  Here v_{ij} is the effect of an event E_j over a compartment P_i. In an individual based model with different compartments this quantity is usually expressed as a matrix having values {-1, 0, +1}.

- Find ϵ_i by identifying the highest order event rate that each population P_i takes part in

- Find the value of τ by using the following relation:

$$\tau = \min_i \left(\frac{\max\{\epsilon_i x_i, 1\}}{|\mu_i|}, \frac{\max\{\epsilon_i x_i, 1\}^2}{\sigma_i^2} \right)$$

  where ϵ_i is derived in the next section.

- Find the number of times K_j each event E_j occur.

$$K_j \sim Poisson(R_j \tau)$$

- Update the model with the state change vector for every compartment.

$$x(t + \tau) = x(t) + \sum_j K_j v_j$$

- Check that no populations P_i became negative.

- Repeat until finished.

B.2 Finding highest order rate

The tau leaping algorithm aims to leap as large as possible, provided that the change in rates R_j is within acceptable range. For a given change in population X_i, the relative change rate R_j should be smaller than the tolerance value ϵ. In all the experiments carried out in chapter 3, the value of ϵ was significantly lower than the tolerance value derived in this section to avoid error in estimating the Gillespie process.

The order of a rate is usually the highest product of populations X_i; $i = 1, 2, 3, ...$ involved. It means that the reactions λH, γR_h and δS_v etc, are all first order. The transmission rates in both models are the ones which includes the product of two population compartments. These are $\alpha I_v\left(\dfrac{S_h}{H}\right)$ and $\beta S_v\left(\dfrac{I_h}{H}\right)$. The order of these reactions is the highest order in both models so the tolerance ϵ is based on their reaction order.

Now, Consider $R_1 = \alpha I_v\left(\dfrac{S_h}{H}\right)$. The change in R_1, i.e., ΔR_1 is written as:

$$\Delta R_1 = \alpha(I_v + \Delta I_v)\left(\frac{S_h + \Delta S_h}{H + \Delta H}\right) - \alpha I_v\left(\frac{S_h}{H}\right). \tag{B.1}$$

By taking LCM, cancelling like terms and using the fact that $H = S_h + I_h + R_h$ and $\Delta H = \Delta S_h + \Delta I_h + \Delta R_h$.

$$\Delta R_1 = \alpha \frac{S_h^2 \Delta I_v + S_h I_h \Delta I_v + \Delta S_h I_h I_v + \Delta S_h R_h I_v + S_h R_h \Delta I_v - S_h I_v\left(\Delta I_h + \Delta R_h\right)}{H\left(H + \Delta H\right)}. \tag{B.2}$$

Here the product of terms like $\Delta x \Delta y$ are ignored. Dividing above equation by R_1 for finding the relative change,

$$\frac{\Delta R_1}{R_1} = \frac{S_h\left(S_h + I_h\right)\Delta I_v + \Delta S_h\left(I_h + R_h\right)I_v - S_h I_v\left(I_h + R_h\right) + S_h R_h \Delta I_v}{S_h I_v\left(H + \Delta H\right)}. \tag{B.3}$$

Equation B.3 can be arranged in the following manner.

$$\frac{\Delta R_1}{R_1} = \frac{H}{H + \Delta H}\left(\frac{\Delta I_v}{I_v}\right) + \frac{H}{H + \Delta H}\left(\frac{\Delta S_h}{S_h}\right) - \left(\frac{\Delta S_h + \Delta I_h + \Delta R_h}{H + \Delta H}\right). \tag{B.4}$$

This equation can be simplified by using the fact that the change in the host population $\frac{H}{H+\Delta H}$ is negligible.

$$\frac{\Delta R_1}{R_1} \leq \frac{H}{H + \Delta H}\left(\epsilon + \epsilon\right). \tag{B.5}$$

Here ϵ is the error control parameter.

$$\frac{\Delta R_1}{R_1} \leq 2\epsilon. \tag{B.6}$$

The reaction is of second order. The value of ϵ should be less than 0.5. In all the analysis undertaken in Chapter 3 and Chapter 5, its value is set to 0.01 and the value is set to 0.001 for the analysis undertaken in Chapter 4.

B.3 Sensitivity analysis

To assess and quantify the influence of each parameter for variability in the persistence threshold, and to target the important areas within the observed patterns, this section discusses the elasticity analysis of both models.

To perform the formal elasticity analysis, the parameters of the model on CCS were sampled by using Latin Hypercube Sampling (LHS) which is a very efficient method for sampling Blower and Dowlatabadi (1994) and Sanchez, M. A., & Blower (1997). LHS is a type of stratified Monte Carlo sampling where sampling of parameter values is done without replacement. LHS table consists of n rows and k columns, where n is the number of samples and k is the number of parameters. This table is generated by sampling the parameters independently of each other from a pre-defined distribution, while every other sample takes into account the rows and columns of previously generated sample points (Mckay et al., 2000). In this way, the LHS scheme achieves the same level of accuracy as random sampling by using fewer samples. This method was used to construct a multi-dimensional space of plausible parameter values from a multivariate distribution, whose minimum and maximum values are given in Table 3.4 on page 69. Uniform probability distribution functions (pdfs) were defined for each parameter, which is preferable when data are not available Marino et al. (2008). Since the starting point of the stochastic simulations was taken as the deterministic equilibrium points in every compartment, parameter combinations which yielded $R_0 \leq 1$ were discarded.

A parameter space containing 100 parameter sets was generated from 100 equiprobable intervals for each parameter in the parameter space and randomly sampled 100 times without replacement. These 100 samples were randomly permuted to yield 100 non-overlapping parameter sets having R_0 lying in the interval $\{1.1, 7.6\}$. The resulting LHS design was again iterated 100 times to reduce the correlation among the variables. Parameter sets were simulated for twenty-five years, starting at the deterministic equilibrium for all sub-populations. Stochastic simulation for every parameter combination was repeated a hundred times, so the whole sensitivity analysis was based on $7 \times 100 \times 100 = 70000$ simulations. A smoothing function was applied on the simulation results to estimate the probability of extinction at different population sizes. Smoothing also helps clear the monotonicity of the patterns of probability of extinction with increasing population sizes. CCS was attained if the infec-

tion in both populations is still retained in half of stochastic simulations at the end of the simulation time.

B.4 Partial rank correlation coefficients (PRCCs)

The order of input matrix obtained from sampling is $n \times k$, where n is number of samples created and k is the number of parameters. In this experiment, the order of output matrix is $n \times 1$ as CCS is the only output variable of interest generated form the stochastic simulations. Ranks are assigned to the data values of input variables x_i and output measures y_j, where i runs from 1 to k, and j runs from 1 to the number of outputs obtained from the experiment (*in this case n=100, k = 7 and only one output measure is of interest* i.e., CCS, *so j=1*). After ranking the input and output variables, two regression models $\bar{I}_k : x_k = f(x_i) \ \ k \neq i$ and $\bar{Y}_k : y = f(x_i) \ i \neq k$, were then made for exploring the relation between a parameter x_k, and CCS. So, $\bar{I}_k$ have x_k as response variable and $\bar{Y}_k$ have CCS as response variable. The remaining parameters were the explanatory variables in both regression equations. The residuals from both models were computed, and then as suggested in (Bishara and Hittner, 2012), Pearson's correlation coefficient was calculated on residuals from both models to obtain PRCC which is shown in Table 3.4.

APPENDIX C

Appendix of Chapter 4

Appendix of Chapter 4

C.1 Solving a non-autonomous system of ordinary differential equations

In RM_{SIER}^s, the system of differential equations is non-autonomous as the birth rate of vectors $\delta_b(t)$ is changing with time (equation set 4.1). Loosely speaking, an autonomous dynamical system does not contains time as an explicit variable.

It is generally possible to convert a non-autonomous system of ordinary differential equations to an autonomous one by expanding the state space of the variables. A new variable z can be assumed as $z(t) = t$, so that the rate of change of z with respect to time, t is constant. This is done to hide the time-dependence of the system. Theoretically the correlation between z and t is always 1. However, there are two trade-offs of using this new variable: (i) the dimension of the system d goes up by $d + 1$, and (ii), the new system has no equilibria (the equilibrium points varies with time). The transformed ordinary differential equations has eight equations.

$$
\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= \Lambda H - \alpha I_v \left(\frac{S_h}{H} \right) - \gamma S_h \\
\frac{\mathrm{d}E_h}{\mathrm{d}t} &= \alpha I_v \left(\frac{S_h}{H} \right) - (\sigma + \gamma) E_h \\
\frac{\mathrm{d}I_h}{\mathrm{d}t} &= \sigma E_h - (\xi + \gamma) I_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \xi I_h - \gamma R_h \\
\frac{\mathrm{d}S_v}{\mathrm{d}t} &= \delta_b(t)\, V - \beta S_v \left(\frac{I_h}{H} \right) - \delta_d\, S_v \\
\frac{\mathrm{d}E_v}{\mathrm{d}t} &= \beta S_v \left(\frac{I_h}{H} \right) - (\rho + \delta_d)\, E_v \\
\frac{\mathrm{d}I_v}{\mathrm{d}t} &= \rho E_v - \delta_d\, I_v \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= 1
\end{aligned}
\tag{C.1}
$$

To solve above system numerically, ode45 solver **??** is used. The maximum step was set to one-tenth of a day. The relative tolerance and absolute tolerance of the solver were in the magnitudes of 10^{-6} and 10^{-7} respectively. The initial conditions are provided by considering two different scenarios. If the invasion is because of an infectious human, then the solver is started with all susceptible population of the mosquitoes. In human population, there are $H - 1$ susceptible, one infectious and no infected and recovered individuals. The human and vector populations change their roles in case of invasion due to one infectious vector. In the stochastic solver, the error control parameter ϵ is kept 0.001, this parameter should be less

than 0.5 (appendix B.2). The lower value is chosen to avoid the chances of error in estimating the fluctuating total vector population.

C.2 Invasion and extinction probabilities

In RM_{SIR} and RM_{SEIR}, the duration of latent and infectious period is exponentially distributed, as the rates of exposure, infection and recovery are constant. Summation of the distribution of secondary infections provides off spring distributions which are geometric with means R_0^{HV} and R_0^{VH} for hosts and vectors respectively (Grassly and Fraser, 2006; Lloyd et al., 2007). This distribution is highly skewed and over-dispersed (variance is more than mean) as compared to Poisson distribution.

After the introduction of infection in naive population, invasion probability, $P(\text{Inv})$ can be estimated by using the results from branching theory; which tell us that the likelihood of disease invasion depends upon the (i) average (R_0) and (ii) distribution of secondary infections around this average. The extinction probability is simply $P(\text{ext}) = 1 - P(\text{Inv})$. Here extinction means that there is no secondary case from an infectious individual. The extinction probability is found by calculating the smallest non-negative root of the equation:

$$G(P(\text{ext})) = a$$

Here G(Pext) is the probability generating function of the distribution of secondary infections. The generating function for geometric distribution is given by:

$$G(P(\text{ext})) = \frac{1}{1 + \mu(1 - a)}$$

In the case of directly transmitted infections, the average number of secondary infections is denoted as R_0. After putting the value, following equation came from simplification:

$$a^2 - \left(\frac{R_0 + 1}{R_0}\right)a + \frac{1}{R_0} = 0$$

solving above equation leads two values for $P(\text{ext})$; 1 and $\frac{1}{R_0}$ which gives a well known result for invasion probability $P(\text{Inv}) = 1 - \frac{1}{R_0}$ for directly transmitted infections, when $R_0 > 1$.

For a host-vector system there are two generating functions for each population.

$$G_H(P(\text{ext}_H)) = \frac{1}{1 + R_0^{HV}(1 - a_H)}; \qquad G_V(P(\text{ext}_V)) = \frac{1}{1 + R_0^{VH}(1 - a_V)} \qquad \text{(C.2)}$$

In order to get extinction probabilities, following composite functions need to be solved for a_H and a_V.

$$G_H(G_V(P(\text{ext}_H))) = a_H; \qquad G_V(G_H(P(\text{ext}_V))) = a_V \tag{C.3}$$

Putting the values from equation C.2 yields the following equation for a_H:

$$a_H + R_0^{VH} a_H + \frac{R_0^{VH} a_H}{R_0^{HV}(a_H - 1) - 1} - 1 = 0$$

The solution is smaller than 1 and

$$\frac{R_0^{HV} + 1}{R_0^{HV}(1 + R_0^{VH})}$$

Above term will be less than 1, if and only if, $R_0^{VH} \times R_0^{HV} > 1$ which equals $R_0 > 1$. For $R_0 > 1$, the probability of major outbreak, after the introduction of one infectious vector is

$$P(Inv) = 1 - \frac{R_0^{HV} + 1}{R_0^{HV}(1 + R_0^{VH})}$$

Similar expression with switching the roles of R_0^{HV} and R_0^{VH} is obtained, if C.3 is solved for a_V.

$\blacksquare$

APPENDIX D

Appendix of Chapter 5

Appendix of Chapter 5

D.1 Basic reproductive number R_0 for both models

D.1.1 VP model

By using the next generation matrix technique (Diekmann et al., 2010), the matrices F and V containing 'gains' and 'losses' respectively to the individuals in the pair of differential equations $\dfrac{dL}{dt}$ and $\dfrac{dI_h}{dt}$ are:

$$F = \begin{pmatrix} 0 & b_{vp} \\ 0 & 0 \end{pmatrix} \tag{D.1}$$

$$V = \begin{pmatrix} \overline{\sigma} + \gamma & 0 \\ -\overline{\sigma} & (\xi + \gamma) \end{pmatrix} \tag{D.2}$$

R_0 is given by the spectral radius of the matrix FV^{-1}. The entries of FV^{-1} provide the rate at which at which infected individuals in L_h produce new infections in I_h, times the average length of time an individual spends in a single visit to compartment L_h. The expression for R_0 is:

$$R_0 = \frac{b_{vp}\overline{\sigma}}{(\xi + \gamma)(\overline{\sigma} + \gamma)}$$

■

D.1.2 Reservoir model

Using the next generation matrix technique (Diekmann et al., 2010), the matrices F and V containing 'gains' and 'losses' to the individuals in the pair of differential equations $\dfrac{dP}{dt}$ and $\dfrac{dI_h}{dt}$ are:

$$F = \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} \tag{D.3}$$

$$V = \begin{pmatrix} \delta & -b_r \\ c & (\xi + \gamma) \end{pmatrix} \tag{D.4}$$

Here, the entries of FV^{-1} provide the rate at which an infected individuals in P produce new infections in I_h, times the average length of time an individual spends in a single visit to compartment I_h. The expression for R_0 is:

$$R_0 = \frac{b_r c}{\overline{\delta}(\xi + \gamma)}$$

■

D.2 Stability analysis of VP and Reservoir model

In the following subsections, the local asymptotic stability analysis of both equilibria, DFE and EE is presented. The goal is to investigate whether the disease free equilibrium point E^0 is locally asymptotically stable when $R_0 < 1$. Moreover at $R_0 > 1$, E^0 becomes unstable and endemic equilibrium point E^* is locally asymptotically stable.

D.2.1 VP model

The host population is closed, so $R_h = H - S_h - I_h$. The differential equation system 5.2 exhibit two distinct equilibrium points, (i) $E^0\Big(H, 0, 0\Big)$ is the disease free equilibrium (DFE) and (ii) $E^*\Big(S_h^*, I_h^*, L_h^*\Big)$ is the endemic equilibrium (EE). DFE is stable at $R_0 < 1$ and becomes unstable at $R_0 \geq 1$. The endemic equilibrium is stable at $R_0 \geq 1$. The basic reproductive ratio R_0 acts as a threshold value for the existence of these equilibria.

The local stability of the equilibrium points is governed by the Jacobian matrix

$$J\Big(S_h, I_h, L_h\Big) = \begin{pmatrix} -(b_{vp}\dfrac{I_h}{H} + \gamma) & -b_{vp}\dfrac{S_h}{H} & 0 \\ 0 & -(\xi + \gamma) & \overline{\sigma} \\ b_{vp}\dfrac{I_h}{H} & b_{vp}\dfrac{S_h}{H} & (\overline{\sigma} + \gamma) \end{pmatrix} \tag{D.5}$$

D.2.1.1 Stability of DFE

At DFE, there is no infection in the population. The number of recovered individuals falls to zero and as a result, the population consists of only susceptible individuals. The Jacobian matrix D.5 at $E^0\Big(H, 0, 0\Big)$ is:

$$J\Big(H, 0, 0\Big) = \begin{pmatrix} -\gamma & -b_{vp} & 0 \\ 0 & -(\xi + \gamma) & \overline{\sigma} \\ 0 & b_{vp} & (\overline{\sigma} + \gamma) \end{pmatrix} \tag{D.6}$$

The characteristic equation is $(\lambda + \gamma)\Big(\lambda^2 + (\xi + \overline{\sigma} + 2\gamma)\lambda + (\xi + \gamma)(\overline{\sigma} + \gamma)(1 - R_0)\Big) = 0$ and the eigenvalues are $\lambda_1 = -\gamma$ and $\lambda_{2,3} = -\dfrac{\xi + \overline{\sigma} + 2\gamma}{2} \pm \dfrac{\sqrt{(\xi + \overline{\sigma} + 2\gamma)^2 - 4\overline{\sigma}(\xi + \gamma)(1 - R_0)}}{2}$.
All of the eigenvalues have negative real part for $R_0 < 1$ and so the $E^0\Big(H, 0, 0\Big)$ is locally asymptotically stable for $R_0 < 1$, according to RouthHurwitz criterion. For $R_0 > 1$, the eigenvalue $\lambda_{2,3} > 0$, hence $E^0\Big(H, 0, 0\Big)$ becomes unstable. ∎

D.2.1.2   Stability of EE

The Jacobian matrix D.5 at $E^*\Big(S_h^*, I_h^*, L_h^*\Big)$ is:

$$J\Big(S_h^*, I_h^*, L_h^*\Big) = \begin{pmatrix} -(b_{vp}\dfrac{I_h^*}{H} + \gamma) & -b_{vp}\dfrac{S_h^*}{H} & 0 \\ 0 & -(\xi + \gamma) & \overline{\sigma} \\ b_{vp}\dfrac{I_h^*}{H} & b_{vp}\dfrac{S_h^*}{H} & (\overline{\sigma} + \gamma) \end{pmatrix} \tag{D.7}$$

The characteristic polynomial is:

$$\left(\lambda + \frac{b_{vp}I_h^*}{H} + \gamma\right)\left((\lambda + \xi + \gamma)(\lambda + \overline{\sigma} + \gamma) - \overline{\sigma}b_{vp}\frac{S_h^*}{H}\right) + b_{vp}^2\overline{\sigma}\frac{S_h^*}{H}\frac{I_h^*}{H} \tag{D.8}$$

By replacing the values of S_h^*, I_h^* and R_0 from equations (5.2) and (5.4), the characteristic polynomial becomes:

$$\lambda^3 + K\lambda^2 + L\lambda + M$$

where $\quad K = \xi + \overline{\sigma} + 3\gamma + \gamma(R_0 - 1), \quad L = \gamma(R_0 - 1)\Big(\xi + \overline{\sigma} + 2\gamma\Big),$ and $\quad M = \gamma(\xi + \gamma)(\overline{\sigma} + \gamma)(R_0 - 1)$

For $R_0 > 1$, the coefficients K, L, and M are positive and according to the RouthHurwithz condition, the characteristic polynomial satisfies the following relation:

$$KL > M \tag{D.9}$$

Therefore, $E^*\Big(S_h^*, I_h^*, L_h^*\Big)$ is locally asymptotically stable.   ∎

D.2.2   Reservoir model

The local stability of the equilibrium points is governed by the Jacobian matrix

$$J\Big(S_h, P, I_h, R_h\Big) = \begin{pmatrix} -(b_r\dfrac{P}{H} + \gamma) & -b_r\dfrac{S_h}{H} & 0 & 0 \\ 0 & -\overline{\delta} & c & 0 \\ b_r\dfrac{P}{H} & b_r\dfrac{S_h}{H} & -(\xi + \gamma) & 0 \\ 0 & 0 & \xi & -\gamma \end{pmatrix} \tag{D.10}$$

D.2.2.1 Stability of DFE

At $E^0\Big(H,0,0,0\Big)$ the matrix D.10 becomes:

$$J\Big(S_h,P,I_h,R_h\Big) = \begin{pmatrix} -\gamma & -b_r & 0 & 0 \\ 0 & -\bar{\delta} & c & 0 \\ 0 & b_r & -(\xi+\gamma) & 0 \\ 0 & 0 & \xi & -\gamma \end{pmatrix} \tag{D.11}$$

The characteristic equation is $(\lambda+\gamma)^2\Big(\lambda^2 + (\xi+\bar{\delta}+\gamma)\lambda + \delta(\xi+\gamma)(1-R_0)\Big) = 0$. $\lambda_{1,2} = -\gamma$ is the eigenvalue of multiplicity two and the other two eigenvalues are found by using the quadratic formula, $\lambda_{3,4} = -\dfrac{\xi+\bar{\delta}+\gamma}{2} \pm \dfrac{\sqrt{(\xi+\bar{\delta}+\gamma)^2 - 4\bar{\delta}(\xi+\gamma)(1-R_0)}}{2}$. $E^0\Big(H,0,0\Big)$ is locally asymptotically stable for $R_0 < 1$, in lieu of the argument mentioned in section D.2.1.1.
∎

D.2.2.2 Stability of EE

The Jacobian matrix D.10 at $E^*\Big(S_h^*,P^*,I_h^*,R_h^*\Big)$ is:

$$J\Big(S_h^*,P^*,I_h^*,R_h^*\Big) = \begin{pmatrix} -(b_r\dfrac{P^*}{H}+\gamma) & -b_r\dfrac{S_h^*}{H} & 0 & 0 \\ 0 & -\bar{\delta} & c & 0 \\ b_r\dfrac{P^*}{H} & b_r\dfrac{S_h^*}{H} & -(\xi+\gamma) & 0 \\ 0 & 0 & \xi & -\gamma \end{pmatrix} \tag{D.12}$$

The characteristic equation for D.12 is:

$$\Big(\lambda+\gamma\Big)\Big(\lambda+\delta\Big)\Big(\lambda+\xi+\gamma\Big)\Big(\lambda+b_r\dfrac{P^*}{H}+\gamma\Big) = 0 \tag{D.13}$$

All eigenvalues are negative as parameters of the models are always positive and the number of individuals in compartment P are always greater than or equal to zero. The expanded polynomial, by replacing $b_r\dfrac{P^*}{H} = \gamma(R_0 - 1)$ is:

$$\lambda^4 + K\lambda^3 + L\lambda^2 + M\lambda + N \tag{D.14}$$

where

$$K = \xi + \bar{\delta} + 3\gamma + \gamma(R_0 - 1),$$

$$L = \left(\left(\gamma(R_0 - 1) + \gamma \right)\left(\xi + \bar{\delta} + \gamma \right) + \delta(\xi + \gamma) \right) + \gamma\left(\gamma(R_0 - 1) + \xi + \bar{\delta} + 2\gamma \right),$$

$$M = \delta(\xi + \gamma)\left(\left(\gamma(R_0 - 1) + \gamma \right) + \gamma\left((\gamma(R_0 - 1) + \gamma)(\xi + \bar{\delta} + \gamma) + \delta(\xi + \gamma) \right), \&\right.$$

$$N = \delta\gamma(\xi + \gamma)\left(\left(\gamma(R_0 - 1) + \gamma \right).$$

The RouthHurwithz conditions for above 4th order polynomial are as follows:

- All the coefficients should be greater than zero.

- $KL > M$

- $KLM > M^2 + K^2 N$

All these conditions are met i.e., the endemic equilibrium point $E^*\left(S_h^*, P^*, I_h^*, R_h^* \right)$ is locally asymptotically stable for $R_0 > 1$. ∎

D.3 Derivation of unknown terms in both models

During the approximations of the full model, some unknown terms arise in VP and Reservoir model. The objective of this section is to derive the unknown parameters in terms of the parameters of the full model. In all derivations it was assumed that: (i) $\gamma R_0 \lll 1$, (ii) $\beta\gamma \lll 1$, and (iii) $x + \gamma \approx x$, where x is any parameter (e.g. If $x = \alpha$ then $\alpha + \gamma \approx \alpha$).

D.3.1 b_{vp} and $\bar{\sigma}$ in VP model

In VP model, there are two unknown quantities, b_{vp} and $\bar{\sigma}$. In order to compare this model to RM_{SIR}, a comprehensive description of these terms is required. A $RM_{SEIR_{SI}}$ model (*without exposed class in vector population*) is constructed and the expressions for exposed and infectious individuals at endemic equilibrium are compared to represent b_{vp} and $\bar{\sigma}$ in terms of the parameters of $RM_{SEIR_{SI}}$ model.

The expression for infectious individuals (I_h^*) at endemic equilibrium for the $RM_{SEIR_{SI}}$ model is:

$$I_h^* = \frac{H^2 \gamma \delta (R_0 - 1)}{\beta(\gamma H + \alpha V)} \tag{D.15}$$

where R_0 for $RM_{SEIR_{SI}}$ model is

$$R_0 = \frac{\alpha\beta\sigma V}{(\xi + \gamma)(\sigma + \gamma)\delta H} \tag{D.16}$$

Setting this expression equal to I_h^* in equation set (5.5)

$$I_h^* = \frac{H^2\gamma\delta(R_0 - 1)}{\beta(\gamma H + \alpha V)} = \frac{\gamma H}{b_{vp}}(R_0 - 1) \tag{D.17}$$

By using (ii), cancelling like terms and making b_{vp} as the subject of the formula

$$b_{vp} \approx \frac{\beta(\gamma H + \alpha V)}{\delta H} \tag{D.18}$$

The same value of b_{vp} is obtained when comparing the equilibrium point of VP model with RM_{SIR}.

In similar fashion, comparing the expression for exposed individuals (E_h^*) in $RM_{SEIR_S I}$ model and L_h^* in VP model yields:

$$\frac{-H^2\gamma\delta(\xi + \gamma)}{\beta\sigma(\sigma + \gamma)(\gamma H + \alpha V)}\Big(\gamma + \sigma H - H R_0(\sigma + \gamma)\Big) = \frac{-H\gamma(\xi + \gamma)}{b_{vp}\overline{\sigma}(\overline{\sigma} + \gamma)}\Big(\gamma + \overline{\sigma} H - H R_0(\overline{\sigma} + \gamma)\Big) \tag{D.19}$$

Using (i) and (iii), the equation becomes,

$$\frac{H^2\delta\sigma(1 - R_0)}{\beta\sigma^2(\gamma H + \alpha V)} = \frac{H\overline{\sigma}(1 - R_0)}{b_{vp}\overline{\sigma}^2} \tag{D.20}$$

Simple algebra yield the equation for $\overline{\sigma}$:

$$\overline{\sigma} \approx \frac{\beta\sigma(\gamma H + \alpha V)}{b_{vp}\delta H} \tag{D.21}$$

In both approximations, D.18 and D.21 the term γH can be dropped as this value has very little contribution ($< 0.007\%$) in the expression $\alpha V + \gamma H$. This disparity arises due the difference in the values of rate of transmission of infection $\alpha = 0.1$, in days versus the average birth / death rate of hosts $\gamma = 4.125 \times 10^{-5}$, in days. So in this expression $\alpha V + \gamma H$, the term γH can be dropped. The new expressions for b_{vp} and $\overline{\sigma}$ are:

$$b_{vp} \approx \frac{\alpha\beta}{\delta}\left(\frac{V}{H}\right) \tag{D.22}$$

$$\overline{\sigma} \approx \frac{\alpha\beta}{\delta}\left(\frac{\sigma}{b_{vp}}\right)\left(\frac{V}{H}\right) \tag{D.23}$$

∎

D.3.2 b_r and $\bar{\delta}$ in Reservoir model

The unknown terms in th Reservoir model can be found by comparing the equilibrium points of RM_{SIR} and Reservoir model. The expression for equilibrium points for susceptible individuals in RM_{SIR}, equation set (2.4) is compared with the expression for equilibrium points for susceptible individuals in the Reservoir model equation set (5.11):

$$S_h = \frac{H^2\Big(\delta(\xi + \gamma) + \beta\gamma\Big)}{\beta(\gamma H + \alpha V)} = \frac{\bar{\delta}H(\xi + \gamma)}{b_r\, c} \tag{D.24}$$

Using (ii) and dropping the value of γH in $\alpha V + \gamma H$ by using the argument stated above

$$S_h = \frac{H^2\delta(\xi + \gamma)}{\alpha\beta V} = \frac{\bar{\delta}H(\xi + \gamma)}{b_r\, c} \tag{D.25}$$

Replacing the value of $c = \beta\dfrac{V}{H}$ leads to:

$$b_r = \alpha \tag{D.26}$$

Hence by substituting the value of b_r in D.25

$$\bar{\delta} = \delta \tag{D.27}$$

∎

BIBLIOGRAPHY

Bibliography

Adams, B. and M. Boots, 2010. How important is vertical transmission in mosquitoes for the persistence of dengue? Insights from a mathematical model. *Epidemics* **2**:1–10.

Adams, B. and D. D. Kapan, 2009. Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics. *PloS one* **4**:1–10.

Anderson, J. R. and R. Rico-Hesse, 2006. Aedes aegypti vectorial capacity is determined by the infecting genotype of dengue virus. *The American journal of tropical medicine and hygiene* **75**:886–892.

Anderson, R. M. and R. M. May, 1981. The Population Dynamics of Microparasites and Their Invertebrate Hosts. *Philosophical Transactions of the Royal Society B: Biological Sciences* **291**:451–524.

Andraud, M., N. Hens, and P. Beutels, 2013. A simple periodic-forced model for dengue fitted to incidence data in Singapore. *Mathematical Biosciences* **244**:22–28.

Andraud, M., N. Hens, C. Marais, and P. Beutels, 2012. Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches. *PloS one* **7**:e49085.

Barbazan, P., M. Guiserix, W. Boonyuan, W. Tuntaprasart, D. Pontier, and J.-P. Gonzalez, 2010. Modelling the effect of temperature on transmission of dengue. *Medical and Veterinary Entomology* **24**:66–73.

Barmak, D. H., C. O. Dorso, M. Otero, and H. G. Solari, 2011. Dengue epidemics and human mobility. *Physical Review E* **84**:011901.

Bartlett, M., 1957. Measles periodicity and community size. *Journal of the Royal Statistical Society. Series A (General)* **120**:48–70.

Bartlett, M., 1960. The critical community size for measles in the United States. *Journal of the Royal Statistical Society. Series A (General)* **123**:37–44.

Bartley, L. M., C. A. Donnelly, and G. P. Garnett, 2002. The seasonal pattern of dengue in endemic areas: Mathematical models of mechanisms. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **96**:387–397.

Beyer, H. L., K. Hampson, T. Lembo, S. Cleaveland, M. Kaare, and D. T. Haydon, 2012. The implications of metapopulation dynamics on the design of vaccination campaigns. *Vaccine* **30**:1014–1022.

Bishara, A. J. and J. B. Hittner, 2012. Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological methods* **17**:399–417.

Blower, S. M. and H. Dowlatabadi, 1994. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *International Statistical Review / Revue Internationale de Statistique* **62**:229.

Blower, S. M., D. Hartel, H. Dowlatabadi, R. M. Anderson, and R. M. May, 1991. Drugs, sex and HIV: a mathematical model for New York City. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **331**:171–87.

Boots, M., 1999. A general hostpathogen model with freeliving infective stages and differing rates of uptake of the infective stages by infected and susceptible hosts. *Researches on Population Ecology* **41**:189–194.

Boyce, W. E. and R. C. DiPrima, 1986. Elementary Differential Equations and Boundary Value Problems. 4th Ed. With Solutions Manual. John Wiley and Sons.

Brownstein, J. S., E. Hett, and S. L. O'Neill, 2003. The potential of virulent Wolbachia to modulate disease transmission by insects. *Journal of Invertebrate Pathology* **84**:24–29.

Butler, D., 2013. Malaria. *Nature* .

Cancrini, G., a. Frangipane di Regalbono, I. Ricci, C. Tessarin, S. Gabrielli, and M. Pietrobelli, 2003. Aedes albopictus is a natural vector of Dirofilaria immitis in Italy. *Veterinary Parasitology* **118**:195–202.

Cao, Y., D. T. Gillespie, and L. R. Petzold, 2006. Efficient step size selection for the tau-leaping simulation method. *The Journal of chemical physics* **124**:044109.

Cao, Y., D. T. Gillespie, and L. R. Petzold, 2007. Adaptive explicit-implicit tau-leaping method with automatic tau selection. *The Journal of chemical physics* **126**:224101.

Castle, M. D. and C. a. Gilligan, 2012. An epidemiological framework for modelling fungicide dynamics and control. *PLoS ONE* **7**.

Cecílio, a. B., E. S. Campanelli, K. P. R. Souza, L. B. Figueiredo, and M. C. Resende, 2009. Natural vertical transmission by Stegomyia albopicta as dengue vector in Brazil. *Brazilian journal of biology = Revista brasileira de biologia* **69**:123–127.

Chan, M. and M. a. Johansson, 2012. The Incubation Periods of Dengue Viruses. *PLoS ONE* **7**:e50972.

Chen, L. H. and M. E. Wilson, 2004. Transmission of dengue virus without a mosquito vector: nosocomial mucocutaneous transmission and other routes of transmission. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **39**:e56–60.

Chen, S.-C. and M.-H. Hsieh, 2012. Modeling the transmission dynamics of dengue fever: implications of temperature effects. *The Science of the total environment* **431**:385–91.

Chikaki, E. and H. Ishikawa, 2009. A dengue transmission model in Thailand considering sequential infections with all four serotypes. *Journal of Infection in Developing Countries* **3**:711–722.

Christophers, S. R., 1960. Aedes Aegpti (L.) the yellow ferver mosquito: Its life history, bionomics and structure. Cambridge University Press, New York.

Chye, J. K., C. T. Lim, K. B. Ng, J. M. H. Lim, R. George, and S. K. Lam, 1997. Vertical Transmission of Dengue. *Clinical Infectious Diseases* **25**:1374–1377.

Collinge, S. K. and C. Ray, 2006. Disease Ecology: Community Structure and Pathogen Dynamics. Oxford University Press, Oxford.

Conlan, A. J. K., P. Rohani, A. L. Lloyd, M. Keeling, and B. T. Grenfell, 2010. Resolving the impact of waiting time distributions on the persistence of measles. *Journal of The Royal Society Interface* **7**:623–640.

Danon, L., A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross, and M. C. Vernon, 2011. Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases* **2011**:284909.

De Benedictis, J., E. Chow-Shaffer, A. Costero, G. G. Clark, J. D. Edman, and T. W. Scott, 2003. Identification of the people from whom engorged Aedes aegypti took blood meals in Florida, Puerto Rico, using polymerase chain reaction-based DNA profiling. *The American journal of tropical medicine and hygiene* **68**:437–46.

De Castro, F. and B. Bolker, 2004. Mechanisms of disease-induced extinction. *Ecology Letters* **8**:117–126.

de Castro Medeiros, L. C., C. A. R. Castilho, C. Braga, W. V. de Souza, L. Regis, and A. M. V. Monteiro, 2011. Modeling the dynamic transmission of dengue fever: investigating disease persistence. *PLoS neglected tropical diseases* **5**:1–14.

Deredec, A. and F. Courchamp, 2003. Extinction thresholds in host-parasite dynamics. *Annales Zoologici Fennici* pages 115–130.

Diekmann, O., J. a. Heesterbeek, and J. a. Metz, 1990. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology* **28**:365–382.

Diekmann, O. and J. A. P. Heesterbeek, 2000. Mathematical epidemiology of infectious diseases : model building, analysis, and interpretation. Wiley series in mathematical and computational biology. John Wiley, Chichester, New York.

Diekmann, O., J. a. P. Heesterbeek, and M. G. Roberts, 2010. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society, Interface / the Royal Society* **7**:873–885.

Ditlevsen, S. and A. Samson, 2013. Stochastic Biomathematical Models, volume 2058 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Dye, C. and B. Williams, 1995. Nonlinearities in the dynamics of indirectly-transmitted infections (or, does having a vector make a difference?). In B. T. Grenfell and A. P. Dobson, editors, *Ecology of Infectious Diseases in Natural Populations*, pages 260–279. Cambridge University Press. Cambridge Books Online.

Effler, P. V., L. Pang, P. Kitsutani, V. Vorndam, M. Nakata, T. Ayers, J. Elm, T. Tom, P. Reiter, J. G. Rigau-Perez, J. M. Hayes, K. Mills, M. Napier, G. G. Clark, and D. J. Gubler, 2005. Dengue fever, Hawaii, 2001-2002. *Emerging infectious diseases* **11**:742–9.

Enserink, M., 2008. A Mosquito Goes Global. *Science* **320**:864–866.

Érdi, P. and G. Lente, 2014. Stochastic Chemical Kinetics. Springer Series in Synergetics. Springer New York, New York, NY.

Erlanger, T. E., J. Keiser, and J. Utzinger, 2008. Effect of dengue vector control interventions on entomological parameters in developing countries: a systematic review and meta-analysis. *Medical and veterinary entomology* **22**:203–221.

Esteva, L. and C. Vargas, 1998. Analysis of a dengue disease transmission model. *Mathematical Biosciences* **150**:131–151.

Esteva, L. and C. Vargas, 1999. A model for dengue disease with variable human population. *Journal of Mathematical Biology* **38**:220–240.

Esteva, L. and C. Vargas, 2000. Influence of vertical and mechanical transmission on the dynamics of dengue disease. *Mathematical biosciences* **167**:51–64.

European Centre for Disease Prevention and Control, 2011. Consultation on mosquito-borne disease transmission risk in Europe. Technical Report November, Stockholm.

European Centre for Disease Prevention and Control, 2012. The climatic suitability for dengue transmission in continental Europe. Technical report.

Fatimil, L., A. Mollah, S. Ahmed, and M. Rahman, 2003. Vertical transmission of dengue: first case report from Bangladesh. *SOUTHEAST ASIAN J TROP MED PUBLIC HEALTH* **34**:800–803.

Feng, Z. and J. X. Velasco-Hernández, 1997. Competitive exclusion in a vector-host model for the dengue fever. *Journal of mathematical biology* **35**:523–44.

Focks, D. A., R. J. Brenner, J. Hayes, and E. Daniels, 2000. Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. *The American journal of tropical medicine and hygiene* **62**:11–18.

Focks, D. a., E. Daniels, D. G. Haile, and J. E. Keesling, 1995. A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. *The American journal of tropical medicine and hygiene* **53**:489–506.

Focks, D. a., D. G. Haile, E. Daniels, and G. a. Mount, 1993a. Dynamic life table model for Aedes aegypti (Diptera: Culicidae): analysis of the literature and model development. *Journal of Medical Entomology* **30**:1018–1028.

Focks, D. a., D. G. Haile, E. Daniels, and G. a. Mount, 1993b. Dynamic life table model for Aedes aegypti (diptera: Culicidae): simulation results and validation. *Journal of medical entomology* **30**:1018–1028.

Funk, S., E. Gilad, and V. a. a. Jansen, 2010. Endemic disease, awareness, and local behavioural response. *Journal of theoretical biology* **264**:501–9.

Gillespie, D. T., 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**:2340–2361.

Grassly, N. and C. Fraser, 2006. Seasonal infectious disease epidemiology. *Proceedings of the Royal Society B* .

Gratz, N. G., 2004. Critical review of the vector status of Aedes albopictus. *Medical and veterinary entomology* **18**:215–27.

Grenfell, B. T., O. N. Bjø rnstad, and B. F. Finkenstädt, 2002. Dynamics of measles epidemics: scaling noise, determinism, and predictability with the TSIR model. *Ecological Monographs* **72**:185–202.

Gubler, D. J., 1998. Resurgent vector-borne diseases as a global health problem. *Emerg Infect Dis* **4**:442–450.

Günther, J., J. P. Martínez-Muñoz, D. G. Pérez-Ishiwara, and J. Salas-Benito, 2007. Evidence of vertical transmission of dengue virus in two endemic localities in the state of Oaxaca, Mexico. *Intervirology* **50**:347–52.

Hanski, I. and O. Ovaskainen, 2000. The metapopulation capacity of a fragmented landscape. *Nature* **404**:755–758.

Harrington, L. C., A. Fleisher, D. Ruiz-Moreno, F. Vermeylen, C. V. Wa, R. L. Poulson, J. D. Edman, J. M. Clark, J. W. Jones, S. Kitthawee, and T. W. Scott, 2014. Heterogeneous Feeding Patterns of the Dengue Vector, Aedes aegypti, on Individual Human Hosts in Rural Thailand. *PLoS Neglected Tropical Diseases* **8**:e3048.

Harrington, L. C., T. W. Scott, K. Lerdthusnee, R. C. Coleman, A. Costero, G. G. Clark, J. J. Jones, S. Kitthawee, P. Kittayapong, R. Sithiprasasna, and J. D. Edman, 2005. Dispersal of the dengue vector Aedes aegypti within and between rural communities. *The American journal of tropical medicine and hygiene* **72**:209–220.

Haydon, D. T., S. Cleaveland, L. H. Taylor, and M. K. Laurenson, 2002. Identifying Reservoirs of Infection: A Conceptual and Practical Challenge. *Emerging Infectious Diseases* **8**:1468–1473.

Heffernan, J., R. Smith, and L. Wahl, 2005. Perspectives on the basic reproductive ratio. *Journal of The Royal Society Interface* **2**:281–293.

Hii, Y. L., H. Zhu, N. Ng, L. C. Ng, and J. Rocklöv, 2012. Forecast of Dengue Incidence Using Temperature and Rainfall. *PLoS neglected tropical diseases* **6**:e1908.

Institute of Medicine, 2008. Vector-borne diseases: understanding the environmental, human health, and ecological connections. The National Academies Press, Washington, DC.

Jansen, C. C. and N. W. Beebe, 2010. The dengue vector Aedes aegypti: what comes next. *Microbes and infection / Institut Pasteur* **12**:272–9.

Johansson, M. a., F. Dominici, and G. E. Glass, 2009. Local and global effects of climate on dengue transmission in Puerto Rico. *PLoS neglected tropical diseases* **3**:e382.

Johansson, M. a., J. Hombach, and D. a. T. Cummings, 2011. Models of the impact of dengue vaccines: a review of current research and potential approaches. *Vaccine* **29**:5860–8.

Joshi, V., D. T. Mourya, and R. C. Sharma, 2002. Persistence of dengue-3 virus through transovarial transmission passage in successive generations of Aedes aegypti mosquitoes. *The American journal of tropical medicine and hygiene* **67**:158–61.

Joshi, V. and R. Sharma, 2001. Impact of vertically-transmitted dengue virus on viability of eggs of virus-inoculated Aedes aegypti. *Dengue Bulletin* **25**:103–106.

Kamgang, B., J. Y. Happi, P. Boisier, F. Njiokou, J.-P. Hervé, F. Simard, and C. Paupy, 2010. Geographic and ecological distribution of the dengue and chikungunya virus vectors Aedes aegypti and Aedes albopictus in three major Cameroonian towns. *Medical and veterinary entomology* **24**:132–41.

Keeling, M. and B. Grenfell, 1997. Disease extinction and community size: modeling the persistence of measles. *Science* **275**:65–67.

Keeling, M. J. and B. T. Grenfell, 1998. Effect of variability in infection period on the persistence and spatial spread of infectious diseases.

Kermack, W. O. and A. G. McKendrick, 1927. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **115**:700–721.

Kermack, W. O. and A. G. McKendrick, 1932. Contributions to the Mathematical Theory of Epidemics. II. The Problem of Endemicity. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **138**:55–83.

Kilpatrick, A. M. and S. Altizer, 2012. Disease ecology. *Nature education knowledge* **3**:55.

Koella, J. C., 1991. On the use of mathematical models of malaria transmission. *Acta Tropica* **49**:1–25.

Kraemer, M. U., M. E. Sinka, K. a. Duda, A. Q. Mylne, F. M. Shearer, C. M. Barker, C. G. Moore, R. G. Carvalho, G. E. Coelho, W. Van Bortel, G. Hendrickx, F. Schaffner, I. R. Elyazar, H.-J. Teng, O. J. Brady, J. P. Messina, D. M. Pigott, T. W. Scott, D. L. Smith, G. W. Wint, N. Golding, and S. I. Hay, 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. *eLife* **4**:1–18.

Kyle, J. L. and E. Harris, 2008. Global spread and persistence of dengue. *Annual review of microbiology* **62**:71–92.

Lambrechts, L., T. B. Knox, J. Wong, K. a. Liebman, R. G. Albright, and S. T. Stoddard, 2009. Shifting priorities in vector biology to improve control of vector-borne disease. *Tropical medicine & international health : TM & IH* **14**:1505–14.

Li, J., D. Blakeley, and R. J. Smith, 2011. The failure of R(0). *Computational and mathematical methods in medicine* **2011**:527610.

Liebman, K. a., S. T. Stoddard, A. C. Morrison, C. Rocha, S. Minnick, M. Sihuincha, K. L. Russell, J. G. Olson, P. J. Blair, D. M. Watts, T. Kochel, and T. W. Scott, 2012. Spatial dimensions of dengue virus transmission across interepidemic and epidemic periods in Iquitos, Peru (1999-2003). *PLoS neglected tropical diseases* **6**:e1472.

Liebman, K. A., S. T. Stoddard, R. C. Reiner, T. A. Perkins, H. Astete, M. Sihuincha, E. S. Halsey, T. J. Kochel, A. C. Morrison, and T. W. Scott, 2014. Determinants of heterogeneous blood feeding patterns by Aedes aegypti in Iquitos, Peru. *PLoS neglected tropical diseases* **8**:e2702.

Lloyd, A. L., 2001. Realistic Distributions of Infectious Periods in Epidemic Models: Changing Patterns of Persistence and Dynamics. *Theoretical Population Biology* **60**:59–71.

Lloyd, A. L., J. Zhang, and A. Root, 2007. Stochasticity and heterogeneity in host-vector models. *Journal of The Royal Society Interface* **4**:851–863.

Lloyd-Smith, J. O., P. C. Cross, C. J. Briggs, M. Daugherty, W. M. Getz, J. Latto, M. S. Sanchez, A. B. Smith, and A. Swei, 2005. Should we expect population thresholds for wildlife disease? *Trends in ecology & evolution* **20**:511–9.

Lotka, A. J., 1912. Quantitative Studies in Epidemiology. *Nature* **88**:497–498.

Maciel-De-Freitas, R., C. T. Codeço, and R. Lourenço-De-Oliveira, 2007. Body size-associated survival and dispersal rates of Aedes aegypti in Rio de Janeiro. *Medical and Veterinary Entomology* **21**:284–292.

Magori, K. and J. M. Drake, 2013. The Population Dynamics of Vector-borne Diseases. *Nature education knowledge* **4**.

Mahalingam, S., B. L. Herring, and S. B. Halstead, 2013. Call to Action for Dengue Vaccine Failure. *Emerging Infectious Diseases* **19**:1335–1337.

Mancy, R., 2015. Modelling persistence in spatially-explicit ecological and epidemiological systems. Ph.D. thesis.

Mancy, R., P. Prosser, and S. Rogers, 2013. Discrete and continuous time simulations of spatial ecological processes predict different final population sizes and interspecific competition outcomes. *Ecological Modelling* **259**:50–61.

Marini, F., B. Caputo, M. Pombi, G. Tarsitani, and a. della Torre, 2010. Study of Aedes albopictus dispersal in Rome, Italy, using sticky traps in mark-release-recapture experiments. *Medical and veterinary entomology* **24**:361–8.

Marino, S., I. B. Hogue, C. J. Ray, and D. E. Kirschner, 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of theoretical biology* **254**:178–96.

Martins, V. E. P., C. H. Alencar, M. T. Kamimura, F. M. de Carvalho Araújo, S. G. De Simone, R. F. Dutra, and M. I. F. Guedes, 2012. Occurrence of natural vertical transmission of dengue-2 and dengue-3 viruses in Aedes aegypti and Aedes albopictus in Fortaleza, Ceará, Brazil. *PloS one* **7**:1–9.

MATLAB, 2014. version 8.3.0 .532 (R2014a). The MathWorks Inc., Natick, Massachusetts, USA.

Mckay, M. D., R. J. Beckman, and W. J. Conover, 2000. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics* **42**:55–61.

McKirdy, S., R. Jones, and F. Nutter, 2002. Quantification of Yield Losses Caused by Barley yellow dwarf virus in Wheat and Oats. *Plant Dis.* **86**:769–773.

McMeniman, C. J., R. V. Lane, B. N. Cass, A. W. C. Fong, M. Sidhu, Y.-F. Wang, and S. L. O'Neill, 2009. Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. *Science (New York, N.Y.)* **323**:141–144.

Medlock, J. M., K. M. Hansford, F. Schaffner, V. Versteirt, G. Hendrickx, H. Zeller, and W. Van Bortel, 2012. A review of the invasive mosquitoes in Europe: ecology, public health risks, and control options. *Vector borne and zoonotic diseases (Larchmont, N.Y.)* **12**:435–47.

Moncayo, A. C., Z. Fernandez, D. Ortiz, M. Diallo, A. Sall, S. Hartman, C. T. Davis, L. Coffey, C. C. Mathiot, R. B. Tesh, and S. C. Weaver, 2004. Dengue emergence and adaptation to peridomestic mosquitoes. *Emerging infectious diseases* **10**:1790–1796.

Muir, L. E. and B. H. Kay, 1998. Aedes aegypti survival and dispersal estimated by mark-release-recapture in northern Australia. *The American journal of tropical medicine and hygiene* **58**:277–82.

Mulyatno, K. C., A. Yamanaka, S. Yotopranoto, and E. Konishi, 2012. Vertical transmission of dengue virus in Aedes aegypti collected in Surabaya, Indonesia, during 2008-2011. *Japanese journal of infectious diseases* **65**:274–276.

Newton, E. a. C. and P. Reiter, 1992. A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics. *American Journal of Tropical Medicine and Hygiene* **47**:709–720.

Nishiura, H., 2006. Mathematical and statistical analyses of the spread of dengue. *Dengue bulletin* **30**:51–67.

Nåsell, I., 1999. On the time to extinction in recurrent epidemics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**:309–330.

Nåsell, I., 2002. Stochastic models of some endemic infections. *Mathematical Biosciences* **179**:1–19.

Nåsell, I., 2005. A new look at the critical community size for childhood infections. *Theoretical population biology* **67**:203–16.

Nåsell, I., 2011. Extinction and Quasi-Stationarity in the Stochastic Logistic SIS Model, volume 2022 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Otero, M., D. H. Barmak, C. O. Dorso, H. G. Solari, and M. a. Natiello, 2011. Modeling dengue outbreaks. *Mathematical Biosciences* **232**:87–95.

Otero, M., N. Schweigmann, and H. G. Solari, 2008. A stochastic spatial dynamical model for Aedes aegypti. *Bulletin of Mathematical Biology* **70**:1297–1325.

Otero, M. and H. G. Solari, 2010. Stochastic eco-epidemiological model of dengue disease transmission by Aedes aegypti mosquito. *Mathematical Biosciences* **223**:32–46.

Otero, M., H. G. Solari, and N. Schweigmann, 2006. A stochastic population dynamics model for Aedes aegypti: Formulation and application to a city with temperate climate. *Bulletin of Mathematical Biology* **68**:1945–1974.

Pandey, A., A. Mubayi, and J. Medlock, 2013. Comparing vector-host and SIR models for dengue transmission. *Mathematical Biosciences* **246**:252–259.

Paupy, C., B. Ollomo, B. Kamgang, S. Moutailler, D. Rousset, M. Demanou, J.-P. Hervé, E. Leroy, and F. Simard, 2010. Comparative role of Aedes albopictus and Aedes aegypti in the emergence of Dengue and Chikungunya in central Africa. *Vector borne and zoonotic diseases (Larchmont, N.Y.)* **10**:259–266.

Peel, a. J., J. R. C. Pulliam, a. D. Luis, R. K. Plowright, T. J. O'Shea, D. T. S. Hayman, J. L. N. Wood, C. T. Webb, and O. Restif, 2014. The effect of seasonal birth pulses on pathogen persistence in wild mammal populations. *Proceedings. Biological sciences / The Royal Society* **281**.

Pinho, S. T. R., C. P. Ferreira, L. Esteva, F. R. Barreto, V. C. Morato e Silva, and M. G. L. Teixeira, 2010. Modelling the dynamics of dengue real epidemics. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **368**:5679–93.

Platt, K. B., K. J. Linthicum, K. S. Myint, B. L. Innis, K. Lerdthusnee, and D. W. Vaughn, 1997. Impact of dengue virus infection on feeding behavior of Aedes aegypti. *The American journal of tropical medicine and hygiene* **57**:119–125.

Polwiang, S., 2015. The seasonal reproduction number of dengue fever: impacts of climate on transmission. *PeerJ* **3**:e1069.

Ponlawat, A. and L. C. Harrington, 2005. Blood feeding patterns of Aedes aegypti and Aedes albopictus in Thailand. *Journal of medical entomology* **42**:844–9.

Reiner, R. C., T. A. Perkins, C. M. Barker, T. Niu, L. F. Chaves, A. M. Ellis, D. B. George, A. Le Menach, J. R. C. Pulliam, D. Bisanzio, C. Buckee, C. Chiyaka, D. a. T. Cummings, A. J. Garcia, M. L. Gatton, P. W. Gething, D. M. Hartley, G. Johnston, E. Y. Klein, E. Michael, S. W. Lindsay, A. L. Lloyd, D. M. Pigott, W. K. Reisen, N. Ruktanonchai, B. K. Singh, A. J. Tatem, U. Kitron, S. I. Hay, T. W. Scott, and D. L. Smith, 2013. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. *Journal of the Royal Society, Interface / the Royal Society* **10**:20120921.

Reiter, P., S. Lathrop, M. Bunning, B. Biggerstaff, D. Singer, T. Tiwari, L. Baber, M. Amador, J. Thirion, J. Hayes, C. Seca, J. Mendez, B. Ramirez, J. Robinson, J. Rawlings, V. Vorndam, S. Waterman, D. Gubler, G. Clark, and E. Hayes, 2003. Texas Lifestyle Limits Transmission of Dengue Virus. *Emerging Infectious Diseases* **9**:86–89.

Richards, S., S. Anderson, and B. Alto, 2012. Vector Competence of Aedes aegypti and Aedes albopictus (Diptera: Culicidae) for Dengue Virus in the Florida Keys. *Journal of medical entomology* **49**:942–946.

Richards, S. L., L. Ponnusamy, T. R. Unnasch, H. K. Hassan, and C. S. Apperson, 2006. Host-feeding patterns of Aedes albopictus (Diptera: Culicidae) in relation to availability of human and domestic animals in suburban landscapes of central North Carolina. *Journal of Medical Entomology* **43**:900–908.

Rodriguesa, H. S., M. T. T. Monteiro, D. F. M. Torres, and A. Zinober, 2012. Dengue disease, basic reproduction number and control. *International Journal of Computer Mathematics* **89**:334–346.

Ross, R., 1911. Some Quantitative Studies in Epidemiology. *Nature* **87**:466–467.

Saifur, R. G. M., A. A. Hassan, H. Dieng, R. C. Salmah, T. Satho, A. R. Saad, R. Enrique, and M. Vargas, 2012. Update on Temporal and Spatial Abundance of Dengue Vectors in Penang , Malaysia. *Journal of the American Mosquito Control Association* **28**:84–92.

Sanchez, M. A., & Blower, S. M., 1997. Uncertainty and sensitivity analysis of the basic reproductive rate: tuberculosis as an example. *American Journal of Epidemiology* **145**:1127–1137.

Service, M. W., 1980. Effects of wind on the behaviour and distribution of mosquitoes and blackflies. *International Journal of Biometeorology* **24**:347–353.

Simmons, C. P., J. J. Farrar, N. van Vinh Chau, and B. Wills, 2012. Dengue. *New England Journal of Medicine* **366**:1423–1432.

Smith, D. L., K. E. Battle, S. I. Hay, C. M. Barker, T. W. Scott, and F. E. McKenzie, 2012. Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS pathogens* **8**:e1002588.

Smith, D. L. and F. E. McKenzie, 2004. Statics and dynamics of malaria infection in Anopheles mosquitoes. *Malaria journal* **3**:13.

Smith, D. L., T. A. Perkins, R. C. Reiner, C. M. Barker, T. Niu, L. F. Chaves, A. M. Ellis, D. B. George, A. Le Menach, J. R. C. Pulliam, D. Bisanzio, C. Buckee, C. Chiyaka, D. A. T. Cummings, A. J. Garcia, M. L. Gatton, P. W. Gething, D. M. Hartley, G. Johnston, E. Y. Klein, E. Michael, A. L. Lloyd, D. M. Pigott, W. K. Reisen, N. Ruktanonchai, B. K. Singh, J. Stoller, A. J. Tatem, U. Kitron, H. C. J. Godfray, J. M. Cohen, S. I. Hay, and T. W. Scott, 2014. Recasting the theory of mosquito-borne pathogen transmission dynamics and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **108**:185–197.

Stedman, T. L., 2005. Stedman's medical dictionary for the health professions and nursing. Lippincott Williams & Wilkins Baltimore, MD, illustrated 5th ed. edition.

Stoddard, S. T., A. C. Morrison, G. M. Vazquez-Prokopec, V. Paz Soldan, T. J. Kochel, U. Kitron, J. P. Elder, and T. W. Scott, 2009. The role of human movement in the transmission of vector-borne pathogens. *PLoS neglected tropical diseases* **3**:1–9.

Swinton, J., M. Woolhouse, M. Begon, A. Dobson, E. Ferroglio, B. Grenfell, V. Guburti, R. Hails, J. Heesterbeek, A. Lavazza, M. Roberts, P. White, and K. Wilson, 2002. Microparasite transmission and persistence, pages 83–101. Oxford University Press.

Tan, P. and G. Rajasingam, 2008. Dengue infection in pregnancy: prevalence, vertical transmission, and pregnancy outcome. *Obstetrics & Gynecology* **111**:1111–1117.

Townson, H., M. B. Nathan, M. Zaim, P. Gullet, R. Bos, and M. Kindhauser, 2005. Exploiting the potential of vector control for disease prevention. *Bull World Health Organ* **83**:942–947.

Valerio, L. and F. Marini, 2010. Host-feeding patterns of Aedes albopictus (Diptera: Culicidae) in urban and rural contexts within Rome province, Italy. *Vector-Borne and Zoonotic Diseases* **10**:291–294.

Vega-Rua, A., K. Zouache, V. Caro, L. Diancourt, P. Delaunay, M. Grandadam, and A.-B. Failloux, 2013. High Efficiency of Temperate Aedes albopictus to Transmit Chikungunya and Dengue Viruses in the Southeast of France. *PLoS ONE* **8**:1–8.

Viana, M., R. Mancy, R. Biek, S. Cleaveland, P. C. Cross, J. O. Lloyd-Smith, and D. T. Haydon, 2014. Assembling evidence for identifying reservoirs of infection. *Trends in Ecology & Evolution* **29**:270–279.

Wearing, H. and P. Rohani, 2006. Ecological and immunological determinants of dengue epidemics. *Proceedings of the National Academy of Sciences of the United States of America* **103**:11802–11807.

WHO, 2004. The World Health Report 2004 - changing history. *World Health* **95**:96p.

WHO, 2009. Dengue: guidelines for diagnosis, treatment, prevention and control.

WHO, 2014. A Global Brief on Vector Borne Disease. Technical report.

World Health Organization, 2014. World malaria report 2014. Technical report.

Yang, H., H. Wei, and X. Li, 2010. Global stability of an epidemic model for vector-borne disease. *Journal of Systems Science and Complexity* **23**:279–292.

Yang, H. M., M. L. G. Macoris, K. C. Galvani, M. T. M. Andrighetti, and D. M. V. Wanderley, 2009. Assessing the effects of temperature on the population of Aedes aegypti, the vector of dengue. *Epidemiology and infection* **137**:1188–1202.