2016

Japanese Encephalitis: Assessing disease risk due to landscape factors at multiple scales

Julia E. Metelka
Wilfrid Laurier University, juliametelka@gmail.com

Japanese Encephalitis: Assessing disease risk due to landscape factors at multiple scales

by

Julia E. Metelka

A thesis
presented to Wilfrid Laurier University
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Geography

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Acknowledgements

First and foremost, I would like to thank my advisor, Colin Robertson. This thesis would not be possible without his guidance and support. Thank you for being patient, understanding, and patient! Secondly, I would like to thank Derek Robinson. Derek provided me with valuable insight on the topic of agent-based modelling. I would also like to thank all the members of the Spatial Lab - Kara, Haydn, Lauren, Ben, Courtney, and Clara - who provided me with support while completing this thesis. I would also like to thank the members of the National Zoonoses and Food Hygiene Research Centre, located in Kathmandu, Nepal. Their knowledge of Japanese Encephalitis in Nepal and hospitality at the research centre, were of great importance for this project.

Dedication

Dedicated to Stephen, Barbara, Andy, Jessica, and Lexie.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Japanese Encephalitis (JE) is the leading cause of viral encephalitis in Southeast and East Asia (Mackenzie et al., 2004; Misra and Kalita, 2010; Campbell et al., 2011; CDC, 2013). It is estimated that approximately 67,900 cases of JE occur annually (CDC, 2013). The case fatality rate is 20–30% and approximately 30–40% of survivors suffer from permanent neurological sequelae, which results from to damage of the central nervous system (CDC, 2013; World Health Organization, 2014b). The disease is primarily acquired by children less than 15 years of age and is more commonly found in rural and agricultural areas (Mackenzie et al., 2004; Campbell et al., 2011; CDC, 2013; World Health Organization, 2014b). LaBeaud et al. (2011) estimated that 57 per cent of the global population is at-risk for contracting JE and that the 2005 disability-adjusted life years (DALYs) is between 107,435 and 1,859,170, thus causing significant burden on the global population.

JE is a vector-borne disease that is transmitted by mosquitoes. Infection and disease are caused by the bite of a mosquito that is infected with the JE virus (JEV). Symptoms of severe JE cases include; rapid onset of high fever, headache, neck stiffness, disorientation, coma, seizures, spastic paralysis, and death (Misra and Kalita, 2010; World Health Organization, 2014b). However, not all incidents of infection result in full-blown JE symptoms. Most JEV infections are either asymptomatic or result in non-specific flu-like illness, which can influence accurate reporting and surveillance due to perceived insignificance of an infection. The estimation of the rate of asymptomatic infections vs. symptomatic infections is difficult to determine and

has varied considerably in the literature (Konishi and Suzuki, 2002; Solomon, 2004). Van den Hurk et al. (2009) stated that the ratio of asymptomatic to symptomatic infections ranges in the literature from 1 in 50 to 1 in 1000. However, the World Health Organization estimates that approximately 1 in 250 infections result in severe disease (full-blown JE) (World Health Organization, 2014b). These ratios can vary geographically and be influenced by the history of JEV prevalence in a particular region. There is currently no antiviral treatment for those with JE. Those who exhibit severe JE symptoms receive supportive care, as they require feeding, assistance breathing, and anti-convulsants for seizures. Thus, identifying and monitoring high risk areas is imperative for public health policy, prevention strategies, and outreach.

Cases of JE are reported to the World Health Organization's Vaccine-Preventable Disease programme (World Health Organization: Country Office for Nepal, 2011). Healthcare providers are responsible for reporting positive cases of JE (World Health Organization, 2014a). However, ensuring the completeness of surveillance data accuracy has been difficult. For example, Kakkar et al. (2013) examined recorded JE cases for the years 2011 and 2012 in the Kushinagar District in Uttar Pradesh, India and found that 216 out of 812 fields in the surveillance forms had to be modified in some way. The majority of fields that were either incorrect or left blank were from 'Date of Symptom Onset' and 'Date of Fever Onset' fields (Kakkar et al., 2013). Kakkar et al. (2013) concluded the quality of this surveillance data was so low that any sort of prevention/control measures could not be based on it. Not only is data collection an issue for accurate surveillance, but the likelihood that JE is reported depends on factors such as access to healthcare and preference for traditional healers (Akiba et al., 2001). Akiba et al. (2001) conducted social surveys in Nepal and found that the majority of rural village dwellers chose a traditional faith healer as their primary preference for healthcare. Akiba et al. (2001) suggested that this finding may suggest why actual reported cases are more common in urban areas as opposed to the more risky (higher mosquito populations) rural areas.

Although surveillance data plays a crucial role in disease control and prevention strategies, it does not portray an exhaustive representation of actual case distribution. Kakkar et al. (2013) demonstrated how selected surveillance records may be incomplete and/or inaccurate and CDC (2013) highlights how, in most countries within the JE geographic distribution, surveillance

sites do not provide full geographic coverage. Disease risk mapping and the identification of high priority areas that incorporate transmission cycle and risk factor knowledge would be of use to fill in this gap. Such methods would better inform control and prevention measures.

The likelihood of contracting JE, like other vector-borne diseases, can be increased by certain landscape factors, for example, the presence of virus amplifying hosts and/or rice paddy fields. JE risk is also influenced by behavioural factors, for example, the amount of time an individual spends in or near the previously mentioned landscape risk factors. JE risk can also be reduced at an individual level if one engages in mosquito avoidance behaviours (e.g. the use of insect repellant and/or window screens). Modelling JE risk can thus be taken from two perspectives; 1) national-, or regional-scale, which allows for population generalizations, and 2) local-, or individual-scale, which allows for individual characteristics to influence disease risk.

A prevailing methodology for mapping disease transmission risk at a small geographic scale has been that of ecological niche modelling. Ecological niche modelling is used to map the environmental covariates associated with a particular disease. With this method, data for large regions can easily be incorporated into a model, highlighting high-risk regions over a large geographic area. In the example of JEV, as in many vector-borne diseases, variables such as climate, landcover, and virus host populations are crucial components of the disease transmission cycle that dictate the spatial and temporal distribution of human cases of the disease. The overarching goal of ecological niche modelling is to identify the geographic disease distribution potential. The concept of ecological niche modelling makes perfect sense for identifying high risk (or low risk) regions, and there has been a plethora of research looking at the geographic distribution of JEV's mosquito vector using these techniques (Hart, 2010; Miller et al., 2012; Sallam et al., 2013; Wang et al., 2014). However, Peterson (2015) suggested that ecological niche modelling may be too simple in some cases and that it fails to examine host population and inter-species interactions. Peterson (2015) stated that modelling complex disease systems, which involve a number of hosts and vectors, has been a challenge in these component-based modelling methods.

Peterson (2015) contrasted ecological niche modelling and other component-based methods

with *black-box* methods. The *black-box'* approach to disease risk mapping involves analyzing solely the spatial pattern of human occurrences of a particular disease. In other words, the results (emergence) of the underlying interactions between environmental variables, landscape characteristics, and inter-species interactions, which lead to increased risk of acquiring said disease. Black-box approaches are useful when knowledge of the disease transmission system and its components are of low-quality or non-existent. However, when surveillance data is inconsistent or incomplete as demonstrated and suggested by previous research, *black-box* methods, or the analysis of reported human cases of a disease, do not capture the entirety of a disease transmission system.

Chapter 3 presents a method for mapping potential JE risk in the near future. Like ecological niche modelling, we present a method that highlights regions of potential risk at a small geographic scale (i.e. large area). However, unlike ecological niche modelling, this Chapter focuses on two primary regional components of JE risk: virus amplifying host populations and the at-risk human population. Motivation for this approach stemmed from being able to estimate future high-risk regions allowing for population generalizations at this scale. Although such generalizations are necessary at this level of scale, other methods that incorporate individual factors are important when looking at local-level risk.

Another prevailing methodology for disease transmission is the SIR (Susceptible-Infected-Recovered) model (Kermack and McKendrick, 1927). This model, and its many variations are used to examine disease transmission within closed, homogeneous populations. Although this method for analyzing disease spread is a dominant method for studying disease transmission, it fails to incorporate critical characteristics for looking at JEV spread. Firstly, SIR models do not account for individual-based characteristics of a population. Populations within each compartment are assumed to be homogeneous, with transmission rates dependent on various set parameters. In other words, the output of an SIR model is the result of a single input value (e.g. population), and a set of static parameters. Secondly, SIR models do not traditionally incorporate spatial relationships, a crucial element of the JEV transmission cycle, although some previous research have applied SIR models to spatial problems (Keeling, 1999; Fuks and Lawniczak, 2001).

A prevailing method for representing and modelling systems that involve individual characteristics is that of individual-based modelling (Grimm et al., 2006). This method evolved from cellular automata models (Ulam, 1950; Wolfram, 1984; Pesavento, 1995), which involve a grid of cells that alter between various states according to a set of rules. At the individual-level, these types of models are able to represent individual characteristics, such as those important in disease spread (e.g. age, occupation, sex). Rules that govern individual-level interactions with the landscape may result in emergent phenomena that are typical in the spatial and temporal distribution of disease outbreak. A study of particular interest, which combines the use of SIR compartment modelling, individual-based modelling, and geographic information systems is that of Hailegiorgis and Crooks (2012). Hailegiorgis and Crooks (2012) use a variation of the SIR, the Susceptible-Exposed-Infected-Recovered (SEIR) model, in an agent-based model for cholera spread in refugee camps. This approach allowed for effective representation of disease characteristics. These methods were appropriate for modelling disease spread at this geographic scale. We propose that such methods may identify regions with higher case counts than what is apparent in surveillance data.

Chapter 4 presents an individual-based model, which can be thought of as a combined-approach to disease risk mapping for JE. In our model we implement the basic components of an ecological niche model to show the potential geographic distribution of JEV's vector, *Culex* mosquitoes, using multiple environmental datasets. We then employ a variation of a Susceptible-Exposed-Infected-Recovered (SEIR) model to capture the human and host dynamics and human-environment interactions. We take an individual-, or agent-based, modelling approach, as it captures age and other individual characteristics that may not be represented as well using other modelling techniques. It is hoped that this model will serve as a simple representation of the JEV-human system and that additional complexity, such as human behaviours regarding mosquito avoidance, can be easily integrated into the model.

Previous research has implemented agent-based modelling to examine vector-borne disease risk. Linard et al. (2009) examined malaria risk in southern France by creating an agent-based model to simulate the spatial and temporal dynamics of mosquito biting rate in relation to landscape characteristics. Gu and Novak (2009) also developed an agent-based model to assess

malaria transmission, but instead of humans being the primary agent, they programmed the behaviour of mosquitoes to examine their movement. Their model was then used to assess mosquito-control measures (Gu and Novak, 2009). de Almeida et al. (2010) also using agent-based modelling to examine the mosquito species *Aedes aegypti* populations and behaviours, which can affect the risk of vector-borne disease transmission. Alexander et al. (2012) highlighted the importance of agent-based modelling in vector-borne and zoonotic diseases, as they are able to systematically incorporate heterogenous populations and various types as agents (e.g. virus hosts, vectors, and vulnerable human populations).

By focusing on disease transmission risk at the individual-level, we develop a simulation that is able to reproduce a similar temporal and spatial distribution as JE surveillance data from the World Health Organization. Human-landscape interactions play an important role in an individual's risk of contracting JE, thus modelling at the same scale is beneficial for finding ways to reduce individual risk. Future work will look at the outcomes of human incidence of JE under various scenarios, such as the increase in the number or change in the spatial distribution of pig farms, increasing temperature due to climate change, and the implementation of vaccination programs.

Both methodologies presented in this thesis identify high-risk areas for contracting the Japanese Encephalitis virus. Countries that have little resources for managing vector-borne diseases will be able to prioritize by focusing their resources on these high-risk regions. For example, irrigation practices that inhibit mosquito growth can be taught and adopted, or educational programs focusing on the reduction of vector-borne disease spread for pig farmers can be implemented in these areas, rather than nation-wide.

Objective 1: Estimating Small-scale Risk

The primary objective of Chapter 3 is to identify future regions of potentially high-risk for JE within the current geographic distribution of the virus. Chapter 3 presents a method for estimating future trends of the incidence of Japanese Encephalitis at a multi-national scale. By taking into account human population trends at a sub-national level, a set of at-risk human

populations are presented. Future trends of the virus's host population are also estimated at the sub-national level, using livestock statistics and projecting forward. Potential future regions of high case counts are thus those with a high at-risk human population and high host populations.

Objective 2: Estimating Large-scale Risk

The primary objective of Chapter 4 is to develop a model that reproduces the spatial and temporal distribution of JE cases in relation to known landscape risk factors and weather patterns. Chapter 4 presents a method for examining the risk of acquiring Japanese Encephalitis at the individual level. Landscape and other environmental factors are taken into account to create an estimated "risk surface" of the Kathmandu Valley, in Nepal. The primary research question this Chapter asks if it is possible to develop a model that takes into account these landscape risk factors to reproduce the spatial and temporal dynamics of the reported JE surveillance data obtained from the World Health Organization. Considering the similarities shared between many vector-borne diseases, it is proposed that such an individual-based method for looking at disease occurrence can be applied to other vector-borne diseases, such as the West Nile Virus, Dengue, malaria, etc..

Chapter 2

Background Information on Japanese Encephalitis

This chapter introduces the JEV's transmission cycle and the various landscape factors that are associated with increased disease risk. The aim of this background information is to emphasize the complexities of this system and how it is intertwined with various human activities (land use and land cover change, mosquito avoidance behaviours, occupation, etc.). A tool that is able to model the transmission cycle of JEV can be used to analyze various scenarios, such as changes in population distribution (of both human and virus host), increases in temperature and altered precipitation patterns due to climate change, and the effects of vaccination programs. The sections of this chapter summarize the primary processes that influence JE risk, specifically. However, it is emphasized that the methods presented in Chapter 3 and Chapter 4 could be applied to the analysis of other vector-borne diseases, as the underlying processes and reliance on the landscape to dictate patterns of disease are roughly similar.

2.1 Transmission Cycle

JEV exists within a zoonotic transmission cycle between animal hosts and insect vectors. Humans are dead-end hosts for the virus, and unlike dengue fever, humans are unable to pass along the virus to uninfected mosquitoes. The most important species within the JEV transmission cycle are those that are able to pass along the virus to uninfected vectors, these

include species from families *Suidae* (swine, both wild and domesticated), *Ardeidae* (wading birds, such as bitterns, herons, and egrets), and mosquitoes from the genus *Culex* (Mackenzie et al., 2004; Erlanger et al., 2009; Misra and Kalita, 2010). Although many species may become infected with JEV (horses, sheep, goats, etc.), viral titers are not high enough in the bloodstream of these species and are thus dead-end hosts.

Swine are considered to be an amplification host of JEV. Van den Hurk et al. (2009) summarized five main reasons for this amplification role of pigs in the transmission cycle of the JEV. These are described as follows; (1) there is a high natural infection rate of pigs, (2) pigs have a high viremia of the virus in the blood stream, (3) viremia in pigs remains high enough for transmission to mosquitoes for about 4 days, (4) pigs are highly prone to be fed upon by mosquitoes, and (5) pigs have a high birth rate, which provides new sources of virus-susceptible pigs each year. Pigs are also frequently located near rice paddy fields or in peri-urban areas where they are close to an at-risk human population (Dhakal et al., 2012, 2014).

Approximately 90 species of birds, domestic and wild, display JEV viremia (Van den Hurk et al., 2009). Of these species, *Ardeidae* play a significant role in virus transmission and are considered to be the primary enzootic hosts of JEV as their habitat overlaps with that of *Culex* mosquitoes (Van den Hurk et al., 2009). Weaver and Barrett (2004) indicated that populations of cattle egrets (*Bubulus ibis coromandu*) increased during the 19th century due to changing agricultural practices and it has been suggested that this has affected the evolution and spread of the five main JEV genotypes. *Ardeidae* birds frequent rice paddy fields to feed on small aquatic animals that reside there (Sibley, 2000). Wading birds become infected with the JEV after they are bitten by a mosquito carrying the virus. After initial infection, *Ardeidae* species have high enough viral titers in their bloodstream to amplify the virus. However, these birds become immune to JEV after virus-specific antibodies have developed, no longer playing a role in the transmission cycle.

Culex tritaeniorhynchus and *Culex vishnui* are the primary vector of JEV (World Health Organization, 2015b; Misra and Kalita, 2010). Although it has been stated that other species are known to carry the virus (Solomon et al., 2003). *Culex* mosquitoes lay rafts of eggs that require clean, stagnant waterbodies (Rozendaal, 1997). *Culex* require temperate to tropical climates

for their life-cycle. Females require the protein obtained from a blood meal to lay their eggs. Eggs hatch after 2-3 days as they enter the larval stage of their life cycle. They enter the pupa stage after about 4-7 days, spending most of their time at the water surface. This pupa stage lasts about 1-3 days and then the adult stage is reached. The entire life-cycle process can range between 7 and 14 day, depending on environmental conditions and food availability. Warmer temperatures typically speed up the life-cycle process between egg-laying and the transition to adult (Rozendaal, 1997).

2.2 Environmental Conditions that Contribute to Japanese Encephalitis Transmission

JEV is a member of the *Flavivirus* genus in the family *Flavivridae*. Most viruses of the genus *Flavivirus* are transmitted via arthropod insects and are thus termed *arboviruses*. Similar *Flaviviruses* include; West Nile, dengue fever, St. Louis, tick-borne encephalitis, and several other such viruses that may cause encephalitis in humans. JEV is the leading cause of viral encephalitis in Asia (World Health Organization, 2014b), causing approximately 67,900 human cases a year.

The current geographic distribution of JE is shown in Figure 2.1. However, its distribution has the potential to expand. Hanna et al. (1996) showed how the virus recently crossed the Torres Strait to reach Australia and Nemeth et al. (2010) examined the seroprevalence of potential JEV antibodies in non-native birds on two islands in Hawaii, highlighting the potential for spread to the Americas. Many factors, acting at various scales, contribute to the possibility of JEV's geographic expansion.

Rice agriculture and pig farming are two common practices that have been associated with heightened JE risk in Asia (Dapeng et al., 1995; Rayamajhi et al., 2007; Erlanger et al., 2009; Misra and Kalita, 2010; Singh et al., 2012; Le Flohic et al., 2013). Rice paddy fields provide breeding habitat for *Culex* mosquitoes and *Culex* populations are more abundant within proximity to paddy fields. Rice paddy fields also provide abundant feeding territory for wading birds, JEV's natural reservoir. In Asia, pig farming is primarily conducted in peri-urban areas (Thorpe

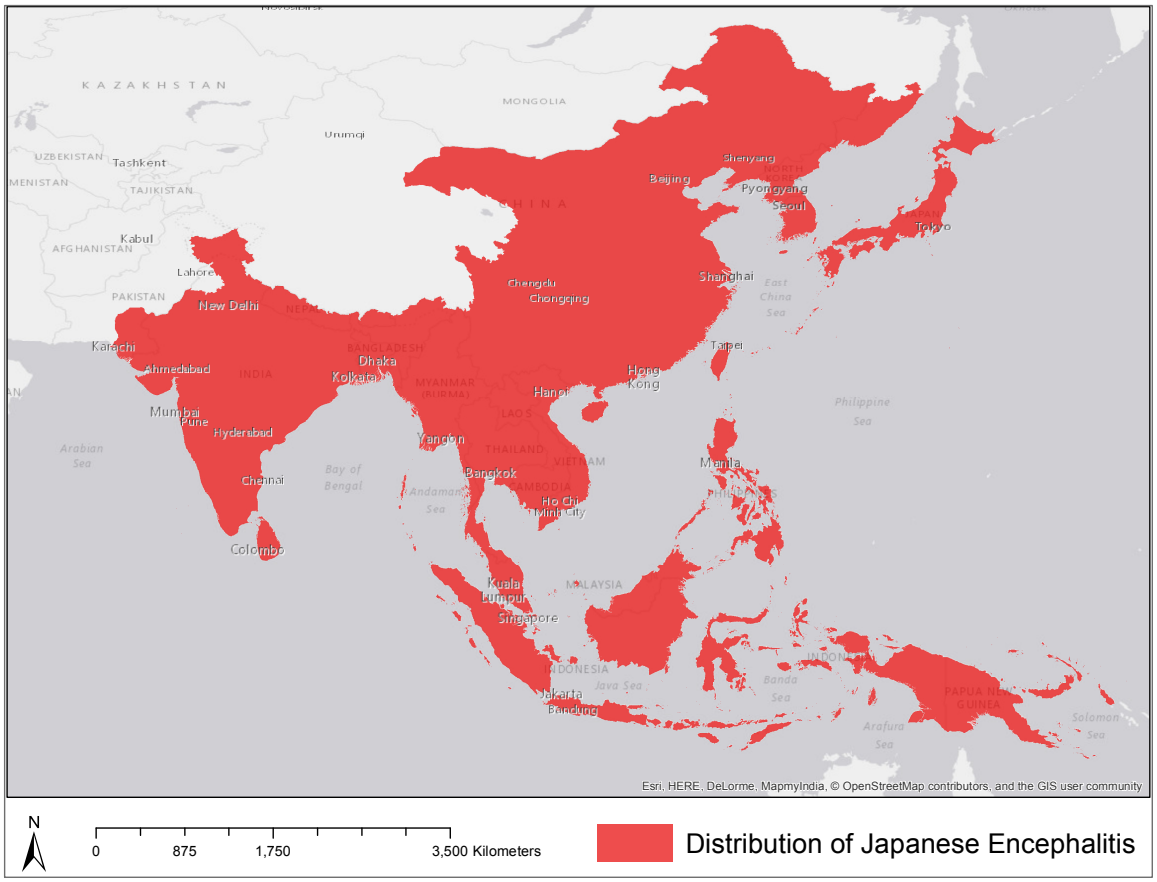


Figure 2.1: The current geographic distribution of Japanese Encephalitis.

and Jemaneh, 2008), or in rural areas, where there is also a higher probability of rice farming located nearby. These conditions influence the endemic nature of JEV in most tropic regions.

The population in Asia is expected to increase from 4.342 billion in 2012 to 5.164 billion by 2050 (Department of Economic and Social Affiars: Population Division of the United Nations, 2014), influencing land-use and land cover change as well as increasing the at-risk population. Land-use/cover change, especially agricultural expansion and increasing swine livestock numbers are expected to intensify as population increases (Coker et al., 2011). Coker et al. (2011) stated that during the period 1990-2008, agricultural land increased in area by over 8% in Southeast Asia. Pig-farming is also intensifying in Asia, increasing the number of potential virus amplifying hosts in many Asian countries. Erlanger et al. (2009) estimated that between 1990 and 2005, Cambodia, China, South Korea, Laos, Myanmar, Nepal, the Philippines, Sri Lanka, Thailand, and Vietnam all experienced increased pork production by between 12 and 381%. Such environmental factors will undoubtedly lead to an increased amount of JE

cases unless otherwise managed.

Figure 2.2 (p. 18) displays the dynamic relationship between human activities/processes and factors that contribute to JEV transmission. Human activities, such as population growth and internal migration patterns, affect the rate of urbanization and peri-urban and rural agricultural activities, potentially increasing the presences of landscape characteristics that contribute to JE risk (e.g. pig livestock counts and rice paddy fields). These processes can then affect mosquito populations by expanding habitat area and fragmenting the landscape (Lambin et al., 2010). It quickly becomes apparent that the JEV exists within a complex, dynamic system in which human and environment interactions play fundamental roles. Since these landscape characteristics are geographically specific, our current research looks at the Kathmandu Valley in Nepal as a case example for the human-environment system.

2.3 Pathogenesis of Japanese Encephalitis Virus

The pathogenesis, the process by which a pathogen (i.e. the JE virus) causes a diseased state (JE) in an individual, of JEV is largely undetermined. Many cases of JEV infection are either asymptomatic or result in non-specific flu-like illness. Very few infections actually result in full-blown incidents of JE. Van den Hurk et al. (2009) stated that the range of asymptomatic to symptomatic cases varies considerably in the literature, from about 1 in 50 to 1 in 1000. However, the exact mechanisms of pathogenesis are not clear (Turtle et al., 2012).

Previous work by Desai et al. (1997) explored the pathogenesis of JEV and co-infections with cysticercosis. Cysticercosis is a parasitic tissue infection cause by larval cysts created by the tapeworm *Taenia solium.* This parasitic disease also resides within a zoonotic cycle between humans and pigs. Human hosts to *Taenia solium* excrete the tapeworm eggs in their feces. When fecal matter reaches rivers and streams, parasite eggs are potentially consumed by pigs inhabiting rivers embankment areas. Infected pigs then manifest cysts throughout their bodies. Larval cysts can then be consumed by human if infected pig meat is consumed undercooked, repeating the transmission cycle of *Taenia solium.*

Desai et al. (1997) found that full-blown encephalitis syndrome, caused by JEV, was more

likely in individuals who were also infected with cysticercosis. Out of 163 sampled confirmed JE cases, 61 were infected with cysticercosis. Desai et al. (1997) suggested that cysticercosis affects the blood-brain barrier, making it easier for JEV to cross and cause encephalitis.

Introduction of JEV into the body elicits an immune response from both the innate and the adaptive immune system. Since many JEV infections do not result in serious illness, the adaptive immune system develops JEV antibodies which recognize the virus for future battles. The combination of non-serious illness and the adaptive immune system response is the reason why younger age groups are more susceptible to contracting serious illness. These age groups have had less time to develop adaptive immunity to the JE virus.

For reasons discussed above, the model we have developed does not differentiate between individual-level probabilities of becoming infected with JEV and contracting full-blown JE symptoms. The likelihood of developing JE, in our model, is solely based on encounters with landscape risk factors, since the precise mechanisms of individually-based likelihood of contracting the disease is unknown.

2.4 Modelling Vector-borne Diseases

Kermack and McKendrick (1927) developed what is known to be the standard methodology for epidemiological modelling. The SIR model, which stands for 'Susceptible-Infected-Recovered', is a deterministic model in which populations in each compartment (either the susceptible, infected, or recovered compartments) are determined by a set of differential equations (Kermack and McKendrick, 1927). There are many variations of the SIR model, used to capture the dynamics of various types of diseases and their spread. The SEIR model, which stands for 'Susceptible-Exposed-Infected-Recovered', accounts for the incubation, or latent period that many pathogens exhibit. The 'Exposed' compartment accounts for this time lag between when individuals are first exposed to a pathogen, and when that individual is determined to be in the compartment 'Infected', and can then spread that disease.

SIR and related models, such as the SEIR model, have been used extensively to study human-transmittable diseases, some examples include; Ebola (Lekone and Finkenstädt, 2006),

influenza (Mills et al., 2004), measles (Bjørnstad et al., 2002), and Chickenpox (Deguen et al., 2000). There are also countless variations of the SIR model used to study hypothetical situations of general disease dynamics; (Beretta and Takeuchi, 1995; Yu et al., 2009), and even to study the spread of rumours (Zhao et al., 2013) and traffic congestion (Wu et al., 2004).

Although SIR models, and their many variations, are well-used for studying disease transmission, they possess inherent characteristics that make them difficult to apply to spatially explicit scenarios. Compartments in SIR models are used to represent a single value for the number of individuals in the corresponding state (either Susceptible, Infected, or Removed). Compartments are unable to capture the heterogeneity of realistic human populations, such as age, sex, and other characteristics that may influence the likelihood of disease transmission. Because of this, they are best used for human-to-human spread diseases that are only dependent upon exposure to an infected individual.

Another shortfall of SIR models is that they do not account for spatial interactions explicitly. Many diseases are highly dependent upon the environment and/or human-animal interactions. For example, Jones et al. (2008) stated that 60% of all emerging infectious diseases are zoonotic, and World Health Organization (2015c) stated that there are over 200 zoonotic diseases. Zoonoses are diseases that are transmittable from animal to human, and whether domesticated or wild, animals play a critical role in the disease transmission system.

The distribution of animal populations is highly dependent on landscape characteristics. Recent work by Rodriguez-Meza (2012) utilized an SIR model to represent the spatial and temporal dynamics of the A/H1N1 virus pandemic in Mexico during the year 2009. They adjusted the SIR model to account for the spatial behaviour of infected individuals. Although, this is a novel contribution to the field of SIR modelling, Rodriguez-Meza (2012) do not vary the characteristics of space, but rather use space and proximity to determine infection spread. Vector-borne disease, such as dengue fever, lyme disease, malaria, and JE are zoonotic diseases that require an insect vector to transfer a pathogen from an infected animal to a human. This in-between step presents a hurdle for SIR models to represent this type of disease transmission system.

Macal and North (2008) defined an agent as 'an identifiable, discrete, or modular, individual

with a set of characteristics and rules governing its behaviours and decision-making capability.' Railsback and Grimm (2012) stated that agents are able interact with other agents, they change, have different stages as they progress through time. Agents also have the potential to make autonomous, adaptive decisions in order to pursue their objectives. Examples of types of agents include; humans, households, businesses, various species of wildlife, institutions or anything that pursues a certain goal (Railsback and Grimm, 2012). Agent-based models simulate these traits of agents to examine a specific or various outcomes.

The ability to model complex systems has become much more feasible in the past decade. Computational power and the capability to gather and store data has increased remarkably. There are also many open-source programs for ABM simulations that are now available. Almost anyone can download and learn software such as NetLogo, which comes with tutorials many demonstration models (Wilensky, 1999). This has facilitated the learning of ABM simulations.

The use of agent-based modelling (ABM) in epidemiology is extensive (Eubank et al., 2004; Fraser et al., 2004; Germann et al., 2006; Balcan et al., 2009). However, the research evaluating the spread of vector-borne diseases using ABMs is not as thorough. ABMs have been used to simulate interactions between people and how these can change the landscape (Deadman et al., 2004; Brown and Robinson, 2011; Murray-Rust et al., 2014), but focus on how land-used/cover change affects mosquito-breeding. The remainder of this section of this paper will examine some studies that have used ABMs to study vector-borne diseases, such as malaria.

Linard et al. (2009) used an ABM to assess the risk of reemergence of malaria in southern France. This region of France is home to the mosquito genus *Anopheles*, which have the potential to transmit malaria. Currently, the south of France is malaria-free, but due to the presence of *Anopheles* mosquitoes, there is a possible risk of malaria moving into this region. To evaluate this risk, Linard et al. (2009) utilize an ABM to represent the factors involved in malaria transmission.

Mosquitoes and humans were the agents in the ABM produced by Linard et al. (2009). They evaluated the contact rate between people and vectors in order to predict the re-emergence of malaria, which varies spatially. They also determined the number of people working (seasonal or regular workers), touring, or enjoying other leisure activities in mosquito habitat areas.

These human activities vary through space and time and are then susceptible to varying vector exposure risks (Linard et al., 2009). The purpose of the model developed by Linard et al. (2009) was to simulate vector-human contacts.

Although the study by Linard et al. (2009) is incredibly useful for understanding these interactions, the human impact on the environment was not considered. The remaining sections of this review will examine human-landscape interactions and how they may lead to the risk of vector-borne diseases. A special focus will the on the disease, Japanese Encephalitis.

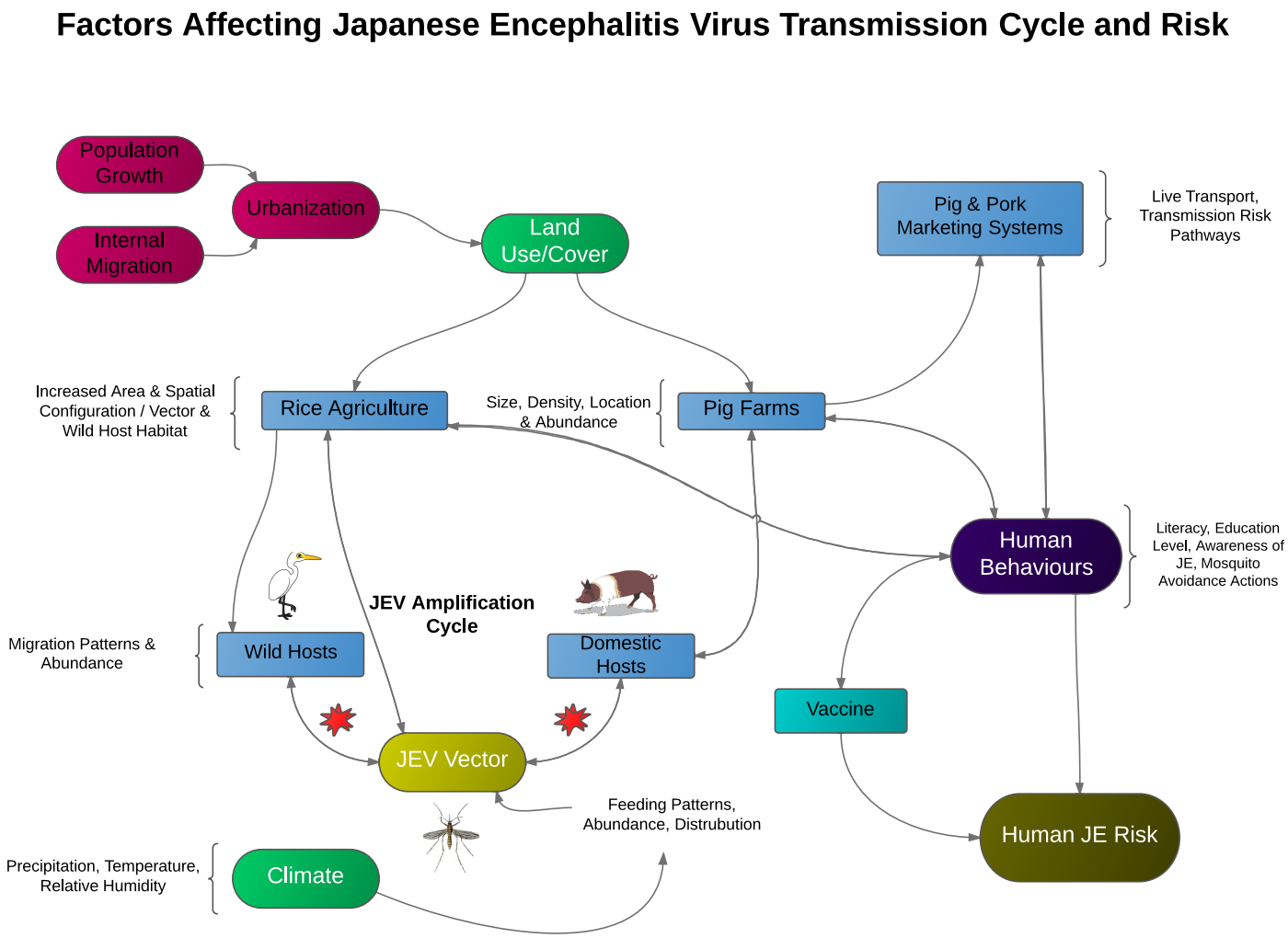


Figure 2.2: Flow diagram depicting factors that affect the JE transmission cycle. Population dynamic processes are shown in red, environmental factors shown in green, entities involved in the JEV amplification cycle shown in blue and yellow (vector), human behaviours in purple, and preventative measures (vaccine) in turquoise.

Chapter 3

Japanese Encephalitis: Estimating Future Trends in Asia

3.1 Introduction

Japanese Encephalitis (JE) is the leading cause of viral encephalitis in Asia (Mackenzie et al., 2004; Campbell et al., 2011). It is estimated that approximately 67,900 cases occur annually (Centers for Disease Control and Prevention, 2013). The case fatality rate is 20-30% and 30-40% of survivors suffer from permanent neurological sequelae (Centers for Disease Control and Prevention, 2013). The disease is primarily acquired by children less than 15 years of age and is historically more common in rural and agricultural areas (Mackenzie et al., 2004; Campbell et al., 2011; Centers for Disease Control and Prevention, 2013).

Confirmed cases of JE are reported to the World Health Organization through the programme for vaccine-preventable diseases (http://apps.who.int/immunization_monitoring/en/). Surveillance is absolutely necessary to guide immunization programs, target surveillance resources, set priorities, and can serve as an early warning system for identifying public health emergencies (World Health Organization, 2014c). However, most human cases of JE are asymptomatic and go unreported, making the spatial distribution of the virus difficult to estimate.

The World Health Organization (2014c) estimates that approximately 1 in 250 of those who

acquire the JE virus display encephalitis symptoms. For those that do experience disease, there is significant variation in health-seeking behaviours across Asia, which distort the sensitivity of disease surveillance systems. Factors such as the structure of the health care system (centralized or not), physical access to health care, propensity to adopt traditional and/or care services outside of the government system all vary across Asia and merge to obscure understanding of the disease burden due to JE. For example, Akiba et al. (2001) noted that many people in Nepal prefer traditional healers/medicine over hospitals, resulting in many unreported cases. Akiba et al. (2001) also mentioned that this preference for traditional healers was more prominent among poor and disadvantaged sectors of the population. There is therefore need to examine not only the reported cases of disease, but also the distribution of disease risk factors when estimating risk, developing health policy, and anticipating future changes.

There are several major contributing factors that lead to high-risk areas for JE. First, landscape factors play an important role in vector-borne disease transmission. Increased area dedicated to land-use types that contribute to vector habitat promote virus transmission and risk to humans. Other landscape characteristics, such as fragmentation and heterogeneity, can also increase risk of vector-borne disease (Lambin et al., 2010). Second, susceptible populations mixed with pig populations are strongly linked to JE risk. Pigs are considered to be an amplifying host for the JE virus. If a pig is infected with the virus, the probability that there will be infected mosquitoes in their vicinity is significantly increased. Finally, Age is also linked to JE, as the virus is typically associated with disease in children for areas where it is endemic. However, where childhood immunization programs exist, risk is approximately equal for adults and children (Wu et al., 1999; Sohn, 2000).

The Morbidity and Mortality Weekly Report conducted by Centers for Disease Control and Prevention (2013) summarized the status of JE surveillance programs in risk areas. The report concluded that as of 2012, 10 of 25 countries with JE risk conducted nation-wide surveillance programs. They reported that six countries do not conduct any form of surveillance. The true incidence of JE is largely unknown due to the lack of sufficient surveillance and diagnostic laboratory testing in many countries (World Health Organization, 2014c; Campbell et al., 2011).

Additional information is required to compensate for limitations in surveillance data to

inform JE preparedness and response planning. In this study, we propose a method that compensates for under-reporting using demographic trends as a forecasting tool. We also estimate potential distribution of pig population density in Asia using livestock head count data from the United Nations' Food and Agricultural Organization (FAO). By estimating where people will be in the near future, and where amplifying host will be, we can identify areas where there may be a larger number of susceptible individuals and thus areas of higher risk for JE.

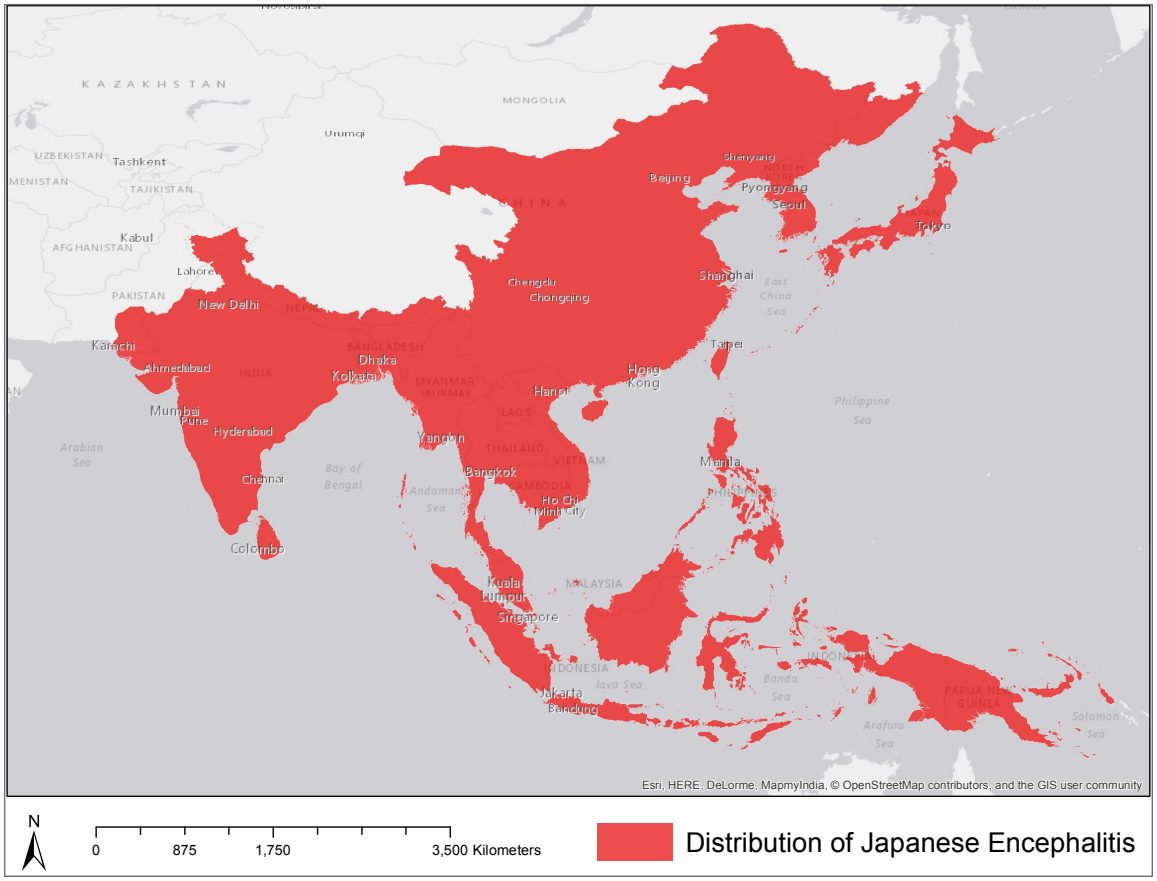


Figure 3.1: The current geographic distribution of Japanese Encephalitis.

3.2 Background

3.2.1 The Japanese Encephalitis Virus

The JE virus (JEV) is a *Flavivirus* transmitted via mosquitoes. The primary vector of JEV is the genus *Culex*, although other mosquitoes are also known to carry the virus (Solomon et al., 2003).

Rice agriculture areas provide ideal habitat for *Culex* mosquitoes, and thus risk of acquiring the virus tends to increase with proximity to paddy fields (Erlanger et al., 2009; Miller et al., 2012). Keiser et al. (2005) estimated that area dedicated to rice agriculture in Asia increased by 22% during the period 1963-2003, which has directly contributed to habitat expansion of *Culex* mosquitoes.

Wild birds and domesticated pigs are the primary hosts of the JEV. Wading birds from the family *Ardeidae* (herons, bitterns, and egrets) are a natural reservoir for the virus. These birds require shallow water bodies as feeding grounds (Sibley, 2000) and thus frequent areas dedicated to rice agriculture. Species from the family *Suidae*, which include domesticated pigs and wild boar, are the main amplifying hosts of JEV (World Health Organization, 2014c; Erlanger et al., 2009; Le Flohic et al., 2013). When pigs are infected with JEV, they are able to infect many additional mosquitoes. Increasing area dedicated to rice agriculture and the growing pork industry in Asia presents a growing concern for JE transmission in the near future. The distribution of risk factors related to pig husbandry, mosquito life cycle development and habitat, and interactions between hosts and vectors is extremely difficult to track over large areas. As such, this paper focuses on human and pig population trends as a way to forecast future changes of JE in Asia.

3.2.2 The State of Japanese Encephalitis in Asia

JE was first identified in 1871 and the first recognized epidemic occurred in Japan in 1924 (Misra and Kalita, 2010). Since then, the geographic distribution of JE has been expanding (Misra and Kalita, 2010). Misra and Kalita (2010) suggested that a changing landscape, involving deforestation and increased agricultural area (especially rice agriculture and pig farming), has promoted the spread of the disease. These landscape variables are associated with population growth and change.

3.2.3 Japanese Encephalitis in Urban Areas

Kumari et al. (2013) recently demonstrated how JE is becoming increasingly urban. For example, urbanized areas in the area of the National Capital Territory of Delhi, India provide

sufficient breeding habitats for *Culex* mosquitoes that transmit the JE virus (Kumari et al., 2013). The presence of migratory *Ardeidae* birds and pig domestication in these urban settings has amplified JE transmission. Kumari et al. (2013) was the first to identify cases of JE in the Delhi area and also provided evidence that JE is becoming increasingly an urban, rather than just a rural issue.

Esteva and Vargas (1999) stated that rapid human population growth leading to crowding and poor sanitation has resulted in the proliferation of *Aedes* mosquitoes, which have the potential to carry the dengue virus (although this disease can be endemic within a human population without any animal reservoirs). These types of urban landscape characteristics have led to increased risk of vector-borne diseases in urban areas. It is for these reasons that this study does not differentiating between rural and urban area in evaluating risk of JE. We do however, differentiate between rural and urban population growth rates in Asia, as urbanization is an increasingly important process transforming exposure opportunities in Asia.

3.2.4 Land-use & Landcover Change

The population in Asia is expected to increase from 4.342 billion in 2012 to 5.164 billion by 2050 (Department of Economic and Social Affiars: Population Division of the United Nations, 2014). Land-use change and the expansion of agriculture, the key driving forces of JE transmission are expected to intensify as population increases (Coker et al., 2011). Coker et al. (2011) stated that during the period 1990-2008, agricultural land increased in area by over 8% in Southeast Asia. Pig-farming is also intensifying in Asia. Erlanger et al. (2009) estimated that between 1990 and 2005, Cambodia, China, South Korea, Laos, Myanmar, Nepal, the Philippines, Sri Lanka, Thailand, and Vietnam all experienced increased pork production by between 12 and 381%. Such environmental factors, as well as population increases, will undoubtedly lead to an increased amount of JE cases unless otherwise managed.

3.2.5 Study Objectives

The objectives of this paper are to;

- Develop forecasts of populations vulnerable to JE across Asia for 2050.
- Develop forecasts under high, medium and low scenarios of amplifying hosts for JE across Asia for 2050.
- Identify locations where overlapping risk reservoirs, host populations, and facilitating contextual factors are present in 2050.

We aim to use to both present a methodology for anticipating future change in emerging diseases, and specific to JE, anticipate future patterns of risk for data-driven health policy formulation and disease management planning.

3.3 Materials and Method

3.3.1 Study Area

The study area includes all countries where JE has been identified. This includes most countries in South Asia, East Asia, and Southeast Asia as shown in Figure 3.1 (p. 21).

3.3.2 Data

Administrative boundaries were obtained from www.diva-gis.org (DIVA-GIS, 2013) and were used for district-level level analyses.

Two sets of population data were obtained. The first set was obtained from www.worldpop.org.uk and is presented as people per pixel for the year 2015 and is adjusted to match the UN Population Division estimates. The gridded population data set was use to estimate districts that were predominantly rural or urban, assuming that higher population densities occur in urbanized areas.

The second population set was obtained from GeoHive at http://www.geohive.com/ (GeoHive, 2013). GeoHive data were retrieved at one level below the national level for each country's most recent census and spatially joined to the administrative boundaries. Population values were matched to the same year (2010) by estimating mid-census values.

The United Nations World Urbanization Prospects project (UN WUP 2014) estimated rural and urban growth rates for five year intervals for each country (http://esa.un.org/unpd/wup/

CD-ROM/). These growth rates were used to project district-level populations until the year 2050, this method will be discussed further in Section 3.3.4.

Two pig population datasets were obtained from the Food and Agriculture Organization of the United Nations (FAO). Yearly pig head counts for the period 1961-2012 were obtained for each country from the the 'Livestock Primary' database (Food and Agriculture Organization of the United Nations, 2014). Data from this database is only available at the country level and was not available for every country in the study area (Figure 3.1). A gridded livestock density map for pigs for the year 2005 was obtained from http://www.fao.org/ag/againfo/resources/en/glw/GLW_dens.html (Animal Production and Health: Food and Agriculture Organization of the United Nations, 2007). The livestock density raster was used to determine district-level pig populations, which was not available using 2013 data from the 'Live Animals' database.

Landcover data derived from MODIS satellite imagery was obtained from the Global Land Cover Facility (GLCF) at http://glcf.umd.edu/data/lc/. The GLCF classifies MODIS imagery, which classifies landcover at spatial resolutions ranging from 250 m to 1000 m, into 16 landcover types at a 5 arc-minute (0.08333 degrees) spatial resolution (Channan et al., 2014). These landcover types were used to define districts as being either rural or urban so that appropriate growth rates could be applied.

3.3.3 District Classification

Weeks (2010) noted that "urban means nonagricultural". Regions that exhibit high population densities may still focus on a rural way-of-life (e.g. agricultural activities), especially in Asia where many countries have rural populations of over 50% of the total population (The World Bank, 2013a). Therefore, to classify districts as being predominantly urban or rural, landcover characteristics from Channan et al. (2014) were considered.

For a district to be classified as urban, two conditions were to be satisfied. The first condition required population densities to exceed 500 persons per square kilometre and that the majority pixel count was not 'croplands' as classified by Channan et al. (2014). The second condition stated that the majority pixel count was 'Urban and built-up' as classified by Channan et al. (2014) within district boundaries.

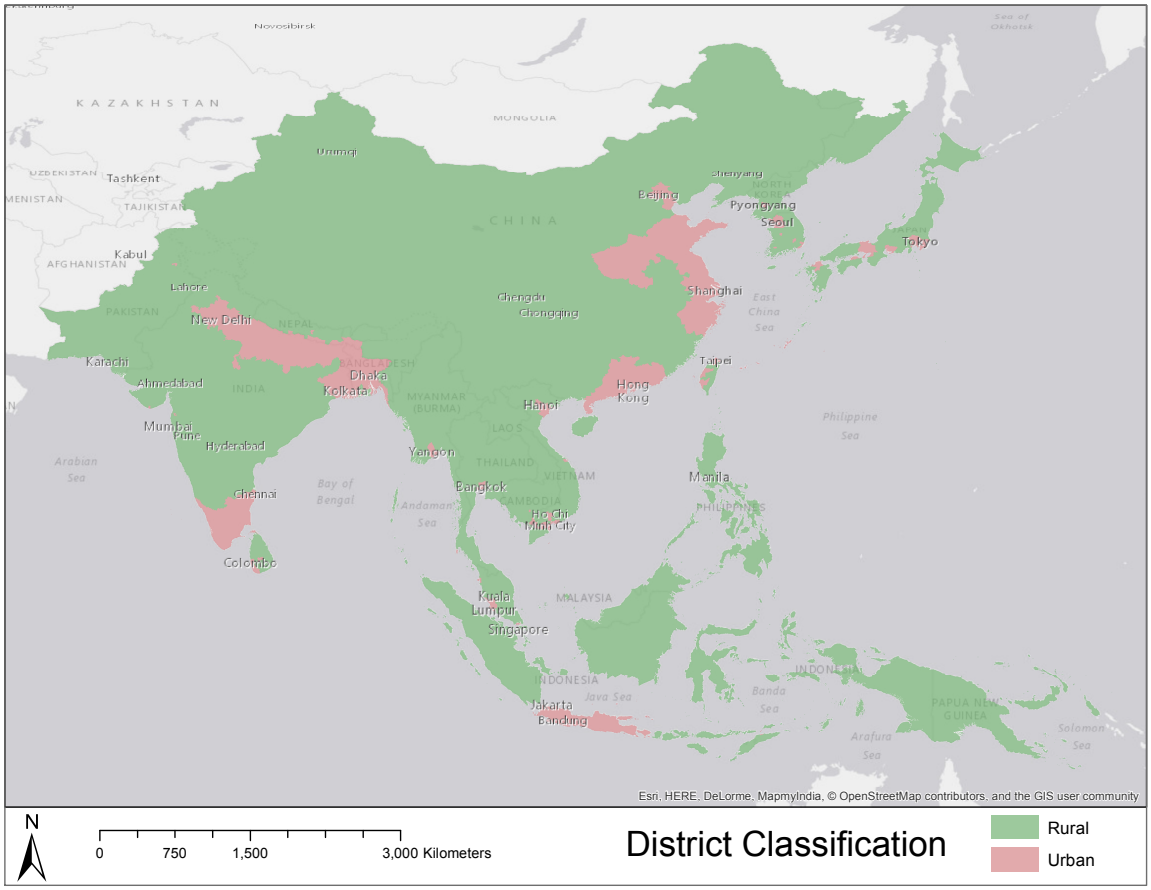


Figure 3.2: District classification as predominantly urban or predominantly rural.

For a district to be classified as being predominantly rural, it failed to meet the criteria of being defined as urban. Landcover characteristics were also considered, and if a district contained a majority of pixels classified as 'croplands' (Channan et al., 2014), then that district was classified as rural. Final district classification is shown in Figure 3.2.

3.3.4 Human Population Projections

The rate of urban population growth and rural population growth often differs. According to United Nations (2002), the urban growth rate for South-central Asia and South-eastern for the period 1950-2000 was 3.34% and 4.02%, respectively. The rural population growth rates for these two regions were 1.84% and 1.53%, respectively (United Nations, 2002). For studies in population projection at a subnational level, it is thus necessary to specify rural and urban areas. However, the distinction between urban and rural areas is not always straightforward. In many Asian countries, a high population density may not always suggest that a region is

dominated by an urban lifestyle. Many Asian countries are still characterized by a rural way of life, even though populations may be high (Tacoli, 1998). In this study, we use the most recent rural and urban growth rates as determined by the Department of Economic and Social Affiars: Population Division of the United Nations (2014) and apply them to districts classified in Section 3.3.3.

Five year average growth rates from WUP 2014 were used to project future population values. Population values were estimated for each 5 year interval's end year, which were then used to estimate the next interval's value. The geometric population growth model was used to estimate future populations as shown in Equation 3.1.

$$Pop_F = Pop_C(1 + \frac{r}{100})^T \tag{3.1}$$

where Pop_F is the estimated population for the designated year, Pop_C is the value of the population for the district's most recent census, r is the population growth rate from UN WUP 2014, and T is the number of years between Pop_F and the country's last census year.

Based on the classification of each district being predominantly urban or rural, each country's most recent census, and population growth rates from the UN's WUP 2014, population at the district level was calculated for the year 2050. We estimated that population in the study area increased from $\approx$ 3.9 billion for the year 2015 to $\approx$ 4.5 billion for the year 2050. Considering that the study area comprises most, but not all of Asia, these estimates are roughly in line with Department of Economic and Social Affiars: Population Division of the United Nations (2014) estimates for the study area.

3.3.5 Pig Population Projections

Pig population data in raster format for the year 2005 were used to estimate district-level values. A zonal statistics summary was performed on this raster dataset and used to estimate pig population proportions at the district-level (i.e. zonal sum in district divided by country total). Tabular data from the FAO's 'Live Animals' database for each year for the period 1961-2013 at the country level were then multiplied by district-level proportions devised from the raster

dataset.

Data for all countries in the study area were not available from the FAO's Live Animals database. Countries for which pig headcount data was obtained were; Bhutan, Brunei, Cambodia, China, Japan, Hong Kong, Macao SAR, Taiwan, North Korea, India, Indonesia, Laos, Malaysia, Myanmar, Papua New Guinea, South Korea, Sri Lanka, Thailand, Timor-Leste, and Vietnam.

At the country-level, pig livestock data from the FAO were used to estimate future values in low, medium, and high scenarios. Projections for the year 2050 were done using the `forecast` package in R, which is used for analyzing univariate time series data. Forecasts were calculated with a 75 percent prediction interval. The upper bounds, mean, and lower bounds of the prediction interval were used for high, medium, and low scenarios, respectively. The `forecast` function fits an exponential smoothing state (ETS) model to the time series data to predict future values (Yeasmin and Hyndman, 2008). Country-level pig projections were then allocated to district administrative boundaries according to their 2005 proportions.

3.3.6 Highly Vulnerable Regions

Highly vulnerable regions (HVR) for JE in 2050 were identified in two ways. Firstly we mapped the spatial overlap of population and pig populations. Population changes were categorized as either declining (% change < 0), stable (%change 0–5), low (%change 6–10), moderate (%change 7–20), and high (%change > 20). Pig populations were categorized using the same classification, and maps of growth classes were created and used to identify highly vulnerable regions.

Additionally, the percent change in the ratio of pig populations to human populations were calculated and mapped. Spatial analysis of the percent change was used to derive hotspots of significant change using the local Getis-Ord spatial statistic. The local Getis-Ord statistic ($G_i(d)$) (Getis and Ord, 1992) takes a moving average of a numeric variable x within a spatial neighbourhood defined by a spatial weight matrix $w_{ij}(d)$ (where d is the maximum distance at which nearby points are included as neighbours) as:

$$G_i(d) = \frac{\sum_j w_{ij}(d) x_j}{\sum_{j=1}^{n} x_j} \tag{3.2}$$

over the total sum of x. In practice $G_i(d)$ is defined as a standard normal deviate, and therefore can be interpreted as a *z-score*. We employ the $G_i(d)$ statistic to identify significant spatial clusters where the relationship between pigs and people are expected to change.

Results of HVR mapping were interpreted relative to contextual variables relevant for JE risk, including landcover classification, vaccination status of populations, and the current distribution of JE incidence.

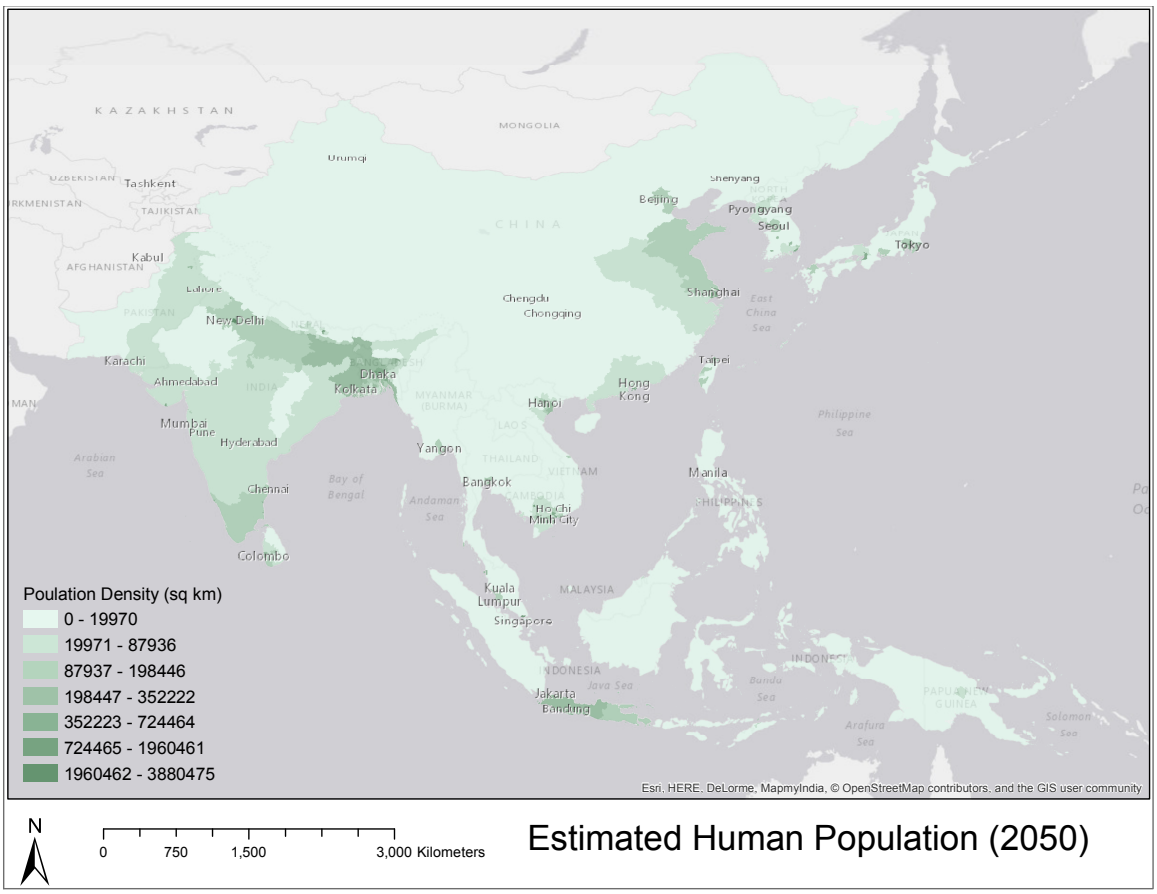


Figure 3.3: Estimated human population density for the year 2050.

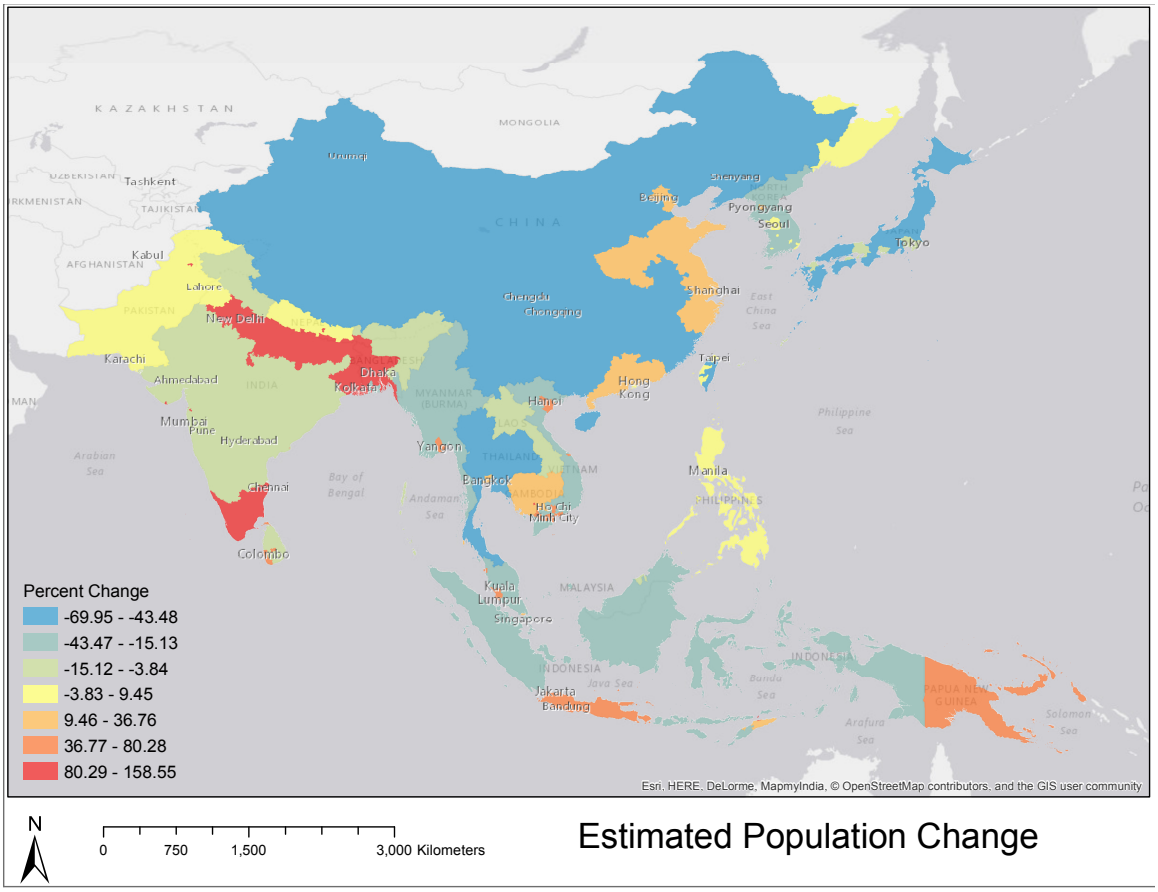


Figure 3.4: Estimated human population change at the district-level. Areas with higher rates of change are those classified as 'urban' in this study, and therefore experience higher rates of change.

3.4 Results

3.4.1 Human Population Change in Asia

Figure 3.3 shows our estimated population density, using urban/rural growth rate distinction, for the year 2050. Figure 3.4 shows our estimation of population change. While large regions of rural China are projected to experience population loss which dominate the map, urban areas in southeast Asia, and southern India are expected to experience positive change in population by 2050, reflecting the regional trend toward urbanization, described by the Department of Economic and Social Affiars: Population Division of the United Nations (2014).

3.4.2 Pig Population Change in Asia

Figure 3.5 (p. 35) displays our estimated pig projection and distribution at the district-level for the year 2050 under the three forecasted scenarios. Here, we see commonalities between the

three forecasts. Central China, and pockets of Nepal, northern India, Japan, and North Korea exhibit high densities in all three scenarios.

3.4.3 Highly Vulnerable Regions in Asia

Maps of pig populations and human populations in 2050 were combined to identify regions that exhibit high human population densities ($>$ 1,000 per sq. km) and high pig densities ($>$ 50 per sq. km) for the year 2050 in Figure 3.6 (p. 36). Parts of eastern coastal China and urban regions in southeast Asia are highlighted as particularly vulnerable in all three scenarios.

The relationship between pigs and people is expressed in Figure 3.7 (p. 37) as the percentage change in the ratio of pigs to people. The Getis-Ord statistic highlights hotspots (significant increases) in pig:human populations and coldspots (significant decreases) based on the local Getis-Ord analysis.

3.5 Discussion

This study considers potential population trends and JEV amplifying host projections at the district-level in order to estimate regions of potential risk for acquiring Japanese Encephalitis. By projecting future spatial patterns of susceptible human populations and risk factors (pigs), we highlight regions that should be of special interest for disease intervention strategies.

Figure 3.6 (p. 36) displays regions where pig population densities and human population densities are expected to be high based on our estimations. It should be noted that these maps do not account for vaccination programs (Figure 3.8, (p. 38)) and countries that are well-developed, such as Japan, may not be of particular risk.

Figure 3.7 (p. 37) highlights regions in which pig populations (under three different scenarios) and human populations are estimated to increase significantly by the year 2050. Figure 3.8 (p. 38) displays the status of immunization programs at the country level (World Health Organization, 2014c; Centers for Disease Control and Prevention, 2013). Regions depicted in green and pink either have no immunization program, or a program that operates at sentinel sites. Many regions that are highlighted as high risk in Figure 3.7 and Figure 3.6 overlap with

regions depicted in Figure 3.8 as having no immunization strategy. Because disease reporting is required for diseases at a national scale, it is important to examine sub-national trends over large areas to reflect the underlying dynamics of the disease processes. Here, we have focused on some key variables of interest, in order to map patterns of risk and risk change over time.

However, the transmission cycle of JE is incredibly dynamic; it is dependent on the presence of vector breeding habitat, agricultural practices, climate, proximity of the population to agricultural areas, as well as various social determinants. The number of JE cases within a country is also dependent on the presence of vaccination programs, the frequency and distribution of these programs within the country (i.e. located in rural vs. urban areas), and effective surveillance. Future large-area mapping efforts may take into consideration some of the main JE drivers at a more granular level of spatial detail, such as land-use change, paddy areas, population movements, and pig raising. Examples of this approach in Nepal Impoinvil et al. (2011); Robertson et al. (2013) have shed light on the regional distribution of risks, however obtaining detailed data for such variables over large areas at granular geographies remains a significant challenge.

Countries in Asia are also at different stages of development. Erlanger et al. (2009) suggested that incidence of JE will increase mainly in low-income areas. Countries such as Cambodia, Laos, and Myanmar will likely see JE outbreaks in the near future due to increases in irrigated rice agriculture and pig rearing (Erlanger et al., 2009). For example, during the period 1990-2005 Myanmar exhibited an increase of 47% in rice agriculture area as well as an increase of 381% in pork production (Erlanger et al., 2009). Centers for Disease Control and Prevention (2013) stated that Laos and Myanmar do not have JE immunization programs in place and that Cambodia only conducts surveillance in three of its twenty-three provinces. Considering these factors, we might expect that the incidence of JE may be greater in these countries than the current incidence as calculated by Campbell et al. (2011). In contrast, incidence may decrease in regions where vaccination programs are in their early stages. Recent progress in JE prevention has caused increased awareness of the disease, funding, and the availability of improved vaccines, although only in more developed countries (Centers for Disease Control and Prevention, 2013).

3.5.1 Limitations

The distinction between urban and rural districts is not always a dichotomy. Weeks (2010) highlights the difficulties with this distinction. In this study, we choose to assign districts to one of these classes based off of landscape and population characteristics in order to project future populations at the district-level. The aim of this was to obtain a fine-scale estimations of susceptible human populations.

This study is hindered by the quality of data used for analysis. We used the best available data to estimate potential trends and distributions of susceptible populations as well as the virus's host. Using livestock data for the period 1961–2013 to project future trends is not always straightforward and there is much uncertainty in these estimations. We have attempted to account for this uncertainty by forecasting pig populations under three scenarios using the upper and lower bounds of a prediction interval. Pig farming is also highly contingent on cultural factors that may not be represented in the data. For example, pig farming in Nepal has significantly increased after the year 1950 due to changing religious views and the tourism industry (Dhakal et al., 2014). Thus, the data shows an upward trend that may stabilize in the near future. However, these relationships are incredibly difficult to map at such a large scale and are beyond the scope of this study. Also, we show how pig populations are expected to decline in India. This may not actually be the case and may be due to low-quality of data obtained by the FAO. It should be noted that pig population estimations were conducted to the best of our knowledge using the best data available.

Land-use and landcover change plays a prominent role in the transmission of vector-borne diseases. This study did not focus on land-use and landcover change as it was beyond the scope of this study. There are many small-scale characteristics that contribute to mosquito populations, which are not visible from large-scale remotely-sensed data such as Channan et al. (2014). For example, Murty et al. (2010) found that certain species of *Culex* mosquitoes are abundant in urban areas of India and suggested that control measures should equally focus on urban and rural areas.

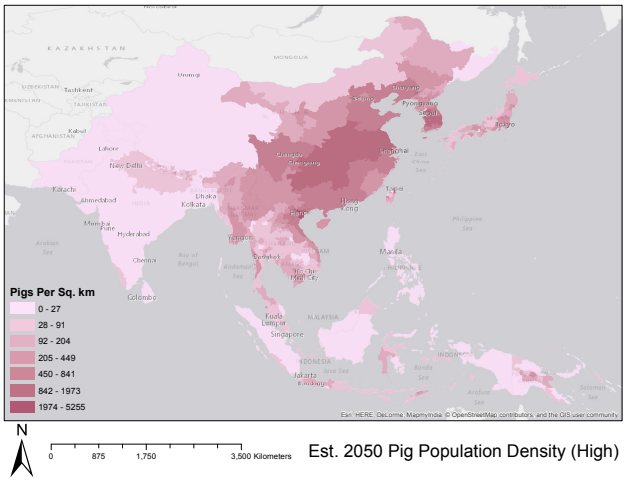
3.6 Conclusion

The (The World Health Organization, 2008; World Health Organization, 2014c) recommends that immunization is the most effective measure for preventing JE and should be extended to regions where JE is a public health problem. For many countries, resource constraints prevent the implementation immunization programs Centers for Disease Control and Prevention (2013). This study attempts to identify regions with high susceptible populations and high host populations. If regions with a high number of cases are identified, this may warrant increased attention for risk surveillance in this regions.

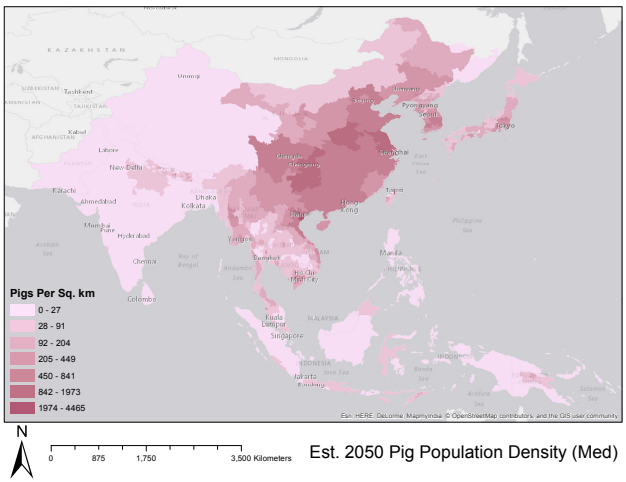
This study examines disease risk factors at a large scale; however, many changes are occurring at a much smaller scale. For example, pig farming in Nepal is increasing due to the reduction of cultural biases against pigs (Dhakal et al., 2014) increasing the number of amplifying hosts in this region. It is evident that much work is needed at a smaller scale to fully understand the spatial dynamics of JE on a larger scale.

Pigs are involved in the transmission of other zoonotic diseases, such as *Streptococcus suis*, which can cause meningitis and other serious symptoms (Gottschalk and Segura, 2000). Using methods presented in this study, it is also possible to highlight regions where interventions on other such diseases should be focused.

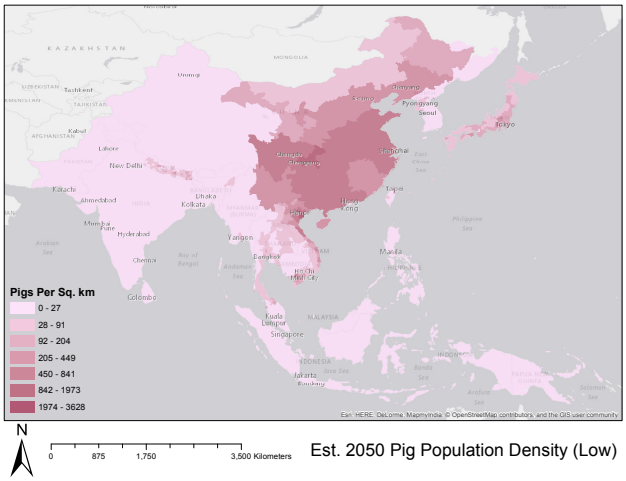
Although this study highlights regions where there could be high increases in both human and pig populations, we do not discourage national-level interventions to prevent JE in other regions in Asia. We attempt to highlight regions that may be of importance in order to appropriate resources to combat JE in the case that resources are limited.



Est. 2050 Pig Population Density (High)



Est. 2050 Pig Population Density (Med)



Est. 2050 Pig Population Density (Low)

Figure 3.5: Estimated pig population density with high, medium, and low forecast scenarios for the year 2050.

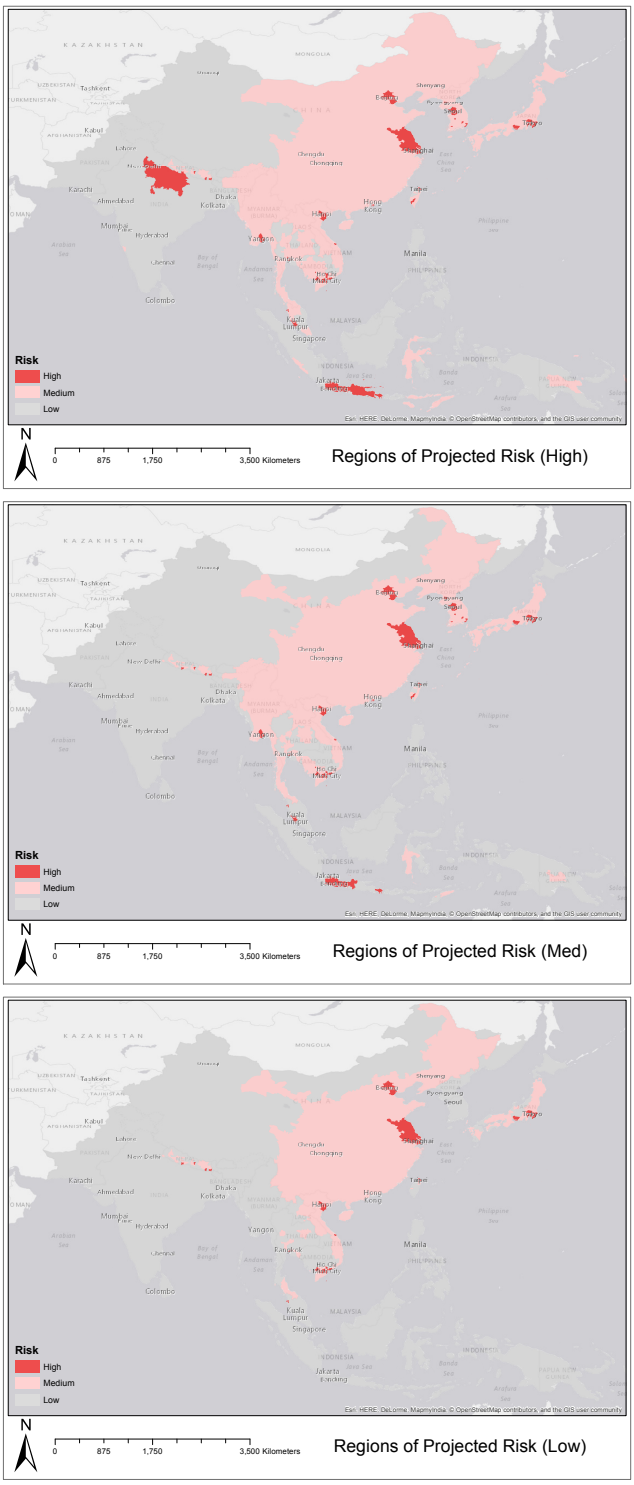


Figure 3.6: Regions with high pig densities (greater than 50/km^2) and high human population densities (greater than 1000/km^2) for the year 2050.

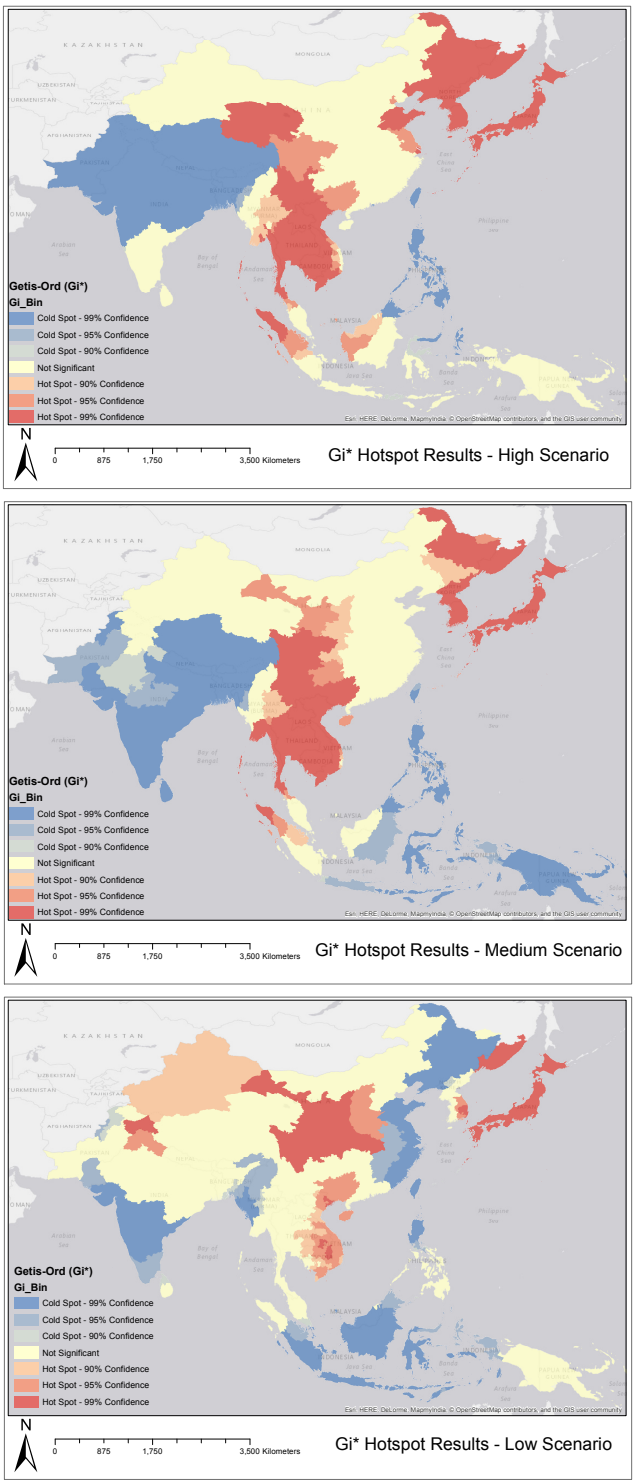


Figure 3.7: Gi^* results. Red regions represent areas with high values of pigs per 100,000 people under the high, medium, and low forecasted scenarios for the year 2050.

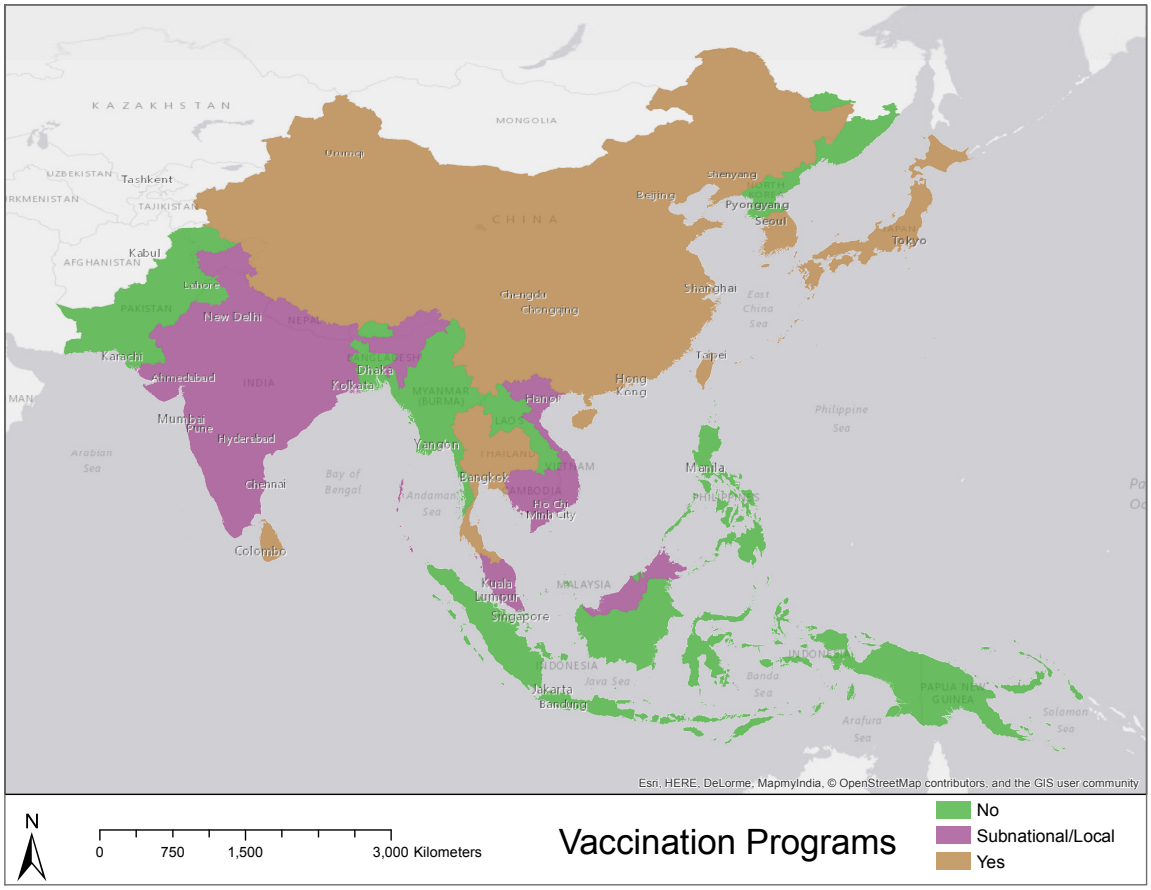


Figure 3.8: Country-level status of immunization programs. Regions depicted in green have no immunization strategies as of the year 2012 (World Health Organization, 2014c; Centers for Disease Control and Prevention, 2013).

Chapter 4

Modelling the Risk Landscape of Japanese Encephalitis in the Kathmandu Valley, Nepal

4.1 Introduction

Japanese Encephalitis (JE) is a mosquito-borne disease prevalent throughout Asia. Globally, the JE virus (JEV) causes approximately 67,900 cases of JE a year (Mackenzie et al., 2004; Misra and Kalita, 2010; Campbell et al., 2011; CDC, 2013). The case fatality rate of the disease ranges between 20–30% and approximately 30–40% of survivors suffer from permanent neurological sequelae, which results from to damage of the central nervous system (CDC, 2013; World Health Organization, 2014b). It is estimated that between 107,435 and 1,859,170 disability-life adjusted years (DALYs) result from JE globally (LaBeaud et al., 2011), and thus JE is a significant burden on the human population. JE is primarily acquired by children less than 15 years of age and is also more common in rural and agricultural areas (Mackenzie et al., 2004; Campbell et al., 2011; CDC, 2013; World Health Organization, 2014b). LaBeaud et al. (2011) stated that 57 per cent of the global population is currently as risk for contracting JEV.

As is the case for other vector-borne diseases, the JEV exists within a complex human-environment system. The primary vector of the JEV are mosquitoes of the genus *Culex*,

specifically *Culex tritaeniorhynchus* and *Culex vishnui* (World Health Organization, 2015b; Misra and Kalita, 2010). These mosquitoes lay their eggs as rafts, which require clean stagnent water. Irrigated rice paddy field provide this ideal habitat for *Culex* mosquitoes, and are thus considered to be high-risk areas for contracting JE. A growing population (Department of Economic and Social Affiars: Population Division of the United Nations, 2014) has lead to increasing area dedicated to these agricultural areas, which are considered to be ideal habitat for *Culex* mosquitoes (Keiser et al., 2005; Erlanger et al., 2009).

Populations of the virus's amplifying host, *Suidae* (pigs, both domesticated and wild), are also highly dependent on human population dynamics. Rising human populations in Asia have increased the demand for meat, leading to rising pig livestock counts (Erlanger et al., 2009; Food and Agriculture Organization of the United Nations, 2014). When species of the family *Suidae* are present, the number of infected mosquitoes can increase dramatically. Van den Hurk et al. (2009) summarized the primary reasons regarding the role of pigs in the transmission cycle of the JEV and are stated as follows: (1) there is a high natural infection rate of pigs, (2) pigs have a high viremia of the virus in the blood stream, (3) viremia in pigs remains high enough for transmission to mosquitoes for about 4 days, (4) pigs are highly prone to be fed upon by mosquitoes, and (5) pigs have a high birth rate, which provides new sources of virus-susceptible pigs each year. Pigs are also frequently located near rice paddy fields or in peri-urban areas where they are close to an at-risk human population (Dhakal et al., 2012, 2014).

The natural host of JEV are birds. Although over 90 species are known to be able to carry the JEV (Van den Hurk et al., 2009), the primary natural hosts are from the family *Ardeidae* (herons, bitterns, and egrets) (Misra and Kalita, 2010). *Ardeidae* birds frequent rice paddy fields as they serve as feeding grounds, and are thus able to spread the virus when they contact high density mosquito populations. Landscape factors that contribute to JE risk are highly dependent on human activities and thus JEV exists within a human-environment complex system. Land-use change, caused by population increases in Asia, will undoubtedly lead to increasing amounts of rice agriculture and pig husbandry (Erlanger et al., 2009) and possibly increase the rate of vector-borne disease trasmission (Keiser et al., 2005; Erlanger et al., 2009;

Lambin et al., 2010).

Surveillance of disease cases is the primary source of data for examining virus transmission within the human-environment complex system. However, previous research has demonstrated that surveillance data does not provide a complete representation of virus distribution (Kakkar et al., 2013; Shrivastava, 2014). There are many factors that contribute to the underreporting, and thus lack of data points within surveillance. This can include such instances as the perceived low severity of symtoms (e.g. asymptomatic or mild flu-like illness) and/or the lack of healthcare access. For example, Akiba et al. (2001) found that rural dwellers in Nepal preferred traditional healers, as opposed to hospitalization. If an ill individual does not seek medical attention at a sentinel surveillance site, then a potential JE case can go unrecorded.

Previous research demonstrates the need for methods that supplement surveillance records. Risk maps and the identification of high priority areas that implement transmission cycle and risk factor knowledge would be of use to fill in this gap. Such methods will provide more informed grounds to inform control and prevention measures. In this paper, we examine the JEV transmission cycle and its risk to the human population in the Kathmandu Valley, Nepal. This region possesses interesting characteristics that relate to JEV transmission. This region has undergone rapid change in recent years, involving remarkable land-use change (Thapa and Murayama, 2008, 2009, 2010). The rate of urbanization in the Kathmandu Valley is among the highest in the Asia-Pacific region, with rates up to 6.6% during the 1990s (Thapa and Murayama, 2009). This has led to increased area dedicated to rice agriculture in the peri-urban areas surrounding the valley (Thapa and Murayama, 2008, 2009, 2010), which has increased the amount of vector habitat and *Ardeidae* feeding area in the region.

Nepal has also exhibited a rapid increase in JEV's amplifying host. Pig husbandry and livestock numbers have increased from about 180,000 in 1960 to about 1.16 million in the year 2013 (Food and Agriculture Organization of the United Nations, 2014). The increase in JEV's amplifying host is due to factors such as increased tourism and associated demand for the diets of other cultures, changing cultural beliefs (there are many castes that do not eat pork), and programs such as the Pro-Poor Livestock Policy Initiative (Maltsoglou and Taniguchi, 2004) that promote small-scale livestock agriculture to reduce poverty levels.

One of the primary methods for mapping vector-borne disease risk is that of ecological niche modelling. Ecological niche modelling involves identifying the environmental covariates (e.g. temperature, precipitation, elevation, landcover, etc.) of a specific vector, such as *Culex* in the case of JE. Some examples of ecological niche models are described in (Peterson, 2001; Hart, 2010; Miller et al., 2012; Sallam et al., 2013) and specifically, (Peterson, 2015). However, this bottom-up approach does not often incorporate species interaction with the environment, or human-environment interactions. Thus, ecological niche models serve as a form of static risk map, where regions of high risk can be assumed.

The majority of research conducted on modelling infectious disease spread involves human-to-human transmission (e.g. SIR and/or SEIR models). However, these models only consider disease transmission within heterogeneous populations and do not explicitly incorporate environmental factors or any other forms of spatial data. Vector-borne disease transmission involves more components than simply a susceptible population, a pathogen, and an infected subset of that population. The environment plays a critical role in vector-diseases transmission, and thus adds considerable complexity to vector-borne disease modelling.

Previous research have implemented agent-, or individual-based modelling to examine vector-borne disease risk. Recent work by Linard et al. (2009) developed and agent-based model to examine the spatiotemporal dynamics of mosquito biting rates in southern France to assess malaria transmission. Malaria transmission was also examined using agent-based modelling techniques by Gu and Novak (2009), who programmed mosquitoes as their primary agents in efforts to predict their mobility patterns. Alexander et al. (2012) highlighted the importance of agent-based modelling in vector-borne and zoonotic diseases, as they are able to systematically incorporate heterogenous populations and various types as agents (e.g. virus hosts, vectors, and vulnerable human populations). Agent-based models to examine vector-borne disease transmission have also been employed in de Almeida et al. (2010); Li et al. (2012); Arifin et al. (2014); Dommar et al. (2014).

In this paper, we develop an individual-based model for analyzing JEV transmission in the Kathmandu Valley, Nepal. Environmental data, including remotely-sensed imagery, daily precipitation and temperature measurements, and pig populations were used to create the

risk-landscape of JEV in the Kathmandu Vally, Nepal. Nepal's national census was used to populate individuals in the model. We collate these data into a realistic model of the most significant JE-risk factors in Nepal. We then define a set of rules that govern agent behaviour and their interactions with the landscape in relation to vector-borne disease risk. Model results and output show that this model represents a realistic system of JEV transmission, as it roughly matches JEV surveillance data collect from the World Health Organization.

4.1.1 Japanese Encephalitis in Nepal

In Nepal, the first case of JE was reported in the district of Rupendehi during the year 1978 (Joshi and Gaidamovich, 1982). Initially, the disease was most prevalent in the Terai region – a belt of relatively low-lying land dominated by grasslands, savanna, and forest, that borders northern India (Akiba et al., 2001; Partridge et al., 2007; Impoinvil et al., 2011). Cases of JE in this region usually peak during and shortly after Nepal's monsoon season, which occurs from June to August (Partridge et al., 2007). This peak of JE cases is attributed to increased mosquito breeding habitat (e.g. standing water) and feeding bahaviour (Impoinvil et al., 2011). In the past, JE has only been a problem in the Terai region. In recent years, however, it has been identified at higher elevations in the hill regions of Nepal (Joshi et al., 1995; Bhattachan et al., 2009; Impoinvil et al., 2011). The Kathmandu Valley, which is the most populated region in Nepal, is located in the hill region of Nepal and it has recently been suggested that JE is endemic in this relatively small region (Partridge et al., 2007).

Joshi and Marg (2013) reported approximately 26,658 confirmed cases of JE during the period 1978–2003 in Nepal. Of those cases, 5,370 resulted in death (Joshi and Marg, 2013). The geographic distribution of JE in Nepal has shifted through time (Impoinvil et al., 2011). Since 1997, cases started to appear in the Kathmandu Valley (Akiba et al., 2001), which is the most populated region in Nepal, home to over 2.52 million people (GeoHive, 2013). In more recent years, Impoinvil et al. (2011) showed how JE cases were clustered in the Terai prior to the year 2006 and then shifted into the Kathmandu Valley region. The geographic shift is most likely due to the dynamics and spatial connection of various environmental factors, land-use and landcover changes, as well as shifts in the human population. For example, rapid urbanization,

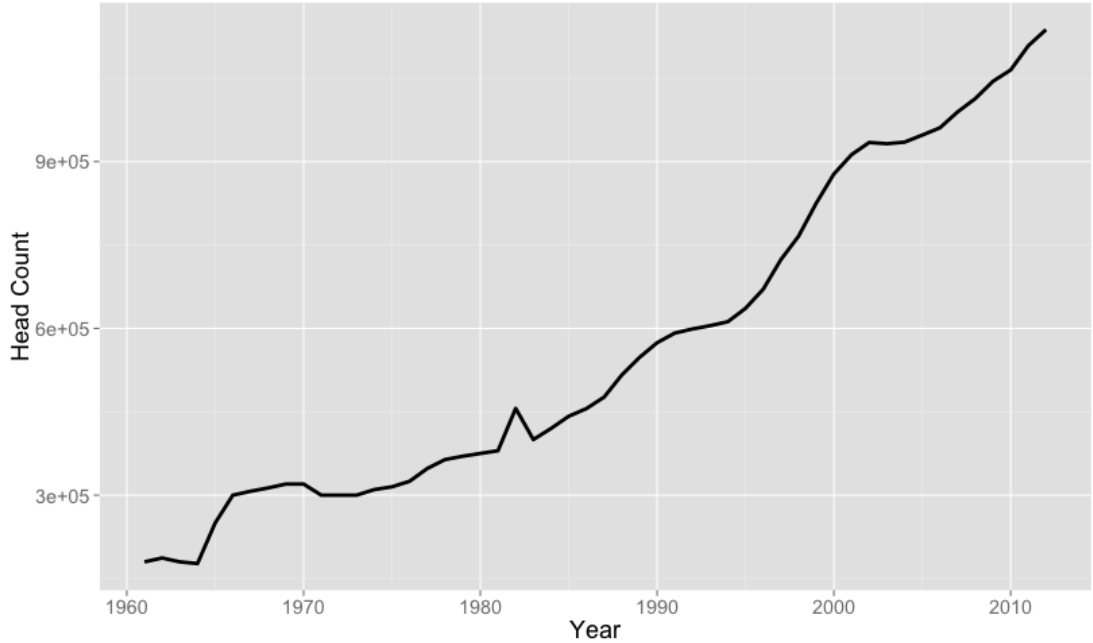


Figure 4.1: The steady increase in the number of pigs by year in Nepal. Data from the Food and Agriculture Organization of the United Nations (2014).

which is influenced by internal immigration processes, with people from higher elevations moving into the Valley. It's likely that people from higher elevations in the mountain region of Nepal have never been exposed to the JE virus, making them more susceptible as they do not possess JEV antibodies.

As mentioned in Section 4.1, pigs play an important role in the JEV transmission cycle. Rayamajhi et al. (2007) demonstrated that an individual's risk of acquiring JE increases if they reside near a pig farm. Rayamajhi et al. (2007) performed a hospital-based prospective cross-sectional study examining the proportion of children in Nepal, aged 1 month to 14 years, diagnosed with Acute Encephalitis Syndrome (AES) as having JE. They determined that the majority of those diagnosed under the umbrella-term 'AES', were actually afflicted with JE. They found that residing near pigs farms what the primary indicator factor related to those who acquired JE (Rayamajhi et al., 2007).

Rice agriculture and pig farming in Nepal has increased considerably in recent years. Figure 4.1 demonstrates the increase for total number of pigs in Nepal (Food and Agriculture Organization of the United Nations, 2014). Dhakal et al. (2014) stated that pig farming in Nepal has increased by 48% in the past 14 years due to reduced cultural biases against the

consumption of pigs as food. Dhakal et al. (2014) also stated that there is an increased demand for pig meat in urban areas, thus increasing the number of small pig farms in peri-urban areas. Area dedicated to agriculture is also increasing (Thapa and Murayama, 2008, 2009, 2010). Thapa and Murayama (2010) quantified land-use change in the Kathmandu Valley and found significant increases in land dedicated to urban areas and agriculture during the period 1967 to 2000.

Table 4.1: Summary of the top ranking risk factors, and their effect on the transmission of JEV, as determined by Hecker et al. (2013). Magnitude was derived by how many experts in the field agreed on the effects of a certain variable, with a maximum possible value of 32.

Variable (increase/decrease)	Effect (increase/decrease)	Magnitude
Vectors+	Pig JE cases+	32
Vectors+	Human JE cases+	31
Pigs+	Pig JE cases+	31
Pigs+	Human JE cases+	31
Poverty rate+	Human JE cases+	29
Pig JE cases+	Human JE cases+	28
Rice fields+	Mosquito net use+	27
Temperature−	Vectors−	27
Pig JE cases−	Human JE cases−	29
Vectors−	Human JE cases−	29
Vectors−	Human JE cases−	30

Hecker et al. (2013) explored expert knowledge on JE risk factors in Nepal using network analysis. They asked participants (those who were considered to be JE experts) to link various JE risk factors with potential effects (e.g. Would there be an increase/decrease/no change if the human population in the area increased by 20%?) (Hecker et al., 2013). Variables were ranked the highest if there was consensus among all experts.

The strongest relationships between JE risk factors and their effects from Hecker et al. (2013) are summarized in Table 4.1. Relationships are described by their magnitude, or the quantity of same responses among researchers, thus a magnitude of 32 indicates that all 32 participants in the study agreed upon that relationship. For example, all but one expert responded that

an increase in vectors would lead to an increase in human JE cases. Authors did not delve into questions as to why there was disagreement between risk factor and effect relationships. Results of Hecker et al. (2013) showed that the most important link between variables was the increase in vectors and their effect on the increase of infected pigs in the area. Many of the decisions and submodels in the individual-based model presented in this research are based on the research conducted by Hecker et al. (2013).

4.2 Methodology

In this section, we introduce the study area – the Kathmandu Valley, Nepal – and summarize its environmental and demographic characteristics. The methodology, including a detailed description of the model in terms of the Overview, Design Concepts, and Details Protocol (ODD) (Grimm et al., 2006) is given in Section 4.2.2, descriptions on how and why specific modelling methods are also introduced, including the employment of a variation of the Susceptible-Exposed-Infected-Removed model and the development of *Culex* mosquito habitat creation (Section 4.2.3).

4.2.1 Study Area

The Kathmandu Valley is located at approximately 27.64° north and 85.36° east (Figure 4.2). It is located in the hill region of Nepal, a belt of transition zone between the low-lying Terai region and the Himalayan mountain region. The Kathmandu Valley is comprised of three districts; Kathmandu, Lalitpur, and Bhaktapur, covering approximately 932 km ^{2}. Terrain in the Kathmandu Valley is highly varied, ranging from about 375 meters above sea level to about 2771 meters above sea level.

There are two primary seasons in Nepal, the dry season (October-June) and the wet season (June-September). The climate is considered subtropical, with mild temperatures ranging from 19–27°C in the summer and 2–20°C in the winter (Mountain GeoPortal, 2012). The monsoon season starts mid-June and reaches a monthly average total of about 363.4 mm for the month of July. It is during this season that Nepali rice farmers start irrigating and planting their rice

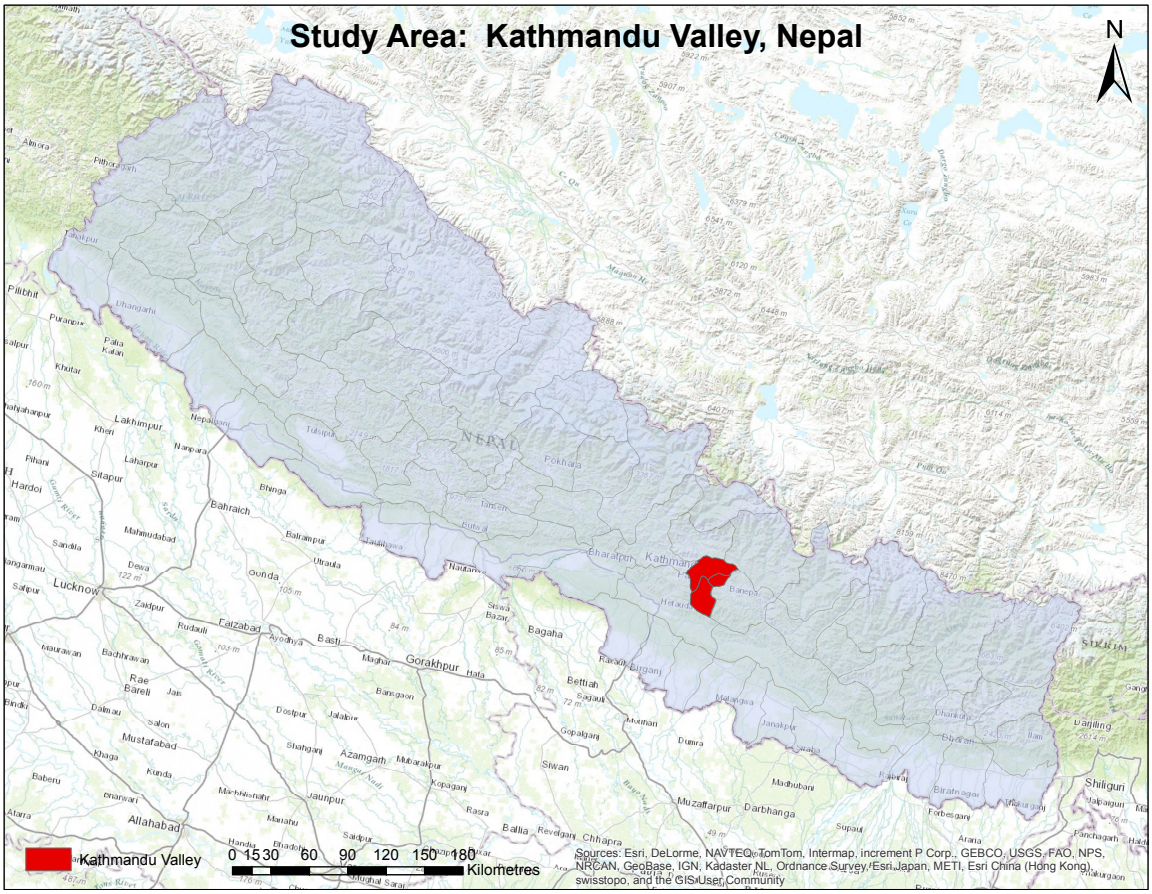


Figure 4.2: Location of the Kathmandu Valley, Nepal.

crops, unintentionally creating perfect breeding habitat for *Culex* mosquitoes and feeding grounds for JEV's natural reservoir, *Ardeidae* birds.

As of 2011, the population of the three districts that make up the Kathmandu Valley was about 2.46 million (Government of Nepal, 2014). Population density is greatest in the district of Kathmandu, and especially in the Kathmandu Metropolitan City, which had a population of about 975,453 in 2011 (Government of Nepal, 2014). The Kathmandu Metropolitan City is the most urbanized region of the valley, and although population growth has slowed since 2001, the urbanization rate has remained at about 3.9% per year (The World Bank, 2013b). Over the past few decades the population of the three districts has grown considerably. Figure 4.3 shows the population increases by district between the years 1981 and 2011. This increase has influenced land-use and landcover change (Thapa and Murayama, 2008, 2009, 2010), expanding agriculture and fragmentation of the landscape, which influences mosquito breeding habitat (Lambin et al., 2010). The growing population not only determines land-use and landcover

change, but increases the number of at-risk people in the valley. This is especially important if people have migrated from the mountainous region of Nepal, where the altitude is too great for *Culex* mosquitoes to survive.

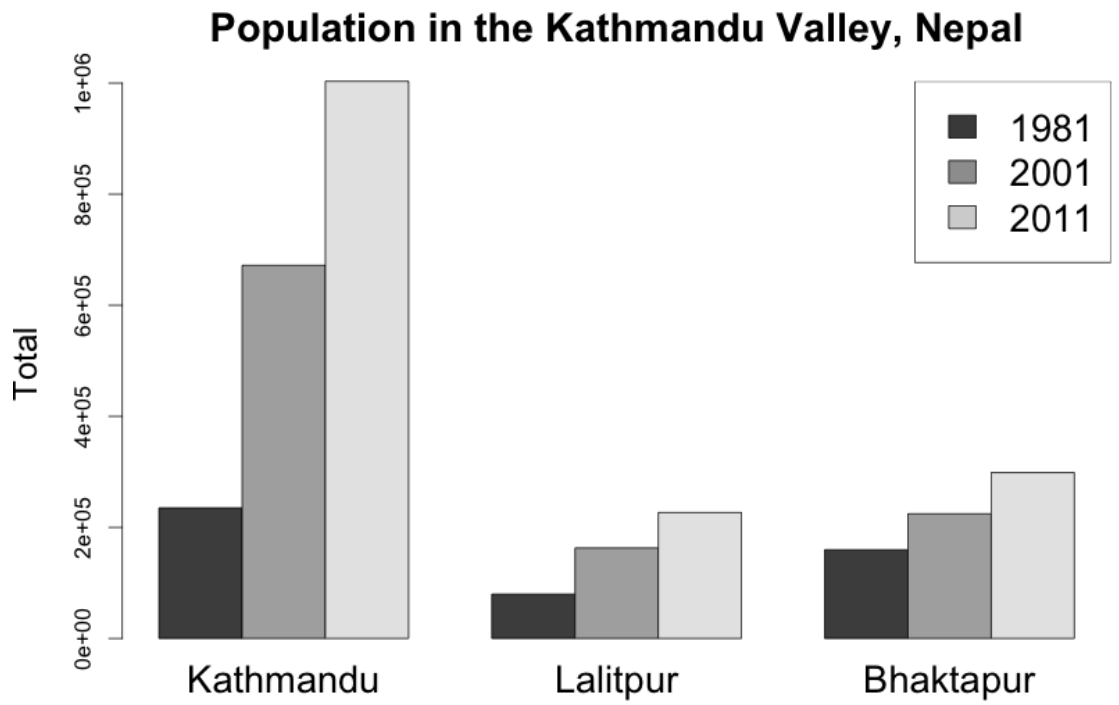


Figure 4.3: Population growth in the three districts that comprise the Kathmandu Valley, Nepal. Bars show populations for the years 1981, 2001, and 2011, respectively.

Pig populations have also increased in Nepal. According to the Food and Agriculture Organization of the United Nations (2014), livestock counts of pigs have increased from about 1.8 million in the year 1960 to over 1.16 billion in the year 2013. Programs such as the 'Pro-Poor Livestock Policy Initiative' have promoted small-scale pig farming to help alleviate poverty in Nepal (Maltsoglou and Taniguchi, 2004). Cultural biases against pig meat have been reduced in recent years, leading to pork being a more acceptable source of food. Figure 4.1 (p. 44) graphically displays the increase in pig population from the year 1961 to 2013.

The recent changes in Nepal, including the rapidly changing landscape due to urbanization driven by a growing population, frame the Kathmandu Valley as an interesting study area for examining JEV transmission. In this research, we create an individual-based model that examines JE landscape risk factors in the Kathmandu Valley and their influence on the human

population. It is hoped that that construction of a basic human-landscape risk model for contracting JEV will aid in the development of public health initiatives and spatial planning of livestock.

4.2.2 The Overview, Design Concepts, and Details Protocol

The current model is described in this section according to the recent textbook by Railsback and Grimm (2012) titled 'Agent-Based and Individual-Based Modeling: A Practical Introduction'. The ODD Protocol is a standard developed by Grimm et al. (2006) as a way to facilitate agent-based model replication. Descriptions of each standard are included in Appendix A (p. 105).

Purpose: The purpose of the current model is to demonstrate how landscape factors contribute to the risk of acquiring the JE virus. A realistic representation of landscape characteristics and population dynamics can then be used to analyze various scenarios, such as changes in climate and implementation of vaccination programs. Although many environmental datasets have been utilized, the model represents a simplified version of landscape complexity.

Entities, State Variables, & Scales: The primary entities, their state variables, and the spatial and temporal scales in this model are summarized in Table 4.2 (p. 50). If state variables were derived from data sources, those sources are labeled in the table. States variables of landscape entities are shown in Table 4.3 (p. 53).

Process Overview & Scheduling:

Overview: There are three primary processes in the model; 1) mobility, 2) agent-environment interaction, and 3) the evaluation of agents' SEIR state. These processes are shown in Figure 4.4 (p. 51). Firstly, human agents move around the landscape, following a few basic mobility rules (e.g. move from current location to random location within a certain radius, returning to home location at the end of the time-step). Because this research lacks mobility data, such as done in Hawelka et al. (2014) and Noulas et al. (2012), we have made assumptions regarding the general mobility behaviour of agents, rather than basing them on actual data. Mobility issues

Table 4.2: Model entities, their state variables, spatial and temporal scales used in the model. Data used to characterize state variables is also specified.

Agent Entities	**State Variables**
Humans	1. Age (2011 Census) 2. Sex (2011 Census) 3. Home VDC (2011 Census) 4. SEIR State (Susceptible, Exposed, Infected, or Recovered) 5. Time exposed - Number of days between contraction of virus and onset of symptoms - used to represent the incubation period of the JEV 6. Time infected - Number of days between onset of symptoms and transition to either state 'Recovered' or termination of that agent
Pigs	1. Age 2. Infected?
Landscape Entities	**State Variables**
Elevation	1. Metres above sea level (WorldClim)
Temperature	1. Weekly mean maximum temperature (Nepal's DHM) 2. Weekly mean minimum temperature (Nepal's DHM)
Landcover Type	1. Agriculture 2. Forest 3. Urban/Built-up
Precipitation	1. Weekly mean precipitation (Nepal's DHM)
Other	1. District 2. VDC boundaries 3. Road Network (ICIMOD) 4. Settlements (ICIMOD) 5. River zone (5 grid cell buffer from rivers, ICIMOD)
Scales	
Spatial	1. Kathmandu Valley ($\approx$ 932.41 km^2) 2. Patch size $\approx$ 65 metres
Temporal	1. Time steps represent daily processes

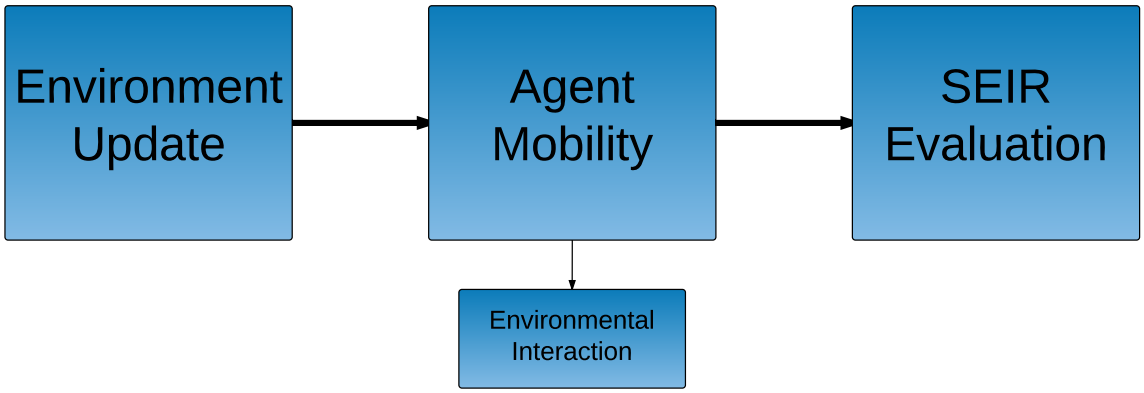


Figure 4.4: The primary process of the model. At each time-step, the environment grid is updated based on precipitation, temperature, and whether or not rice fields are planted. Agents then interaction with those landscape characteristics, which may alter their SEIR state. Each of these processes involves submodels that are defined in proceeding sections.

are further discussion in Section 4.11.1 (Limitations, p. 92) of this thesis.

The second main process is human agent interaction with the environment. An agent's transition from SEIR stated 'Susceptible' to 'Exposed' is determined by the types of landscape variables it encounters. Agents traverse the landscape and interact with landscape patches, which possess a *risk probability value*. At each agent-patch interaction, the agent generates a random float value, which is then compared with the patch's *risk probability value*. If the random float is less than the risk value, then that agent moves from the 'Susceptible' compartment of the SEIR model to the 'Exposed' compartment. The transition to other states is guided by incubation period estimation and the likelihood of displaying serious JE symptoms.

The transition between SEIR states ('Susceptible' to 'Exposed') represents an agent contracting the JE virus. Because the likelihood of full-blown JE (encephalitis symptoms) is approximately $1 : 250$ as stated by World Health Organization (2014c), another evaluation of an 'Exposed' agent's SEIR state is conducted between 5 and 15 days, which is the virus's incubation period. If the result of this evaluation returns TRUE, then that individual's SEIR state is changed to 'Infected'. If the result returns FALSE, then that individual is assumed to possess JEV antibodies and it's SEIR state is set to 'Removed', meaning that it cannot become infected with JEV at a later period. The date at which an agent moves into the 'Infected' state is then recorded and used for model validation and spatiotemporal analyses.

Environmental Processes: As stated in the *Overview* section of the Process Overview & Scheduling, agents interact with the environment at each time-step and potentially transition to new SEIR compartments depending on a landscape *risk probability value.* The landscape *risk probability value* can be thought of as the probability that an agent contracts JEV on a given day. This value is calculated by considering landscape attributes, such as landcover and a particular region's temperature and precipitation conditions for the current, and previous two weeks (three weeks in total). This is meant to capture adequate conditions for mosquito life cycle processes (e.g. *Culex* development from egg, to larva, to pupa, to adult takes approximately 7-13 days with specific climatic conditions (Rozendaal, 1997)). How this value is determined is explained in this section.

At the start of each iteration of the model, each patch's '*risk-probability-value*' is updated. The '*risk-probability-value*' of each patch is dependent on four factors: 1) distance to rice paddy field, 2) temperature, 3) precipitation, and 4) distance to pig farm. These environmental variables are represented as patch state variables, and shown in Table 4.3.

To compute the each patch's *distance-to-rice* value, a distance to rice paddy field raster was created using the Euclidean distance tool in ArcMap v.10.2. This tool uses Euclidean distance to compute a raster of distance values to the source cells (source cells = rice paddy fields). Because rice paddy fields are where *Culex* mosquitoes breed, population density of the vector is assumed to be highest in patches that are classified as 'Agriculture'. Risk values of patches that were not classified as agriculture, were computed using the inverse-square law, which states that as distance increases from a source the intensity (risk according to rice paddy field) is inversely proportional to the square of the distance and is denoted by:

$$\text{Risk}_{\text{Rice}} = \frac{1}{\text{distance}^2} \times 0.04. \tag{4.1}$$

To compute each patch's precipitation and temperature components of risk, the maximum possible value in each dataset was used to normalize it's relative probability. The maximum weekly precipitation value for the year 2007 was 347.98mm and the maximum weekly mean temperature was 39.70°C. Values from the current and the past two weeks were used to determine risk, as conditions need to be within a certain range for at least a week to account

Table 4.3: Table displaying all patch state variables, which were derived from variable environmental datasets.

State Variables	Definition
VDC	Represents the Village Development Committee (VDC) to which it belongs. Determines the number of agents initialized within VDC boundaries.
District	Represents the district to which a certain patch belongs.
Landcover-type	Determined by landcover classification of RapidEye imagery, there are three prominent landcover types; 'Agricultural' (high-risk areas due to higher densities of mosquito populations), 'Urban' (low-risk areas due to lack of breeding habitat for mosquitoes), and 'Forested Area' (not associated with significant increases/decreases in risk).
is-river?	Determined by the rivers shapefile obtained from ICIMOD (International Centre for Integrated Mountain Development, 2013). All patches are initially set to *is-river? FALSE*, then patches that intersect this shapefile are set to *is-river? TRUE*.
river-zone?	If a patch lies within 50 meters of a river (patch with *is-river? TRUE*), that patch is determined as being *river-zone? TRUE*. This variable is important in determining the initial locations of pig farms, which require the use of river water.
elevation-m-asl	A numeric value which represents the altitude (meters above sea level) of a particular location.
weekly-maxtemp	The average maximum temperature ($^{\circ}C$) for a given week at that location according to the interpolation of station maximum temperature data from Nepal's Department of Meteorology and Hydrology.
weekly-mintemp	The average minimum temperature ($^{\circ}C$) for a given week at that location according to the interpolation of station minimum temperature data from Nepal's Department of Meteorology and Hydrology.
weekly-precipitation	The amount of rain (mm) that a particular location received according to the interpolation of station precipitation data from Nepal's Department of Meteorology and Hydrology.
max-temp-list	A three item list which holds maximum temperature values for the current, and previous two weeks. With the first item holding the value for two weeks previous, the second item being the previous week, and the third item holding the value for the current week: $[T\max_i(t_1), T\max_i(t_2), T\max_i(t_3))]$.
min-temp-list	A three item list which holds minimum temperature values for the current, and previous two weeks. With the first item holding the value for two weeks previous, the second item being the previous week, and the third item holding the value for the current week: $[T\min_i(t_1), T\min_i(t_2), T\min_i(t_3))]$
rain-list	A three item list which holds the cumulative precipitation values for the current and previous two weeks. With the first item holding the value for two weeks previous, the second item being the previous week, and the third item holding the value for the current week: $[P_i(t_1), P_i(t_2), P_i(t_3)]$
rice-paddy-planted?	A logical value (TRUE/FALSE), that, when set to TRUE, increases the risk probability of a patch and increases risk. This value is set to TRUE during the monsoon months, and Nepali farmer have planted their rice fields, which promote mosquito habitat, breeding, and biting behaviour.
distance-to-rice	If a patch is *not* classified as 'Agriculture', then it possesses a *distance-to-rice* value that is greater than 0, which represents the distance of that patch to the closest 'Agriculture' patch.
risk-probability-value	A probability value which is computed by incorporating *weekly-precipitation*, *weekly-temperature*, *landcover-type*, *distance-to-rice*, *rice-paddy-planted?* variables.

for the minimum amount of time needed for *Culex* development. Weekly precipitation and temperature (minimum and maximum) values were stored as a list of length three for each patch ([week1, week2, week3]). As time progressed, the list was updated to include a new week's ptrecipitaion or temperature value at the end of the list, and the first value was removed.

Risk values associated with precipitation were calculated using the following rules and formulas;

If

$$P_i(t_1) \geq 100, \tag{4.2}$$

then;

$$\text{Risk}_{\text{Precipitation}} = \frac{P_i(t_2) + P_i(t_3)}{347.98 \times 2}. \tag{4.3}$$

If

$$P_i(t_1) < 100, \tag{4.4}$$

then;

$$\text{Risk}_{\text{Precipitation}} = \frac{P_i(t_2) + P_i(t_3)}{347.98 \times 2} \times 0.5. \tag{4.5}$$

where $P_i(t_1)$, $P_i(t_2)$, and $P_i(t_3)$ are the precipitation values for the first, second, and third items in the precipitation list corresponding to the weekly precipitation values, where $P_i(t_3)$ is the current week.

Risk values associated with temperature were calculated using the following rules and formulas;

If

$$T\text{max}_i(t_1) < 25, \tag{4.6}$$

then;

$$\text{Risk}_{\text{Temperature}} = \frac{T\text{max}_i(t_2) + T\text{max}_i(t_3)}{39.7(2)} \times 0.75, \tag{4.7}$$

if

$$T\text{max}_i(t_1) \geq 25, \tag{4.8}$$

then;

$$\text{Risk}_{\text{Temperature}} = \frac{T\text{max}_i(t_2) + T\text{max}_i(t_3)}{39.7(2)}. \tag{4.9}$$

If the average weekly temperature for the last three weeks is less than 10°C;

$$\frac{T\text{min}_i(t_1) + T\text{min}_i(t_2) + T\text{min}_i(t_3)}{3} < 10, \tag{4.10}$$

then $\text{Risk}_{\text{Temperature}}$ is updated using;

$$\text{Risk}_{\text{Temperature}} = \text{Risk}_{\text{Temperature}} \times 0.1, \tag{4.11}$$

and if the average weekly temperature for the last three weeks is less than, or equal to 20°C;

$$10 < \frac{T\text{min}_i(t_1) + T\text{min}_i(t_2) + T\text{min}_i(t_3)}{3} \leq 20, \tag{4.12}$$

then $\text{Risk}_{\text{Temperature}}$ is updated using;

$$\text{Risk}_{\text{Temperature}} = \text{Risk}_{\text{Temperature}} \times 0.75, \tag{4.13}$$

and if the average weekly temperature for the last three weeks is greater than 20°C;

$$\frac{T\text{min}_i(t_1) + T\text{min}_i(t_2) + T\text{min}_i(t_3)}{3} > 20, \tag{4.14}$$

then $\text{Risk}_{\text{Temperature}}$ is updated using;

$$\text{Risk}_{\text{Temperature}} = \text{Risk}_{\text{Temperature}} \times 1.2, \tag{4.15}$$

where P_i and T_i are the respective *weekly-precipitation* and *weekly-temperature* values of a patch.

The Equation 4.16 include model parameters; rain-weight (10^{-3}), temperature-weight (10^{-8}), and rice-weight (10^{-5}). These values were chosen to reduce each patch's *risk-probability-value* so that at each time step in the model run, the randomly generated number for each individual would likely be greater than *risk-probability-value*. Parameter values were chosen due to their affect on

To calculate each patch's overall risk value (Risk, as shown in Equation 4.16), the previous three risk values are then summed:

$$\text{Risk} = \text{Risk}_{\text{Rice}} \times 10^{-5} + \text{Risk}_{\text{Precip}} \times 10^{-3} + \text{Risk}_{\text{Temp}} \times 10^{-8}. \tag{4.16}$$

To incorporate the risk of contracting JEV in proximity to pig farms, the *risk-probability-value* of patches within 2400 meters of a pig farm were increased. This increase was dependent upon how many pigs were JEV seropositive at a given pig farm (Total_{pigs}), which was then added to the previously calculated risk value shown in Equation 4.16. Risk associated with the number of infected pigs on a pig farm was obtained by;

$$\text{Risk}_{\text{Pig}} = \text{Total}_{pigs} \times 0.001 + \text{Risk}, \tag{4.17}$$

where Total_{pigs} is the number of pigs at a given pig farm that are JEV seropositive. The number of JEV seropositive pigs is determined by user input (set 50% of all pigs to 'infected?' = TRUE), which is randomized at each model initialization.

Design Concepts: Grimm et al. (2006) discussed several design concepts that should be clarified when using the ODD Protocol. The design concepts – emergence, sensing, interactions, and observation – are discussed in the following enumeration.

1. *Emergence:* The phenomena of interest is the spatial and temporal distribution of agents with SEIR state set to 'Infected'. This trait will be validated to the disease incidence data collected from the World Health Organization's reporting system.
2. *Sensing:* Individuals in the model will 'sense' environmental conditions. Encounters

with environmental conditions that contribute to JEV transmission (i.e. rice paddy fields and/or pig farms) will lead to a greater probability of an individual contracting the JE virus.

3. *Interaction:* Individual's in the model interact with the environment. At each time-step, an agent will move to a different location in the environment. The environmental characteristics that determine landscape risk will determine the new location's risk of contracting the JEV as a *risk probability value.* Therefore, if environmental determinants of JEV risk are ideal, then there will be a greater probability of an agent becoming exposed to the JE virus.
4. *Observation:* The model will run daily time-steps and evaluate weekly cases of JE. Weekly totals will be stored in a comma separated value file that also contains agent/individual attributes (e.g. age, sex, home VDC). These data will then be imported into ArcMap v.10.2 for further spatial analysis and validation.

Initialization: The initial state variables of human agents are based on data from the 2011 Nepal Census (Government of Nepal, 2014). Table 4.4 (p. 58) shows the state variables of human agents and how they are initialized. These variables are not intended to change for each model run.

Pig agents possess two state variables, 'Age' and 'Infected?'. The 'Age' state variable is initialized as a random floating point number below 8, as the average age of domesticated pigs is approximately 8 years. The 'Infected?' state variable is set to either TRUE or FALSE based on user input. Data regarding pig population numbers by VDC was obtained via the National Zoonoses and Food Hygiene Research Centre (NZFHRC), who obtained VDC values from the Ministry of Agriculture Development (MOAD), Nepal. Pig location is initialized according to VDC counts from MOAD data. Pig farms are setup at random locations, but are in proximity (5 grid cells) to rivers, in attempt to represent the preference of pig farmers to locate to areas with water access.

Table 4.5 (p. 59) displays the model parameters. Human and pig populations were based on Government of Nepal (2014) and VDC level pig population data.

Table 4.4: Human agent initialization state variables. Population-related variables were initialized according to trends estimated from Government of Nepal (2014). State variables relating to the epidemiology of JE were estimated from literature and previous work regarding the JEV.

State Variables	**Definition**
Age	Determined by the number of agents within each age group according to the census. Due to knowledge of the number of people within each age interval (i.e. 0-4, 5-9, 10-14...), agents are randomly assigned a floating point age value between the end points of that interval. This state variable governs how long the agent will live within the environment.
Sex	Attribute which is determined by the 2001 Nepal Census (Government of Nepal, 2014).
Home VDC	The 'Home VDC' state variable is assigned in accordance with the 2011 census and is initialized in a similar manner as 'Age' and 'Sex' attributes of individuals.
SEIR	Initialization of this state is in accordance to the incidence rate as determined by Campbell et al. (2011). Agents are born with SEIR state set to 'Susceptible'.
time-since-exposed	This state variable represents the number of days since an individual's SEIR state was changed to 'Exposed'. The virus's incubation period is 5-15 days, so at a randomly generated number between 5 and 15, the infection calculation will be performed.
time-infected	If an individual's state is 'Infected', the state variable *time-infected* will increase by one each time-step.

Table 4.5: Initial parameters of the model.

Parameter	Value	Source
Agent Parameters:		
Human Population	21235	Author's estimation
Symptomatic to asymptomatic ratio	1:10	Author's estimation
Pig Population	201	Author's estimation
Pig Infection Rate	Kathmandu: 41.95% , Lalitpur: 91.4%, Bhaktapur: 75%	Pant (2006)
Landscape Parameters:		
Number Pig Farms	50	Author's estimation
rain-weight	1×10^{-3}	Author's estimation
temperature weight	1×10^{-8}	Author's estimation
rice weight	4×10^{-5}	Author's estimation
pig-weight	1×10^{-8}	Author's estimation
Rice field irrigation weight	1.0008	Author's estimation

Input Data: Numerous environmental datasets were obtained from a variety of sources in order to develop the model's landscape. The 'Landscape Entities' section of Table 4.2 (p. 50) gives a brief summary of these data and their sources. Data preparation is discussed further in Section 4.6. Input data includes all data that build the model's environmental conditions, such as; precipitation, temperature, elevation, and landcover type (Table 4.2). These environmental variables contribute to the imposed dynamics of the model run.

Static landscape data include rivers, roads, and settlement locations, which were obtained as ESRI shapefiles from the the International Centre for Integrated Mountain Development's Mountain GeoPortal at http://geoportal.icimod.org/ International Centre for Integrated Mountain Development (2013). These data did not require processing and thus will not be mentioned in the chapter entitled 'Data Preparation' (Section 4.6).

Submodels: Individual behaviour is dependent on two primary submodels. The first submodel is shown in Figure 4.5 (p. 60) and controls how individual's move. Individuals perform mobility behaviours dependent on their age. In our model, we assume that young children (those under the age of 5) stay at home. School-aged individuals make daily commutes to urban areas

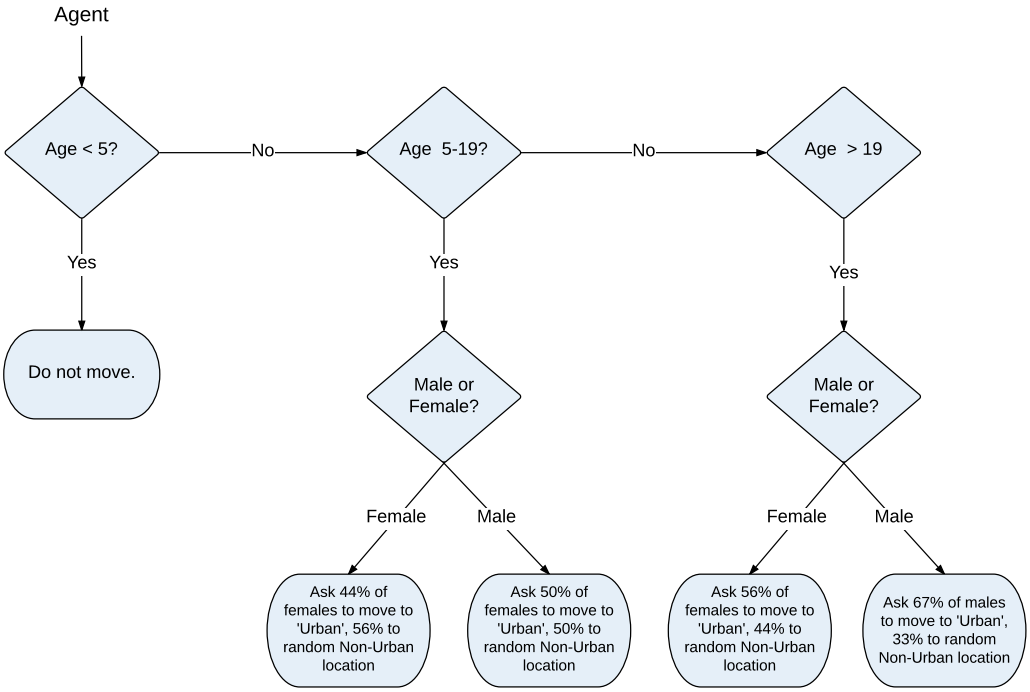


Figure 4.5: Mobility submodel that determines how agents traverse the landscape.

for school. The number of individuals in an age group that make these daily commutes is based on data from UNdata (2015) at http://data.un.org/CountryProfile.aspx?crName=Nepal#Social. We also use statistics regarding the percentage of males and female within the labour force from the same source. The model assumes that locations of employment are within urban areas and that the remaining population work in agricultural areas.

Figure 4.6 (p. 61) displays the primary process of the model. Agents traverse the landscape and encounter various types of landscape variables that influence the probability of an individual contracting the JEV.

4.2.3 Epidemiological Modelling

A variation of the Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model is used in this research to represent the dynamics of the JEV transmission cycle and it's effect on disease incidence. Other geographic disease evaluation that utilizes this epidemiological model can be seen in research conducted by Hailegiorgis and Crooks (2012). The SEIR model is

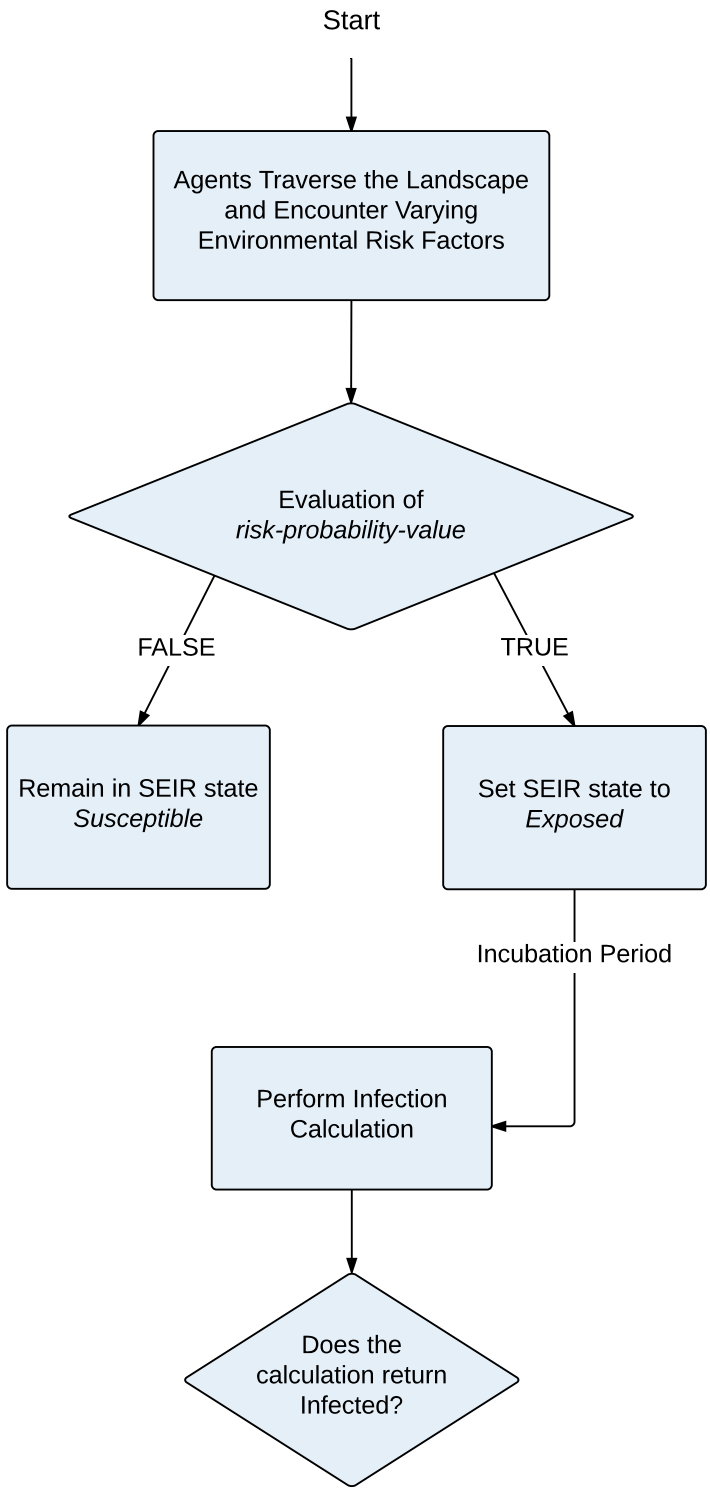


Figure 4.6: Submodel that determines an agent's SEIR state. Agents move throughout the landscape, then evaluate a randomly generated float against their current patch's *risk-probability-vale* (see Equation 4.16). If the float is less than the patch's probability value, then that agent sets it's SEIR state to 'Exposed'. It also allows for the virus's incubation period to pass, then performs infection calculation, which determines that agent will move to either SEIR state 'Infected' or 'Removed'.

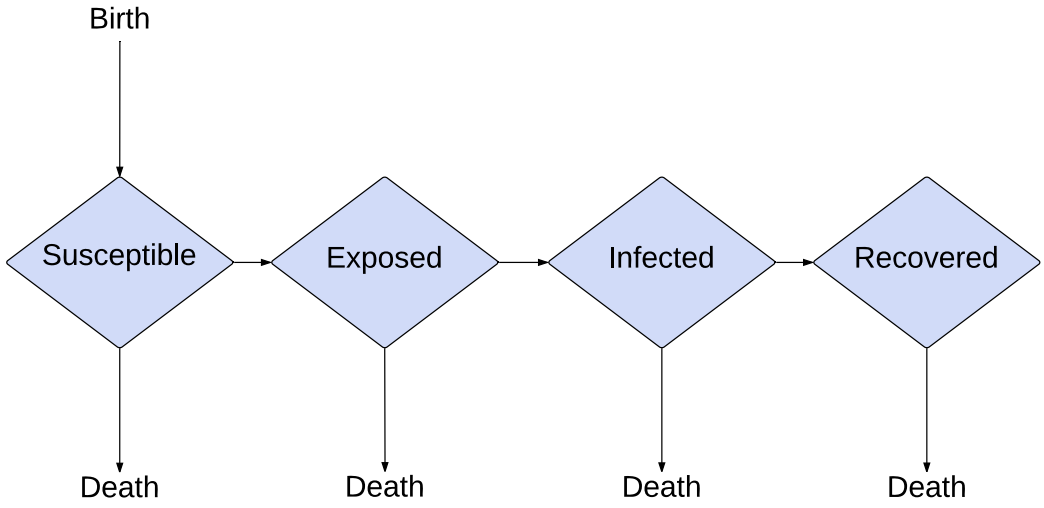


Figure 4.7: Epidemiological Susceptible-Exposed-Infected-Removed model flow. Agents move from the 'Susceptible' compartment to the 'Exposed' compartment if the evaluation

used instead of the simpler SIR model due to it's ability to capture the incubation period of the virus, which is capture in the 'Exposed' compartment of the model. Figure 4.7 shows the flow between the model compartments and Table 4.6 summarizes the variation between SEIR states.

4.3 Ecological Niche of *Culex* Mosquitoes

There are approximately 550 species of *Culex* mosquitoes worldwide (Rozendaal, 1997), of which about 34 have been seen for in Nepal (Darsie Jr et al., 1990). *Culex tritaeniorhynchus* is the primary vector of the Japanese Encephalitis virus, although there has been evidence that other species are able to transmit the virus (Solomon and Winter, 2004). Like all other arthropods, *Culex* mosquitoes require specific environmental conditions to proliferate. Temperature, precipitation, and landcover/use are primary characteristics of the environment on which *Culex* mosquitoes depend on certain ranges and/or types. Previous research, such as Miller et al. (2012), have modelled the distribution of *Culex* mosquitoes throughout Asia using ecological

Table 4.6: Explanation of how an individual moves through SEIR compartments in relation to the Japanese Encephalitis virus.

SEIR State	Determinants of SEIR State
Susceptible	All individuals are born with SEIR state set to Susceptible. Individuals each possess a probability state variable which is increased as they traverse the landscape and encounter JEV risk factors. Individual's remain in this state if the Infection Calculation, as shown in Figure 4.6 (p. 61), returns false.
Exposed	If the infection calculation returns TRUE, then an individual's SEIR state is changed to 'Exposed'. Individuals remain in the 'Exposed' stated for a random number of days between 5 and 15, which is the virus's incubation period. When this period is over, another infection calculation is preformed that represents the World Health Organization's estimate of 1:250 of those infected with JEV displaying full-blown JE symptoms World Health Organization (2014c).
Infected	If the second infection calculation returnsTRUE, then that individual's state is changed to 'Infected'. This state assumes that the individual is displaying full-blown JE symptoms and it's condition is recorded as a case in output results.
Removed	If the second infection calculation returns FALSE, then that individual's state is changed to 'Removed'. It is assumed that this individual has been exposed to the virus, did not display encephalitis symptoms, and now possess JEV antibodies that cause an individual to be immune to JEV.

niche modelling techniques. In this section, the ecological niche of *Culex* mosquitoes is described as well as it's potential geographic distribution in the Kathmandu Valley, Nepal.

Reeves et al. (1948) showed how various species of *Culex* mosquitoes can travel up 4 kilometers but generally occupy a range of about 325 meters from their habitat of origin. Masuoka et al. (2010) stated that the typical flight range for *Culex tritaeniorhynchus* is less than 1km and Verdonschot and Besse-Lototskaya (2014) showed how *Culex tritaeniorhynchus* have an average flight distance of about 68.4 meters and an average maximum flight distance of about 2214 metres, but generally stay within 1km of their breeding grounds.

Hecker et al. (2013) found that, among JE experts (researchers, medical professionals, and veterinary personelle) mosquito vector abundance was the most important factor influencing the seroprevalance of JEV in both the human population and pig populations. Other factors

that were deemed the most important in the JE transmission cycle are shown in Table 4.1 (p. 45).

4.4 NetLogo

NetLogo is a platform for developing ABM simulations that has increased in popularity in recent years. Developed by Wilensky (1999), NetLogo is an open source software that uses a Logo language dialect, which is easy to learn for new modellers. NetLogo can also incorporate GIS data, such as raster grid files (.asc and .grd) and ESRI vector files (.shp).

This research applies NetLogo 5.2 due to its accessibility and its ability to implement GIS data. The NetLogo online community is also extensive. Sites such as https://www.openabm.org/ provide open access to hundreds of agent-based models in NetLogo, which aid learning and sharing of code.

4.5 Model Interface

Figure 4.8 (p. 4.8) displays the model's interface. On the left are the controls for setting up, starting, and clearing the model. The date and number of agents are also shown. To the right of the main display (in NetLogo terminology, the "world"), are the landscape variables, which can be turned on or off, depending on user preferences. To the right of the landscape are the plots that show the number of agents in each SEIR compartment. Environmental variables, such as temperature and precipitation are also plotted.

4.6 Data

This chapter describes all data used in the current model, where it was obtained, and how it was prepared. There are two primary sections of this chapter; 1) environmental data and 2) population data. The environmental data section summarizes how the model's landscape was created and the population data describes how agents were created.

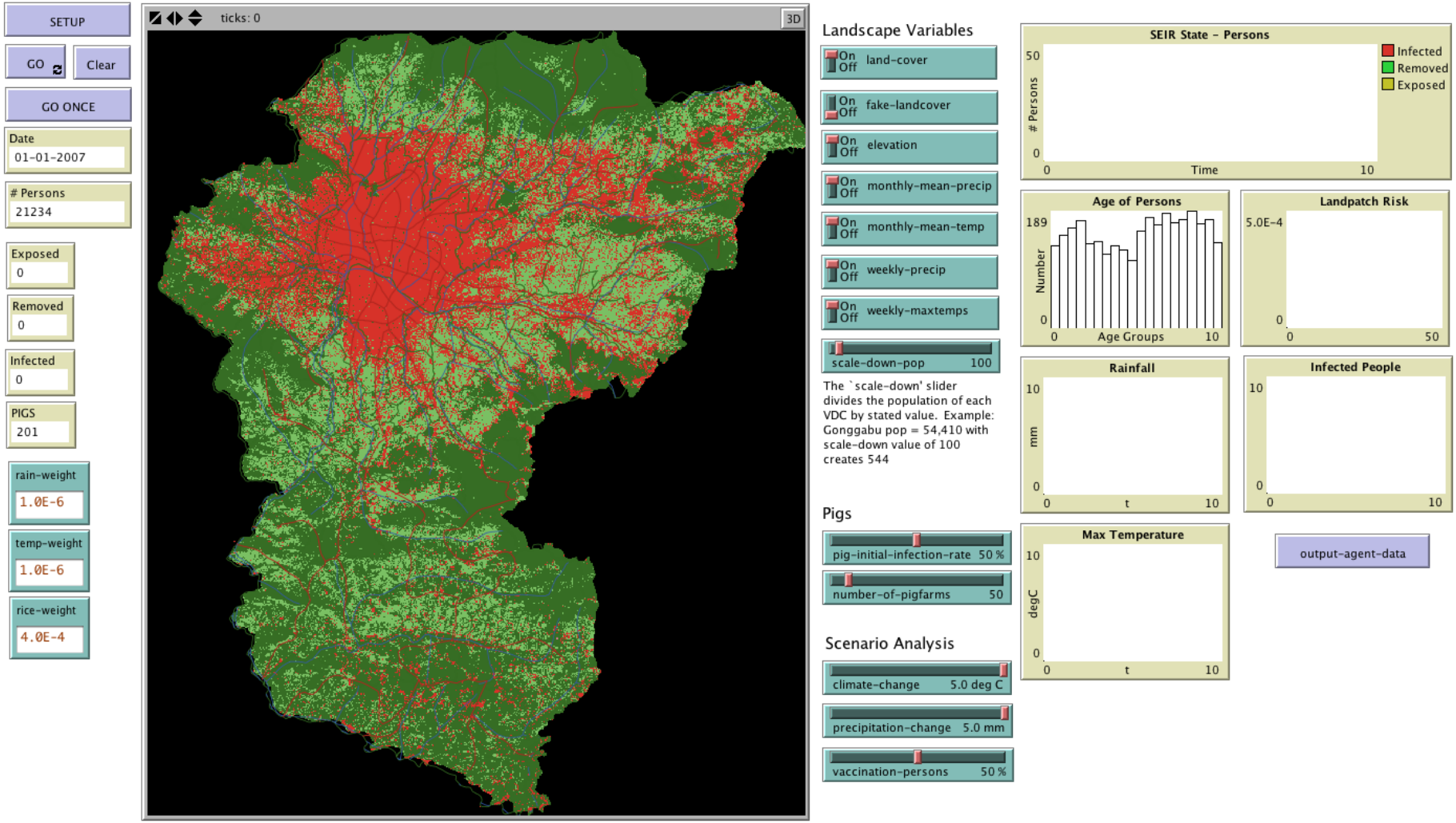


Figure 4.8: Model interface.

4.6.1 Environmental Data

RapidEye Imagery

RapidEye imagery was used to create a two-dimensional grid to represent the environment for the model. The RapidEye constellation consists of five satellites carrying a Jena-Optronik multi-spectral imager. This sensor records high-resolution, multi-spectral imagery in five bands (blue, green, red, the red-Edge, and near-infrared), which are displayed in Table 4.7. RapidEye imagery was chosen for this project due its ability to acquire spectral data in the red edge band (690-730nm).

Table 4.7: Table displaying spectral range of RapidEye bands.

	Band	Wavelength (nm)
1	Blue	440-510
2	Green	520-590
3	Red	630-690
4	Red Edge	690-730
5	Near Infrared	760-880

Previous research has shown the benefits of using the Red Edge band when identifying areas of rice agriculture. Recio et al. (2011) demonstrated higher accuracy when identifying areas of cropland and grassland when conducting a landcover classification with RapidEye imagery. Recio et al. (2011) attributed this high accuracy to the availability of data in the Red Edge region of the spectrum. Other work by Ichikawa et al. (2014) exemplified the use of RapidEye imagery to accurately identify rice paddy fields in rural areas of Japan and work by Conrad et al. (2012) showed that using metrics composed of the Red Edge, Red, and NIR bands were very effective to delineate between various crop types.

A total of eight RapidEye images were needed to cover the entire Kathmandu Valley. To use only cloud-free images, the eight images were from different dates. Dates ranged from years 2010 to 2013, all from the month of October. Information regarding these images is shown in Table 4.8 (p. 67).

Table 4.8: Tile numbers and dates for images of RapidEye images.

Image	Tile ID	Date (yyyy/mm/dd)
1	4551908	2010/10/30
2	4551708	2011/10/22
3	4551808	2012/10/19
4	4551907	2013/10/23
5	4551807	2013/10/23
6	4551707	2013/10/23
7	4551909	2013/10/24
8	4551809	2013/10/24

A mosaic of the images in Table 4.8 was created using PCI Geomatica's OrthoEngine Automatic Mosaicking tool. ArcMap v.10.2 was then used to perform a supervised image classification of the RapidEye mosaic. A total of 51 training sites were used to classify three different landcover types; 'Urban/Built-up', 'Agriculture', and 'Forest/Hills'. The resulting raster is shown in Figure 4.9 (p. 68). The maximum likelihood classifier was used to perform image classification.

To determine the accuracy of the supervised image classification, a confusion points matrix was created in ArcMap v.10.2. References points from the confusion matrix were then combined into a single raster to calculate the number of accurately assigned pixels. Accuracy assessment results of the classified image are shown in Table 4.9. A minimal number of landcover classes were used to generalize the environment. Only landcover types that were considered to be a primary influence on JEV transmission were included in this study.

Table 4.9: Accuracy assessment of the supervised image classification for the Kathmandu Valley.

Landcover Class	Total Reference Pixels	Accurate Reference Pixels	Accuracy
Urban/Built-up	50	45	90%
Forest/Hills	65	64	98%
Agriculture	45	41	91%

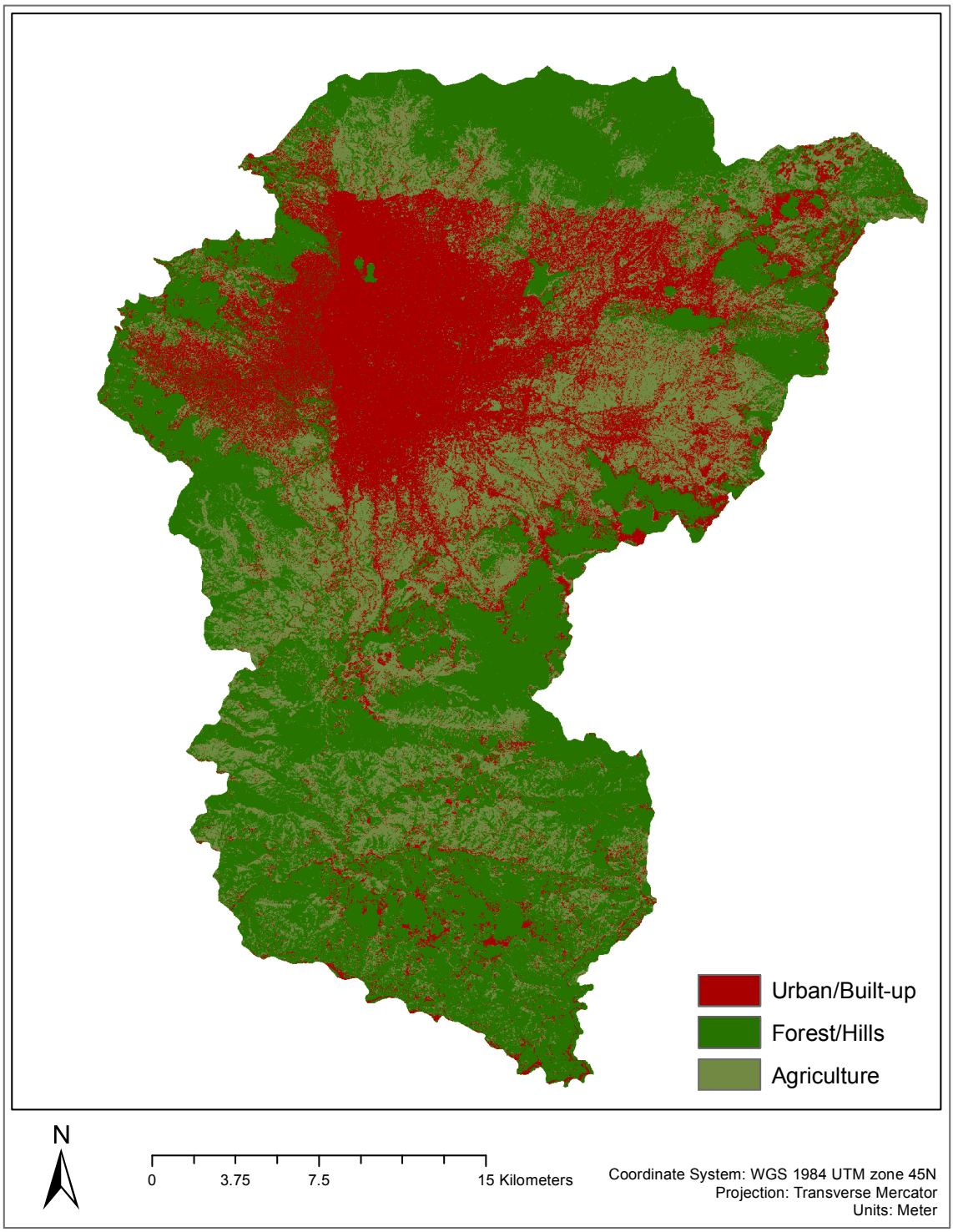


Figure 4.9: Results of RapidEye image classification into four landcover types (Agriculture, Urban/Built-up, Forested Land, and Water.)

Temperature & Precipitation

Climate variables from Nepal's Department of Hydrology and Meteorology, including daily rainfall and daily minimum and maximum temperature, were collected from 27 meteorological stations (Figure 4.10, p. 70) in and around the Kathmandu Valley. Daily measurements were used to interpolate surface values throughout the study area. These daily measurements were aggregated into weekly totals (rainfall) and averages (temperature) to lessen computation during the simulation.

A leave-one-out cross validation was performed for each interpolated surface. This method removes a data point (station), performs the surface interpolation, then compares the observed value to the predicted value. Appendix B (p. 109) shows the mean of the residuals and the variance of the Krigging cross validations.

Daily Rainfall Daily rainfall measurements from 27 stations located in or near the Kathmandu Valley were obtained from Nepal's Department of Hydrology and Meteorology. Data for the period 01 January 2007 to 30 September 2014 were used to interpolate raster grids for each week. Daily rainfall data were aggregated into weekly totals and then interpolated using cokriging techniques to account for the variety of elevation in the valley. Figure 4.11 (p. 71) shows the weekly precipitation dynamics from the meteorological station located at the Tribhuvan International Airport, in the district of Kathmandu (about 6 kilometres east of the city centre).

A cross-correlation analysis between weekly recorded AES cases and weekly precipitation values at Tribhuvan International Airport, Kathmandu and is shown in Figure 4.12 (p. 71). Figure 4.12 (p. 71) shows that the highest correlation value (0.288) occurs at a lag of 13 weeks. This relationship coincides with the time it takes to plant rice for agriculture, the virus's incubation period, and reporting of a JE case.

Elevation Data

Elevation data for the Kathmandu Valley was obtained from WorldClim.org at http://www.worldclim.org/current. The global altitude ESRI grid at a spatial resolution of 30 arc second

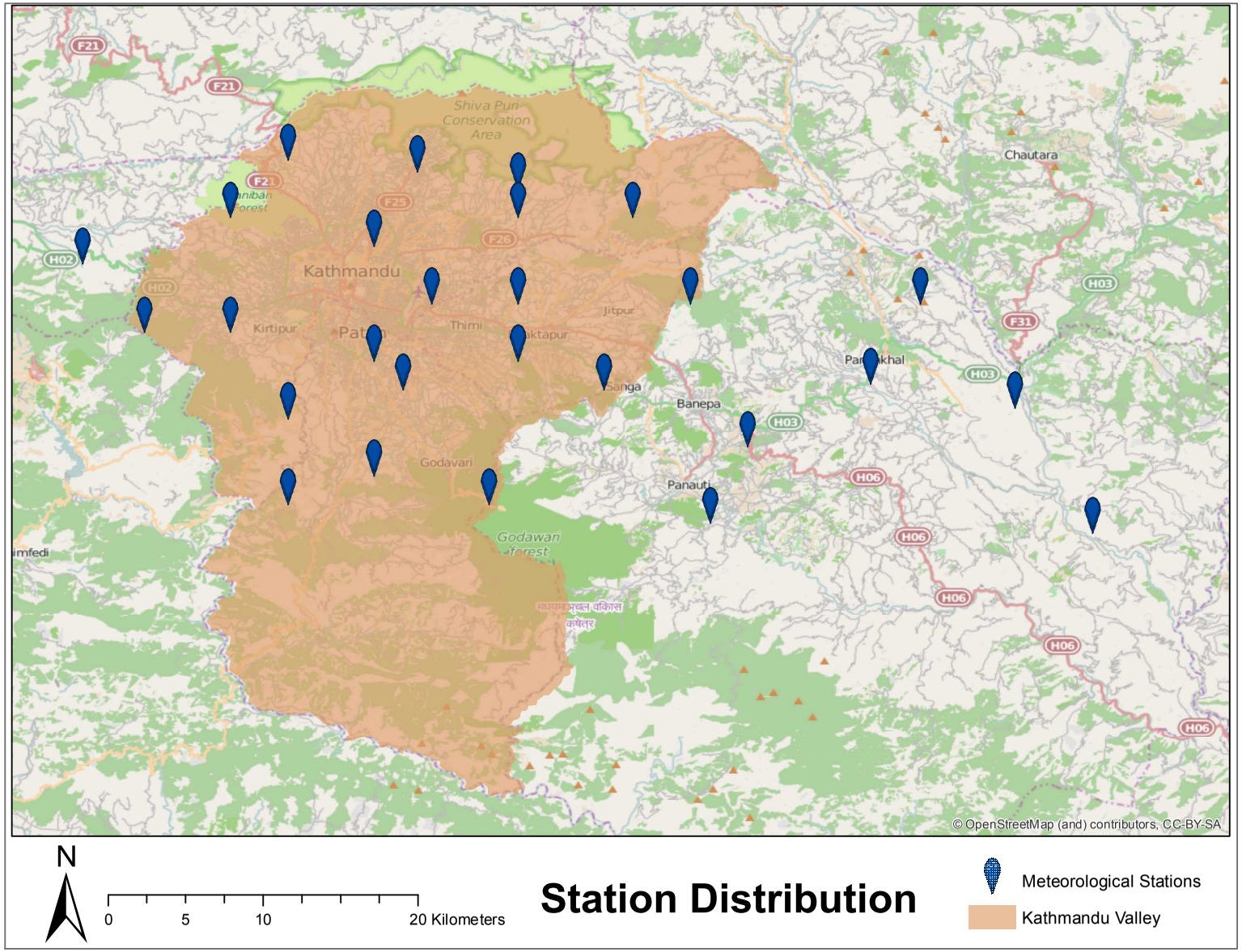


Figure 4.10: Locations of meteorological stations in and around the Kathmandu Valley. Data from Nepal's Department of Hydrology and Meteorology.

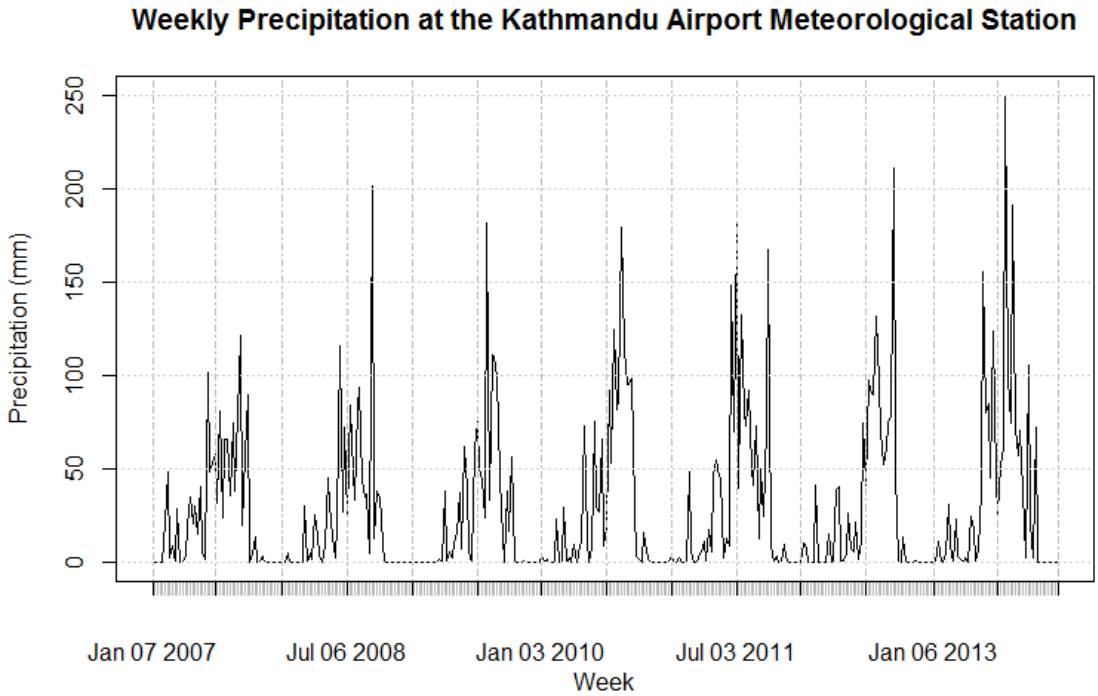


Figure 4.11: Weekly precipitation dynamics recorded from Tribhuvan International Airport, Kathmandu. Data from one station is used to demonstrate the seasonal dynamics of rainfall in Kathmandu, Nepal.

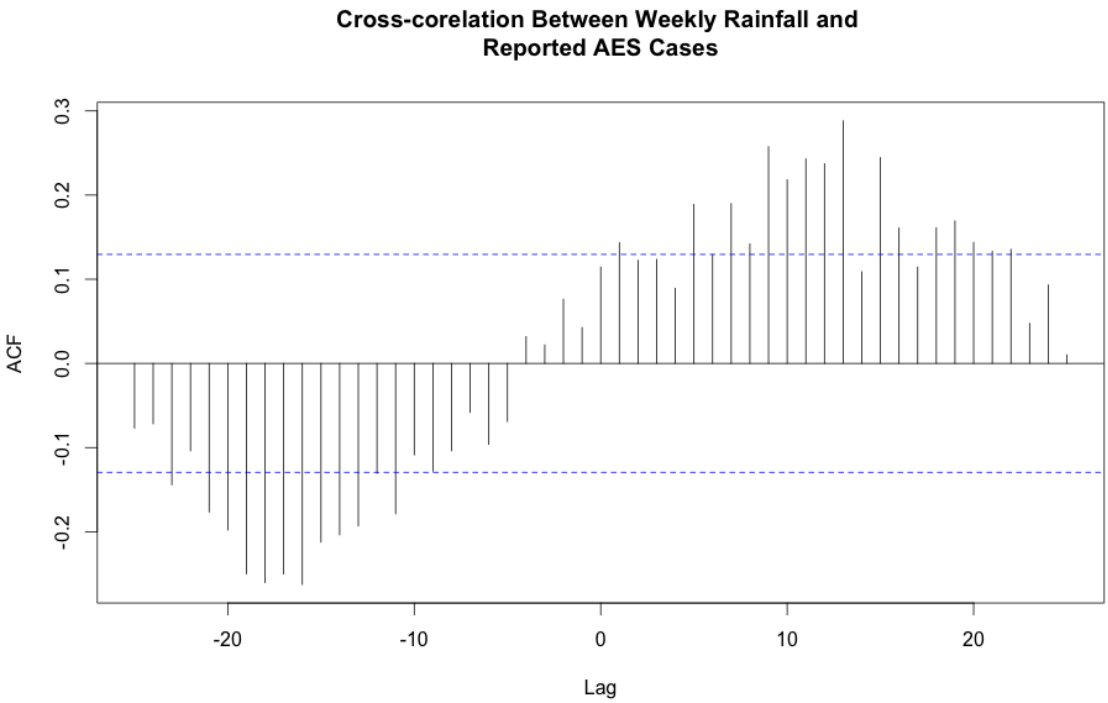


Figure 4.12: Cross-correlation analysis of weekly rainfall and weekly reported AES data. Results show that the strongest correlation value (0.288) appears at a temporal lag of 13 weeks between rainfall and reported AES cases.

(≈ 1km) was downloaded and clipped to the spatial extent of the study area. Due to NetLogo processing speeds, the elevation raster was then converted into a shapefile to be used in the model.

Other Landscape Data

Other landscape data, including rivers, road networks, and settlement locations where retrieved from the International Centre for Integrated Mountain Development (ICIMOD) (International Centre for Integrated Mountain Development, 2013). These datasets were downloaded as shapefiles and remained unmodified for implementation into the model. These data were used to define agent behaviour. For example, pig farmers are known to set up their farms along river banks in peri-urban areas. For this reason, it is estimated in this research that most (95%) of pig farms will locate within 50 metres of a river. The remaining 5% will initialize at random locations. This estimation attempts to account for the uncertainty of the exact locations of some pig farms, although common knowledge in Nepal states that pig farms are most likely to be located on, or very close to, a riverbank.

4.6.2 Population Data

Human Data

Human agents in the model were populated at the VDC level using Nepal's 2011 census (Government of Nepal, 2014). During model initialization, a proportional (actual population / 100) number of agents populate each VDC at a random location within that VDC. The actual population value (≈ 2.5 million) was not used due to computing power limitations.

The age and sex of each agent is also specified from the census data. A probability density function was created for each sex group. The 2011 Nepal census reports the number of people in each age group (0-4, 5-9, 10-14, 15-19, etc.)

Pig Data

Pig population data was obtained at the VDC level from Nepal's Ministry of Agricultural Development (MoAD). Data was collected from MoAD by the National Zoonoses and Food

Hygiene Research Centre in Kathmandu, Nepal and passed along to us. Because pig population was obtained at an administrative boundary level, rules are defined in the model to determine their initial locations (mostly near the banks of rivers).

4.7 Surveillance Data

In this section, we briefly describe the data used to validate the model. Acute Encephalitis Syndrome (AES) and JE data were obtained from the AES Surveillance System as part of the World Health Organization's Programme for Immunization Preventable Diseases (http://www.who.int/immunization/monitoring_surveillance/en/). In the following subsections, the temporal and spatial distribution of both these datasets are described.

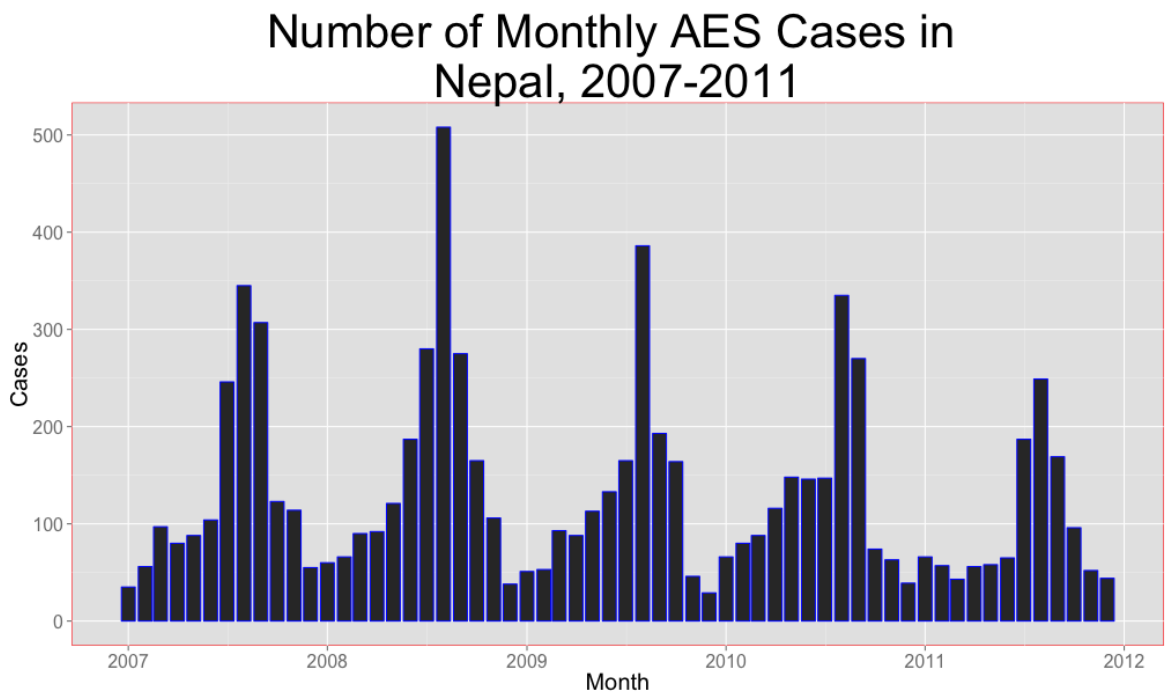


Figure 4.13: Country-level temporal dynamics of reported Acute Encephalitis Syndrome cases for Nepal during the period 2007-2011. Number of cases are shown as monthly totals.

4.7.1 Acute Encephalitis Syndrome Data

Acute Encephalitis Syndrome (AES) is an all-encompassing term used to describe the sudden onset of clinical symptoms of encephalitis. Symptoms may include; fever, headache, nausea,

changes in mental state, such as confusion and/or disorientation, seizures, and loss of consciousness. Symptoms are caused by inflammation of the brain, which may be caused by a wide range of pathogens, of which viruses are the primary cause. Viruses such as rabies, West Nile, measles, and the Human papillomavirus, along with others, may cause encephalitis in rare/extreme cases. Bacterial infections may also cause AES. However, recent research has shown that cases of AES in Nepal may actually be unconfirmed cases of JE (Robertson et al., 2013). Although the ratio of viral infections to full-blown encephalitis cases caused by such viruses is quite low (Whitley, 1990), it is a life-threatening condition that must be treated.

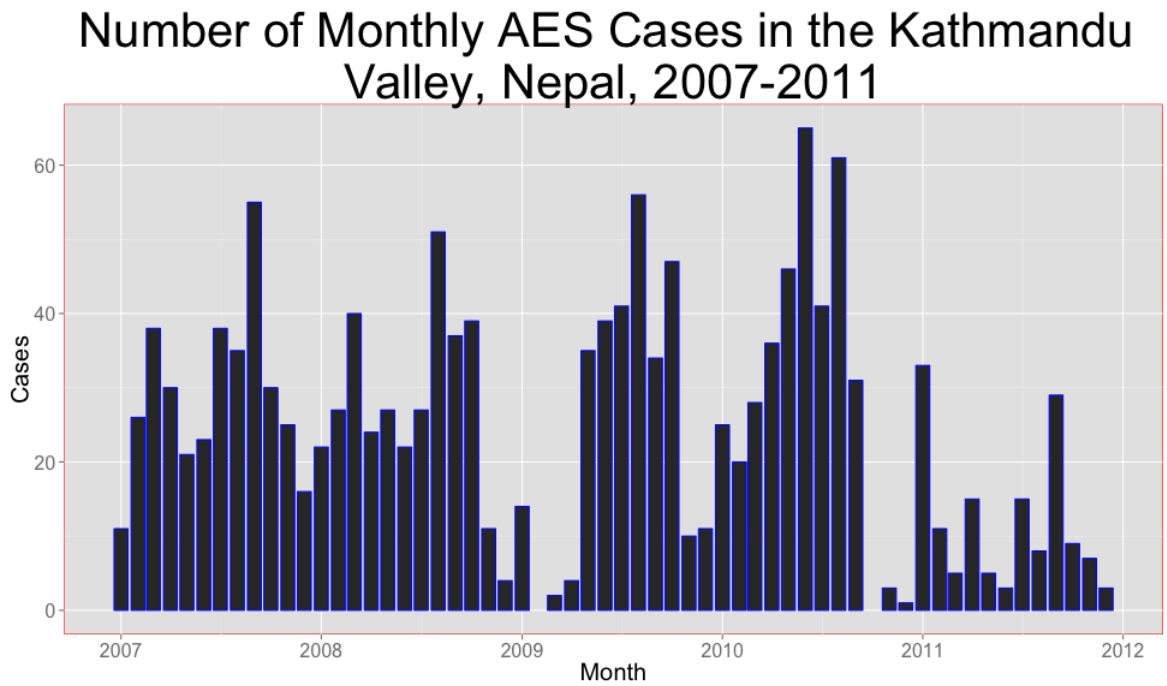


Figure 4.14: Temporal dynamics of reported Acute Encephalitis Syndrome cases for the Kathmandu Valley (districts of Kathmandu, Lalitpur, and Bhaktapur) for the period 2007-2011. Number of cases are shown as monthly totals.

Over the period 2007 to 2011, there were a total of 7993 reported cases of AES. The mean age was 15.22 ($SD = 18.69, range = 0 - 545, median = 9$), with more males (4703) than females (3288). Most cases were not vaccinated (6597), with 819 cases with the vaccination status of 'Unknown', and 441 stating that they were vaccinated. The Kathmandu district reported the highest number of cases at 1048 with the district of Morang reporting the second highest value at 538. There were a few missing values/errors in the data. For example, three Case ID's occurred twice in the data with differing entries. There was one missing 'Sex' classification,

136 missing vaccination status entries, 127 'Date Onset' entries, and a maximum age of 545 years of age.

Figure 4.13 (p. 73) shows the temporal dynamics of reported AES cases at the country-level for Nepal. The graph shows the number of monthly reported cases of AES. There is a peak in the number of cases around month 8 (August), which roughly corresponds to the time it takes to plant rice, the life-cycle of *Culex* mosquitoes, and the virus's incubation period. These events and their time periods are shown in Table 4.10 (p. 75).

Figure 4.14 (p. 74) shows the temporal dynamics of reported AES cases in the three districts that comprise the Kathmandu Valley; Kathmandu, Lalitpur, and Bhaktapur. Seasonal dynamics are less apparent due to the reduced number of reported cases at the smaller spatial scale.

Table 4.10: Rice agricultural events and their corresponding time periods.

Agricultural Event	Time Period
Rainy Season/Monsoon	June to August
Rice Production	Mid-June to the first week of September
AES Peak	August

Figure 4.15 (p. 76) shows the spatial distribution of reported AES cases in Nepal during the period 2007-2011. AES data was available only at the district-level, making spatial analysis less accurate than when analyzing JE data. It is apparent that the district of Kathmandu contains the most cases of AES (1048 reported cases for this time period). This is partially due to it's high population density and availability of healthcare facilities in the most urban region of Nepal. It is in the more populated regions where individuals are more likely to seek healthcare due to greater accessibility.

4.7.2 Japanese Encephalitis Surveillance Data

Japanese Encephalitis case data was also obtained from the WHO's AES Surveillance System. Suspected cases of JE must undergo laboratory testing for confirmation. Laboratory testing for JE involves a JE virus-specific Immunoglobulin M (IgM) capture enzyme-linked immunosorbent assay (ELISA) test on the cerebrospinal fluid (CSF) or serum of a patient. The ELISA test uses

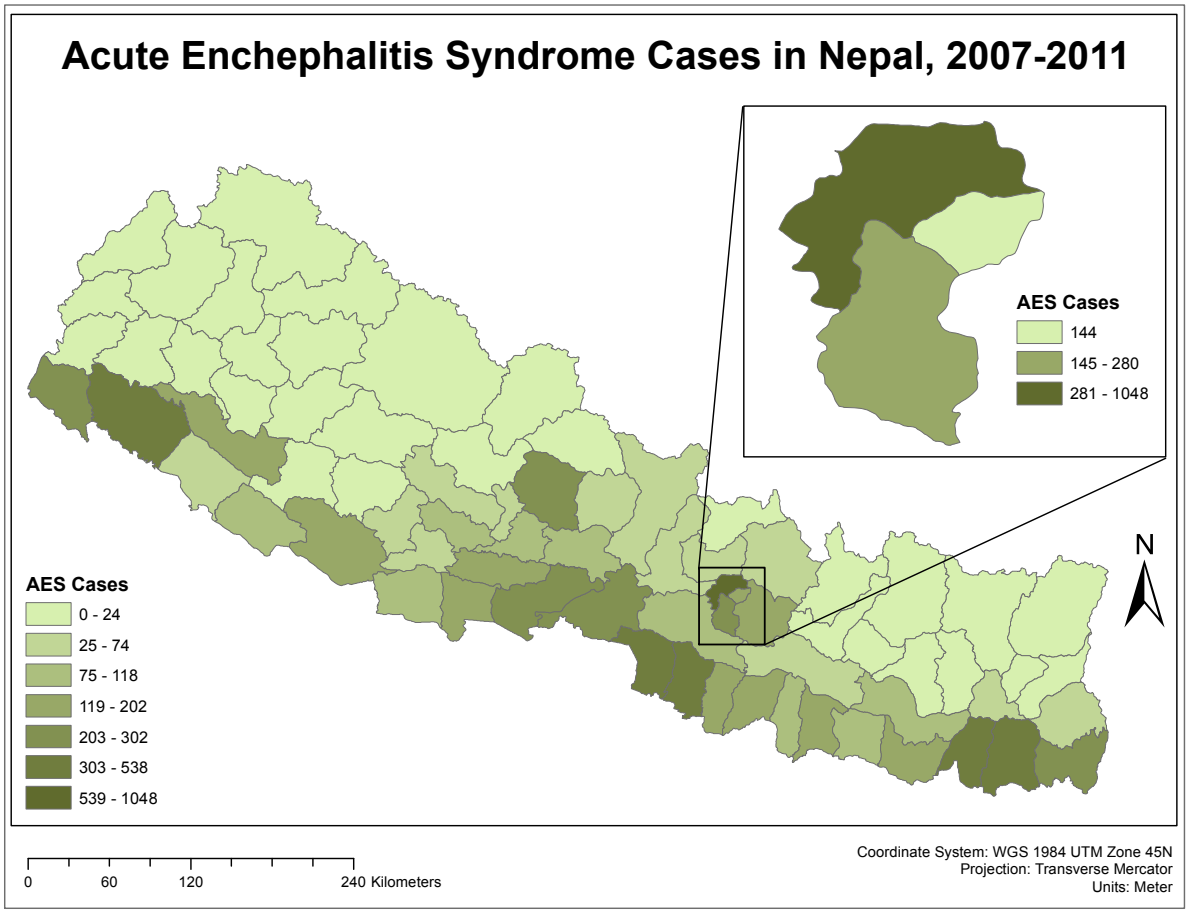


Figure 4.15: The spatial distribution of reported acute encephalitis syndrome cases in Nepal, highlighting the number of reported cases in the Kathmandu Valley, during the period 2007-2011.

antibodies to test for the presence of the pathogen (JEV). This test is often costly and involves technical training (Joshi and Stephen, 2013).

Data for the years 2007-2013 are used for analysis, two more years than the AES data. For this period, a total of 1576 JE cases were reported for all of Nepal, 248 of these occurred in the three districts that make up the Kathmandu Valley. The mean age for reported cases was 18.17 ($SD = 19.7, range = 0 - 90, median = 5$), with more males (928) than females (648) testing positive for JE.

Figure 4.16 (p. 77) displays the temporal dynamics of reported JE at the country-level for Nepal cases during the period 2007-2013 and Figure 4.17 shows the temporal dynamics of reported cases in the Kathmandu Valley for the same time period. Just as shown in the AES data, there is a peak in the number of cases in the months of August and September for all years, exemplifying the seasonal dynamics of the disease. This seasonality is more apparent when country-wide statistics are analyzed. Reported cases in the Kathmandu Valley do peak

in the late summer months, although the seasonal trends is less apparent due to the reduced number of cases.

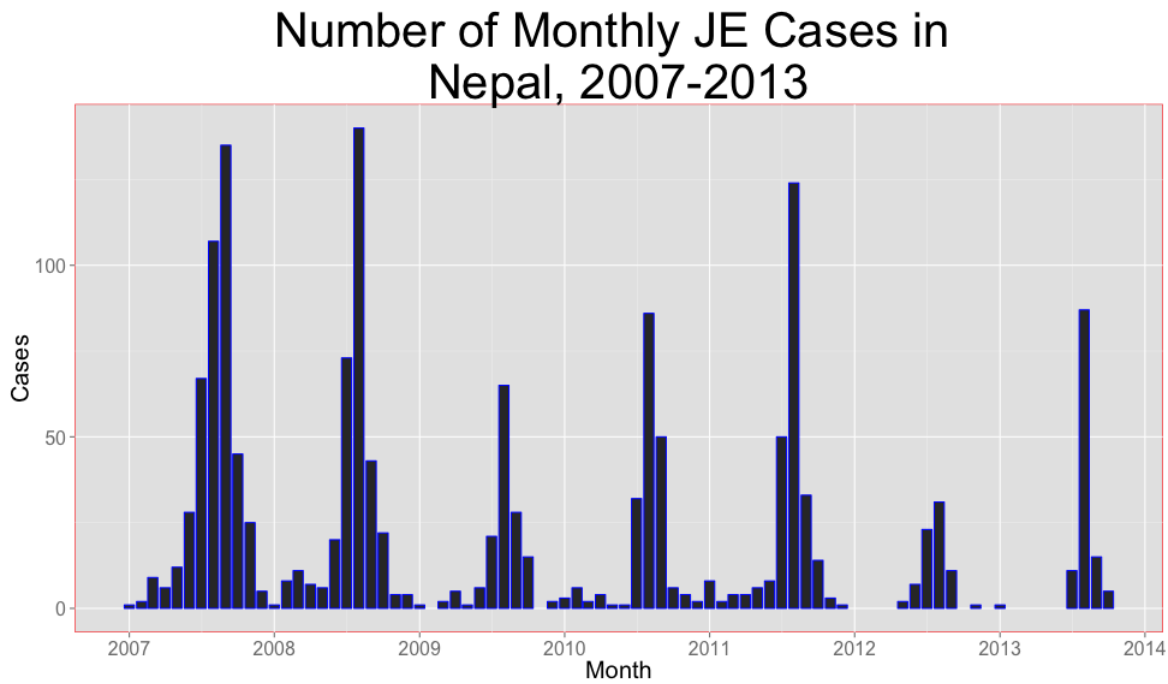


Figure 4.16: Temporal dynamics of reported Japanese Encephalitis cases at the country level for Nepal for the period 2007-2012. Number of cases are shown as monthly totals.

Figure 4.18 (p. 84) shows the spatial distribution of reported JE cases at the VDC level. It is apparent that most reported cases of JE are within the most populated VDC, the Kathmandu Metropolitan City. The VDCs surrounding the Kathmandu Metropolitan City also have higher population densities and greater access to healthcare.

4.8 Model Output

In this section, the temporal and spatial characteristics of our model output are described and compared to the AES and JE surveillance data. Section 4.8.1 will evaluate the temporal characteristics of the model's output. To do this, the spatial variation/distribution is excluded. The primary goal of this analysis is to view the seasonal variation in the number of reported cases per week, which is associated with increases in vector prevalence. Section 4.8.2 will show the spatial distribution of model output. The number of cases for the entire model run period (the year 2007) are shown spatially in Figure 4.21 (p. 86).

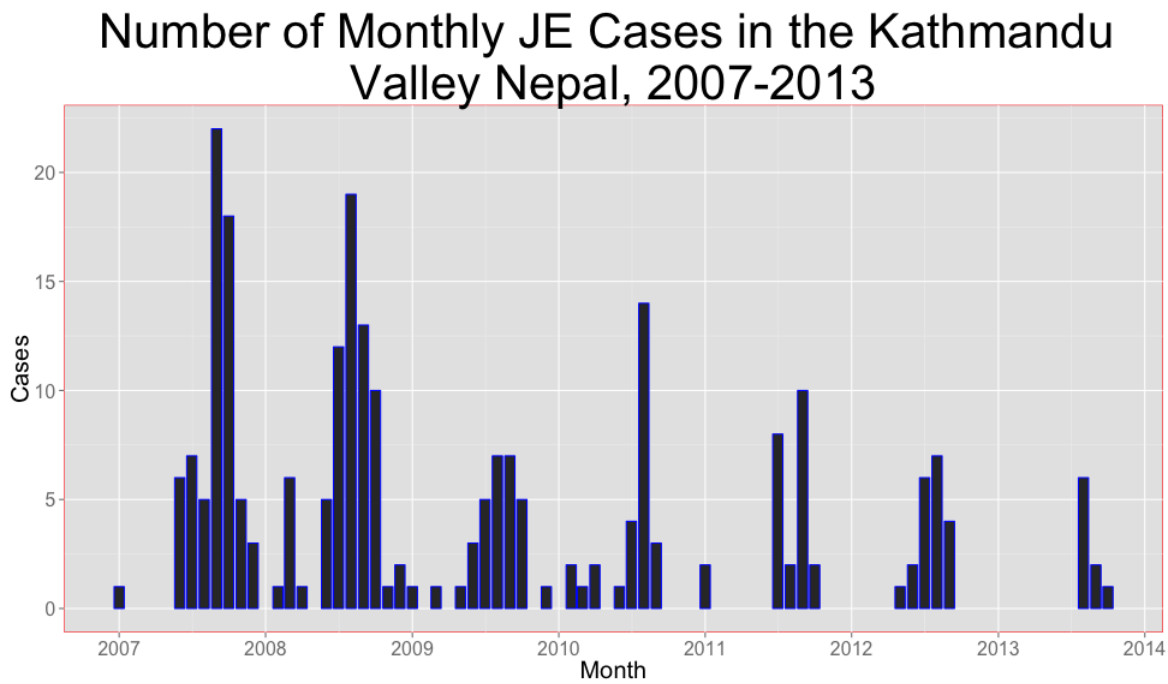


Figure 4.17: Temporal dynamics of reported Japanese Encephalitis Syndrome cases for the Kathmandu Valley (districts of Kathmandu, Lalitpur, and Bhaktapur) for the period 2007-2013. Number of cases are shown as monthly totals.

4.8.1 Temporal Analysis

Figure 4.19 (p. 85) displays the temporal distribution of 50 runs of the model. Initial values of the model were stated in Section 4.2.2, and were static for all model runs. The plot displays the number of cases reported at a weekly time scale. It can been seen that there is a seasonal variation of cases. Figure 4.20 (p. 85) shows the results for four random model runs, to give readers a sense of the variability.

The root mean square error (RMSE) is defined as the sample standard deviation of the differences between predicted and observed values. It is generally used to report a measure of the differences between model output and observed data. The RMSE is obtained by;

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{x}_i - x_i)^2}, \tag{4.18}$$

where n is the number of different predictions, $\hat{x}_i - x_i$ represents the residuals between actual observed values and model output ($\hat{x}_i$ defined as the predicted values from the model output and x_i defined as the observed value).

To compute the RMSE for both AES and JE surveillance data, the reported cases were aggregated to totals per week, using the `stats` package in R and compared to the aggregated weekly model output cases. Mean number of weekly cases for all 30 runs were used to compute the RMSE, instead of computing the RMSE for each model run. The `rmse` function in the `hydroGOF` package in R were used to compute the RMSE. This measure is shown in Table 4.11 and was used to measure the difference between the model output and the surveillance data. An RMSE of 0 would indicate that there is no difference between the observed surveillance values and predicted model values. Table4.11 indicates the magnitude of difference between the model output and the surveillance data.

Table 4.11: The root mean square error and the mean squared error between temporal model output (cases per week) and the weekly number of reported Japanese Encephalitis and Acute Encephalitis Syndrome cases.

	JE Surveillance Data	AES Surveillance
RMSE	3.171	4.395

4.8.2 Spatial Pattern Analysis

Figure 4.21 (p. 86) shows the spatial distribution of the mean number of cases at the VDC level for 50 model runs. To analyse the spatial pattern of the output distribution, a Moran's I, which represents the degree to which features are spatially autocorrelated, was calculated. In typical analyses of spatial pattern, the null hypothesis states that features exhibit Complete Spatial Randomness (CSR), of either the features themselves and/or their associated values. In this section, we will look at the spatial autocorrelation of the number of JE cases reported at the VDC level for both the JE surveillance data (AES data is aggregated at larger administrative boundaries, so will not be used for this analysis).

Spatial Autocorrelation

Moran's I is a measure of the degree to which a set of features, and their associated data values, tend to be clustered together in space (Moran, 1950)). Moran's I is obtained by;

$$I = \frac{n}{S_O} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2}, \tag{4.19}$$

where z_i is the deviation between an attribute's value from its mean, $w_{i,j}$ is the spatial weight between *i* and *j*, *n* is the total number of features (number of VDCs), and S_O is the aggregate of the spatial weights (Moran, 1950). Moran's I value was calculated for both the JE surveillance data and the model's mean of 50 model runs.

The 'Spatial Autocorrelation (Moran's I)' tool was used in ArcMap v.10.2 to compute the Moran's I value. The 'contiguity edges only' method of spatial relationship conceptualization was used. This method computes the Moran's I based on feature adjacency, rather than a distance threshold. Table 4.12 shows the Moran's I values, and associated spatial autocorrelation statistics for both datasets.

Table 4.12: Spatial autocorrelation statistics for JE surveillance data and the model output.

	JE Surveillance Data	Model Output
Moran's Index	0.152082	0.321913
Variance	0.000977	0.001429
z-score	5.137046	8.739759
p-value	0.00000	0.00000

Moran's I values and their associated statistics are shown in Table 4.12 for both JE surveillance data and the model's output. Both z-scores and p-values returned from the Moran's I analysis for both JE surveillance data and model output data indicate that the null hypothesis can be rejected. In other words, the spatial pattern of the number of cases reported at the VDC level is not considered to be due to CSR. In both datasets (JE surveillance data and the model's output), there is some degree of positive spatial autocorrelation. There is a slightly higher degree of spatial autocorrelation for the model's output than the JE surveillance data. The Moran's Index was used here to give the reader a sense of similarity, or difference, between the surveillance data and the model's output.

Hot Spot Analysis

The local Getis-Ord statistic ($G_i(d)$) (Getis and Ord, 1992) takes a moving average of a numeric variable x within a spatial neighbourhood defined by a spatial weight matrix $w_{ij}(d)$ (where d is the maximum distance at which nearby points are included as neighbours) as;

$$G_i(d) = \frac{\sum_j w_{ij}(d) x_j}{\sum_{j=1}^{n} x_j} \tag{4.20}$$

over the total sum of x. In practice $G_i(d)$ is defined as a standard normal deviate, and therefore can be interpreted as a *z-score*. We employ the $G_i(d)$ statistic to identify significant hotspots in both the model's output and the JE surveillance data.

The Getis-Ord statistic was used to identify spatial clusters of significantly high disease in both the surveillance data and the model's output. Figure 4.22 (p. 87) shows the results of the Getis-Ord analysis. Results of both analyses show clustering around the VDC of the Kathmandu Metropolitan City and that the significant hot spot clusters are roughly similar for both model output and the JE surveillance data.

Figure 4.23 (p. 88) shows VDCs where our model predicts the likelihood of JE cases, or possibly any form of landscape-influenced case of AES, which was not apparent in the JE surveillance data obtained from the World Health Organization. This shows that these areas are of potentially high-risk due to their landscape, climatic, and population characteristics.

4.9 Alternative Scenario Analysis

To assess model parameters shown in Table 4.5 (p. 59), alternative analyses were examined by modifying author-estimated parameters. These parameters include the rain-weight, temperature weight, and the rice weight, which were introduced in Table 4.5. Table 4.13 displays how these initial parameters where adjusted and how the adjustments affected output results of the model. For each alternative scenario, the models was run 30 times and the mean total number of infected individuals, as well as other measures of central tendency, were reported (Table 4.13).

Table 4.13: Alternative scenario analysis parameter adjustment and subsequent model output. For each scenario, the model was run 30 times and measures of central tendency were reported for the number of infected individuals. The mean, median, and standard deviation are reported for the number of infected agents per model run. All parameter values have been scaled by 10^5.

Scenario	rain	temp	rice	Mean	Median	SD
Initial Parameters	100	0.001	1	207.07	203.5	15.78
Scenario 1	0.01	0.001	1	179.0	180.5	13.04
Scenario 2	0.001	0.01	1	176.46	179.5	17.33
Scenario 2	0.01	0.01	1	177.80	172.5	21.94
Scenario 3	0.001	0.001	0.1	64.50	63.0	13.21
Scenario 4	0.001	100	1	432.67	435.5	22.45
Scenario 5	0.001	0.001	1	176.03	178.5	16.96

Table 4.13 shows the initial parameters with which the model was run and adjusted values for various scenarios, which were used to assess the sensitivity of these parameters. The columns labelled 'Mean', 'Median' and 'Standard Deviation' show how modifications to the model parameters influence the model output.

Initial values of the model parameters are shown in row one of Table 4.13. It is apparent that rain receives a much greater weight than does temperature. This is due to the raw values of each variable and how they were normalized. For example, throughout the year in Nepal, rain is generally low (e.g. 25-50 mm per week). To normalize, these values were divided by the maximum value for that variable (e.g. 25 / 350), resulting in a very low value. Temperature, on the other hand, was less variable and normalized values were generally larger (e.g. 22 / 37).

Hence, the weighting scheme shown in Table 4.13.

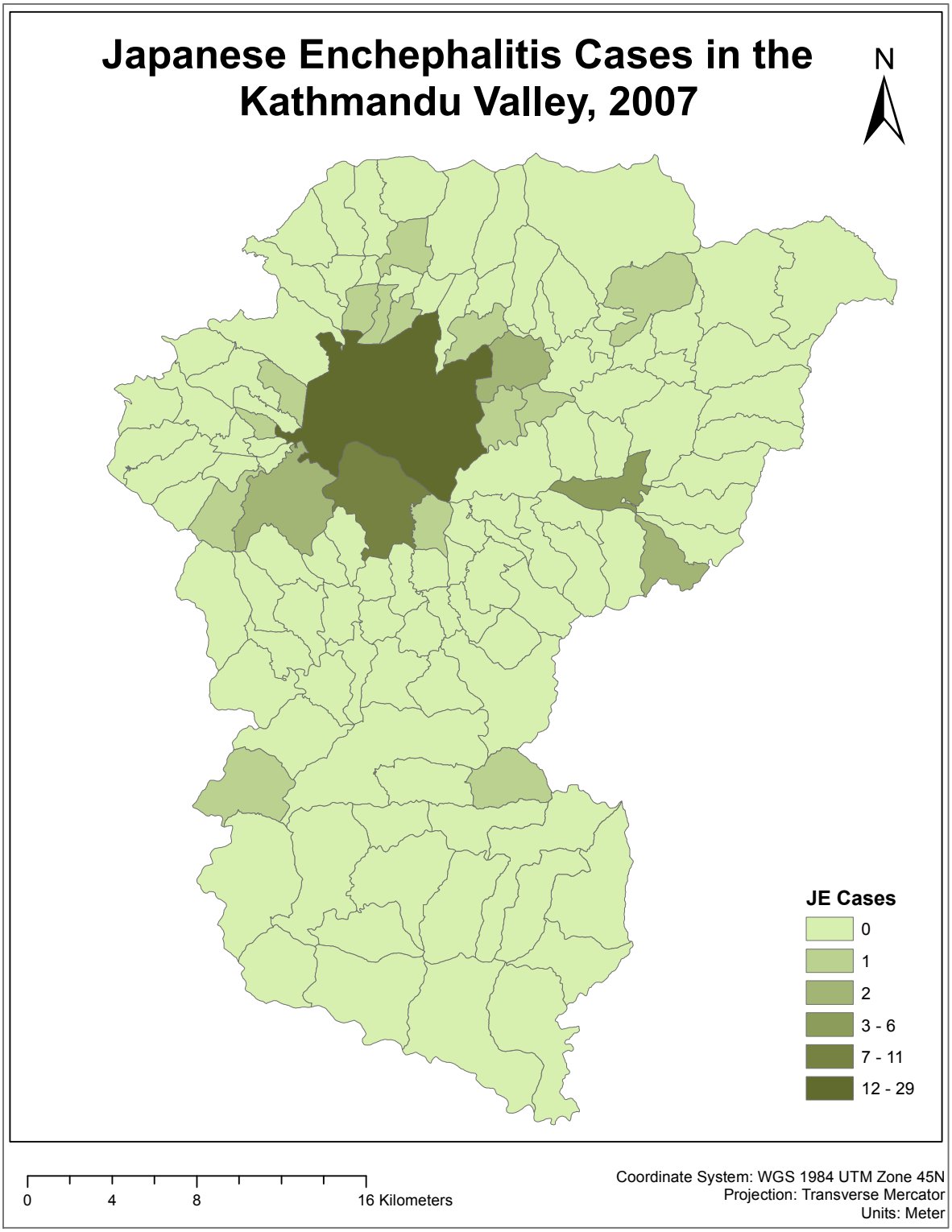


Figure 4.18: Spatial distribution of Japanese Encephalitis cases reported to the World Health Organization's Programme for Vaccine Preventable Diseases for the year 2007.

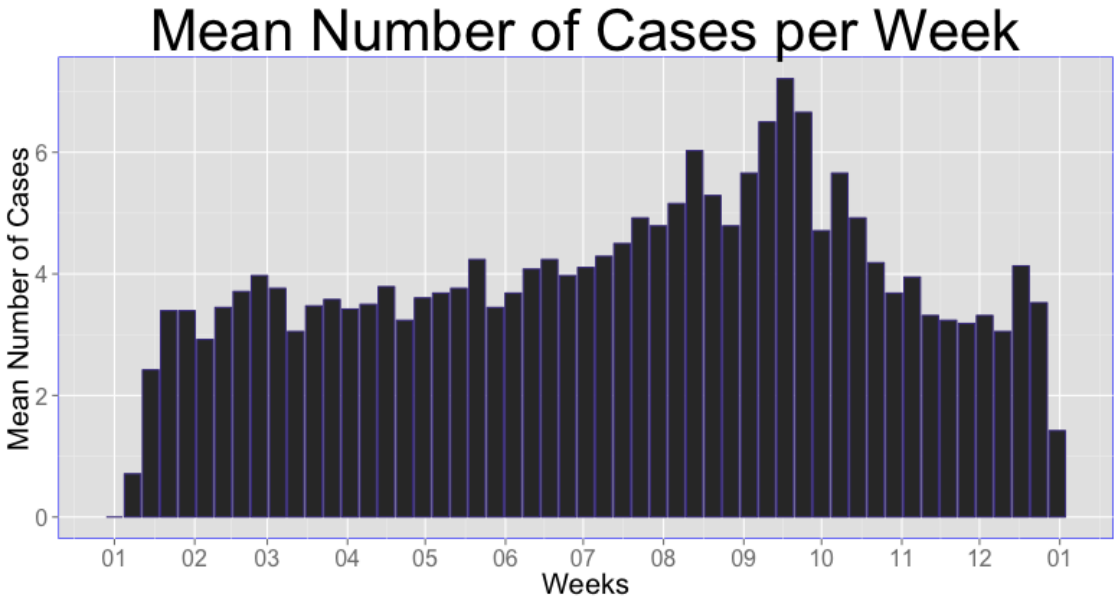


Figure 4.19: Temporal model output. Graph displaying the mean number of cases per week from 30 model runs.

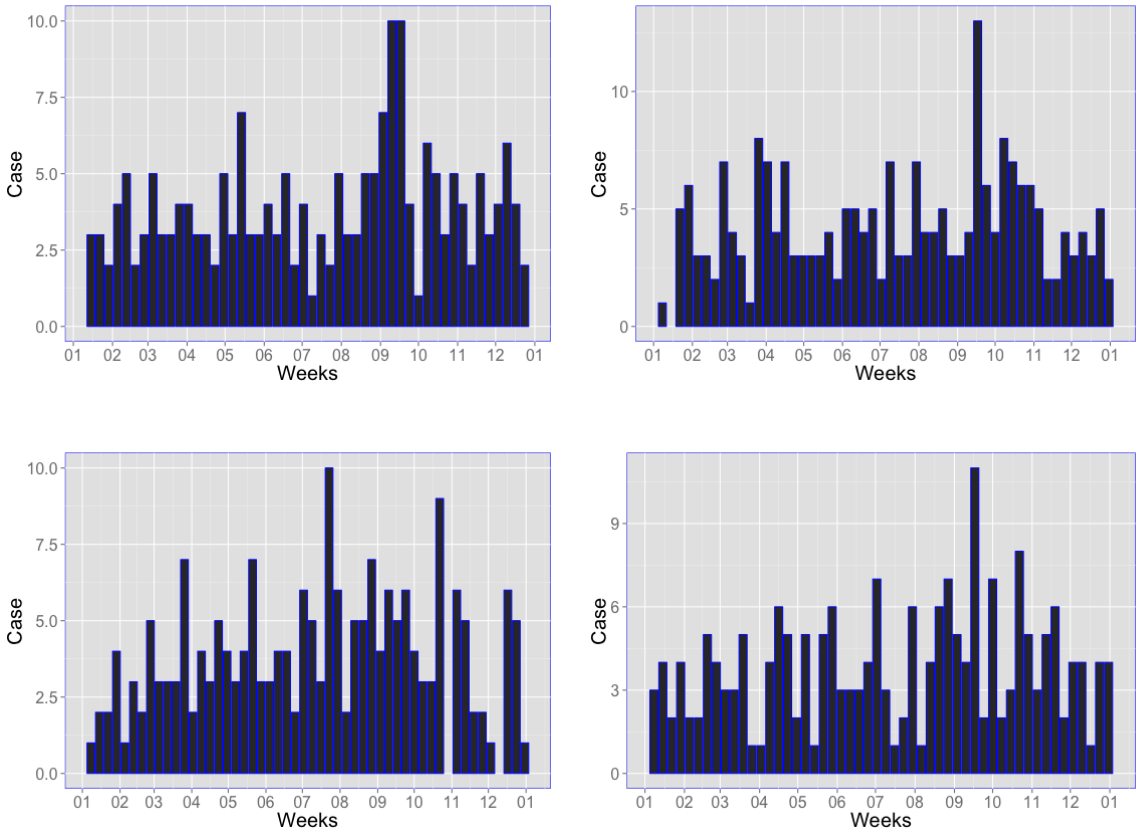


Figure 4.20: Results of randomly chosen set of 4 model runs. Temporal patter shows an increase of infected individuals in the later months, which roughly corresponds to the surveillance data.

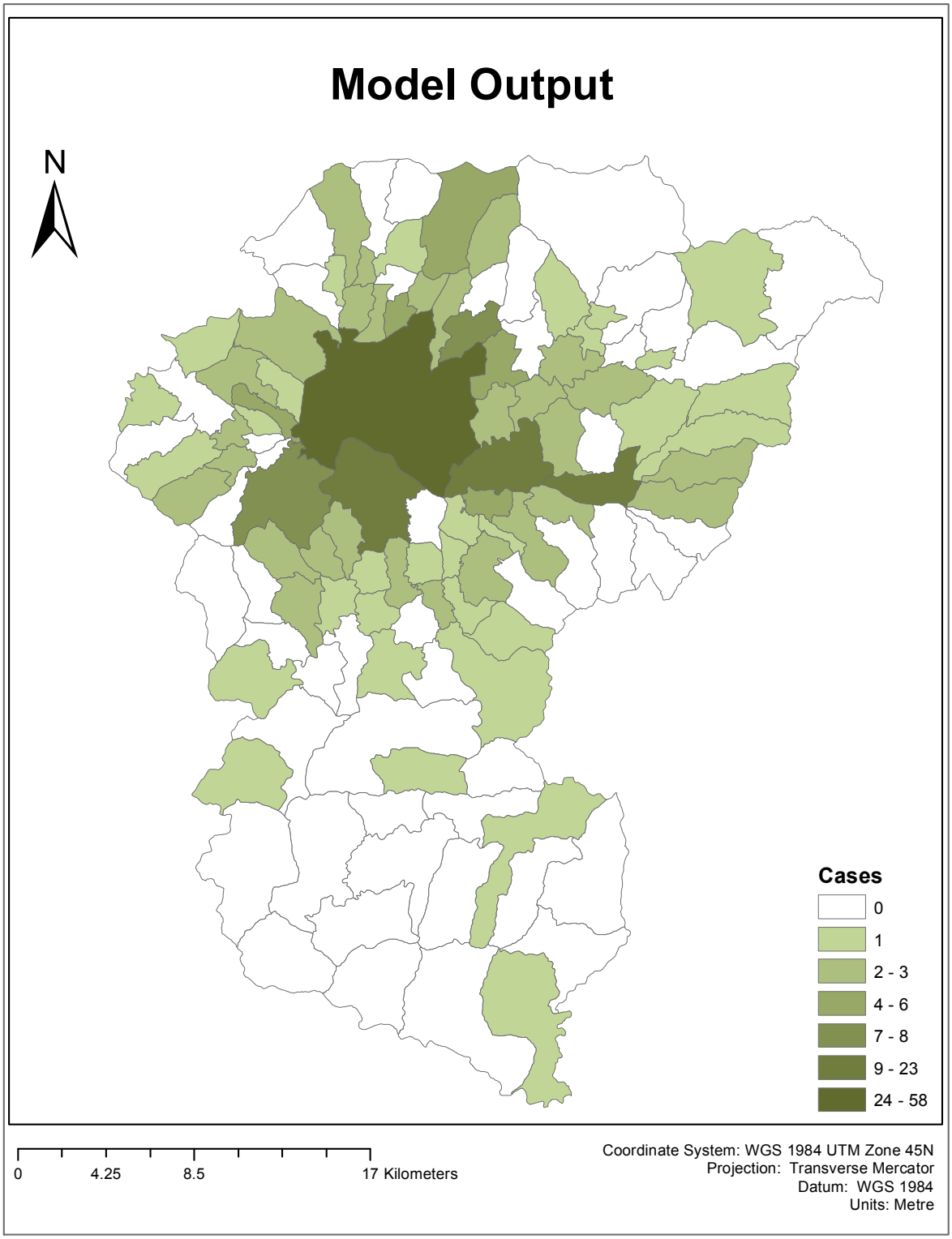


Figure 4.21: Spatial distributions of model output results. The current maps shows the mean number of cases at the VDC level for 50 model runs.

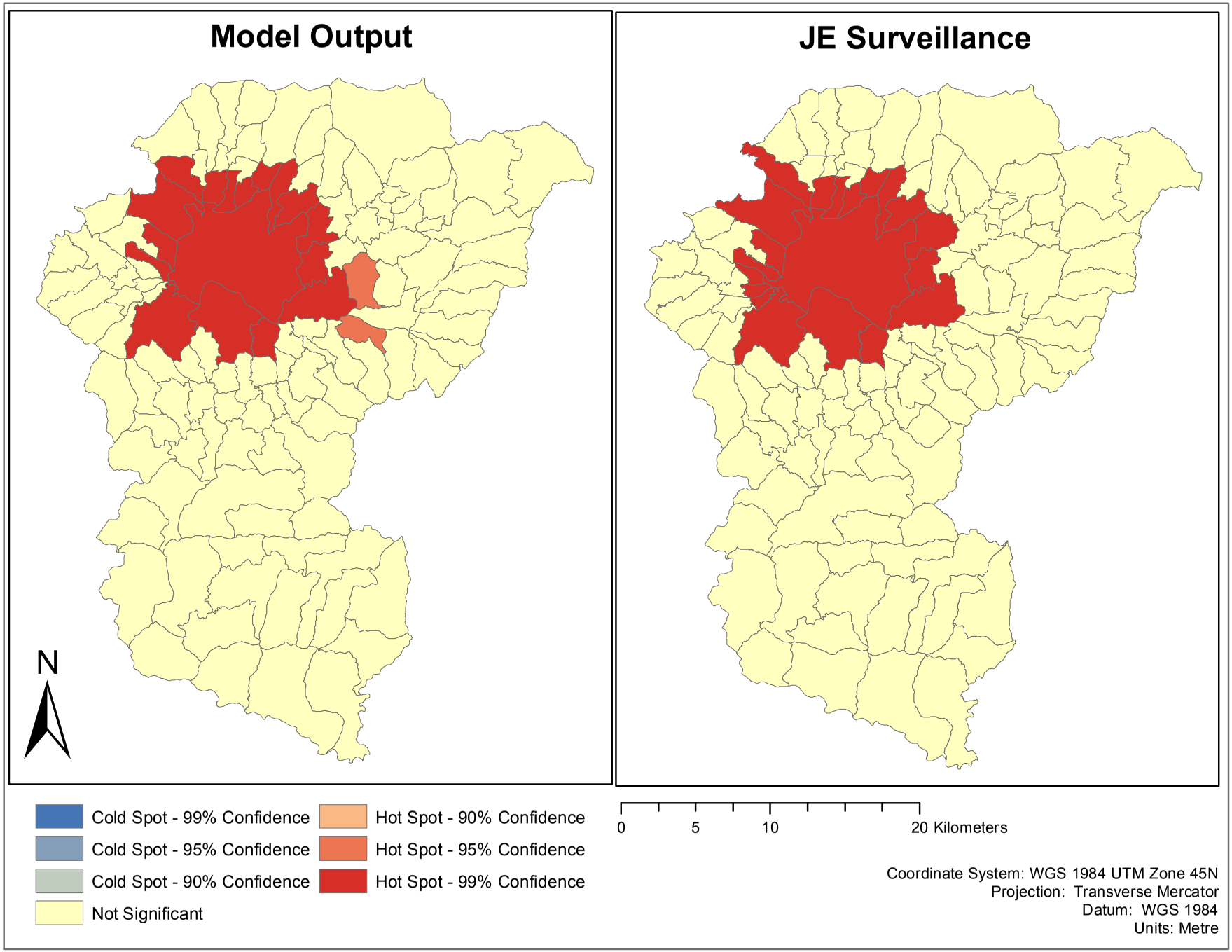


Figure 4.22: Getis-Ord analysis results for model output and JE surveillance data. Significant hotspots are clustered around the Kathmandu Metropolitan City and are shown in red.

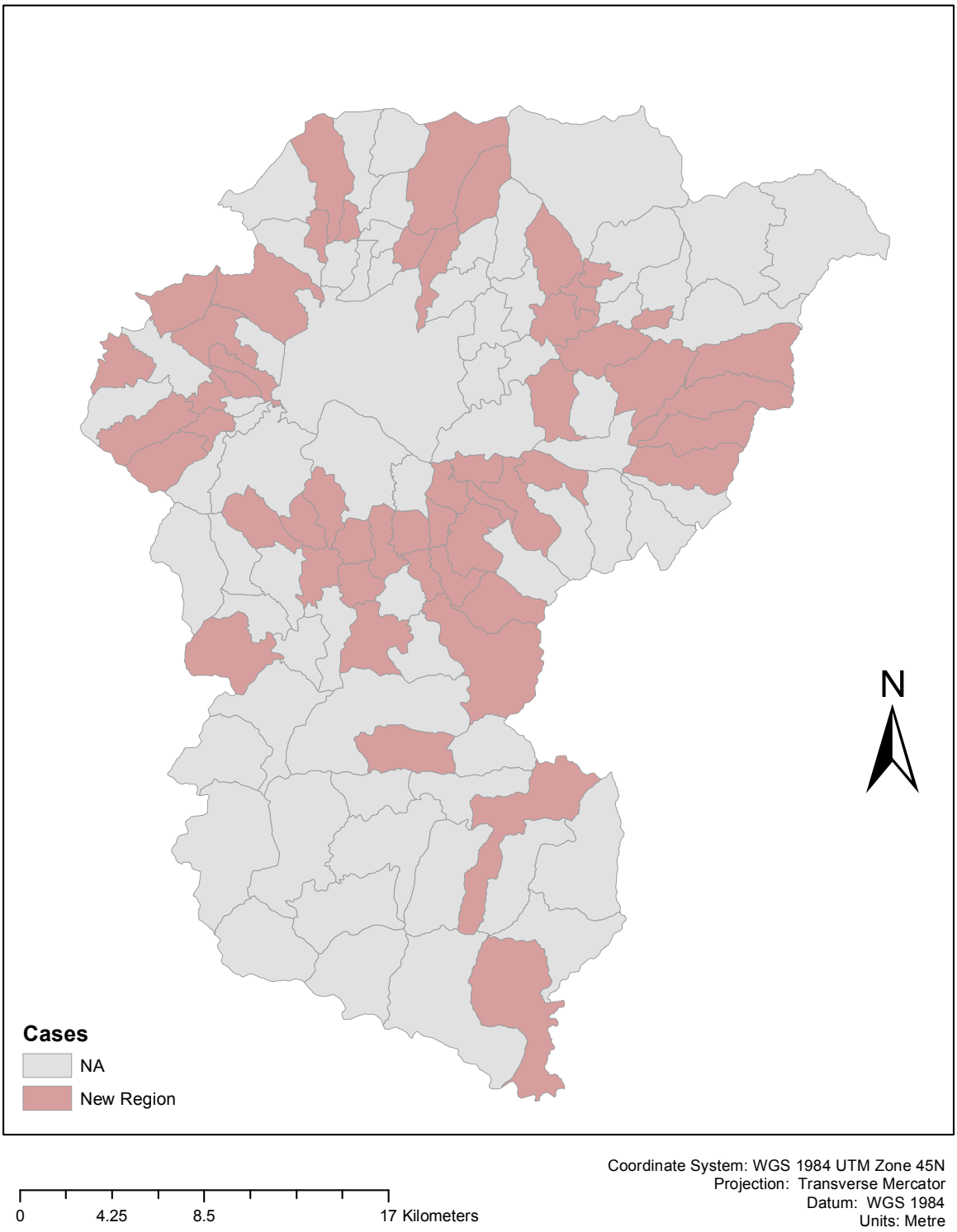


Figure 4.23: VDCs that exhibit cases of JE in the model run, but not in the JE surveillance data. VDCs represented by gray colour are those in which both model output and the surveillance data repot cases of JE (or lack of), and VDCs represented by red are those that our model predicts cases, but the JE surveillance data shows no cases.

Alternative scenarios show what happens to the model output, which is shown in Figure 4.25 on page 90. For example, when the weighting for rain and temperature variables are switched, the mean number of cases per model run, as well as variability, dramatically increases.

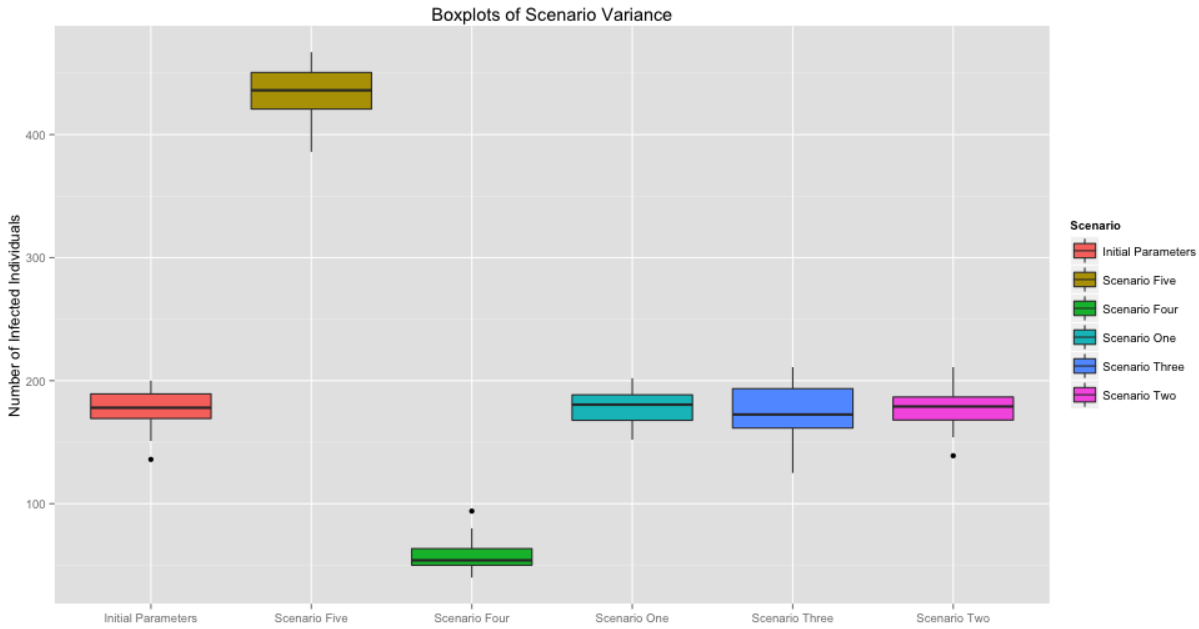


Figure 4.24: Boxplots of sensitivity scenarios displaying output variance (number of infected individuals per model run) for 30 iterations of the model with adjusted parameter values.

4.10 Null Model Evaluation

To assess the quality of the model presented in this study, a null model was employed. The 'risk surface', which was discussed in Section 4.2.2 (p. 49), was set to a constant value and did not change according to temporal dynamics of temperature and precipitation values. All grid cells for this analysis were set to the mean *risk-probability-value* for all patches, for the entire year, for 30 model runs (1.6205×10^{-4}). This method was employed in order to test the validity of the dynamic cell values, which were influenced by weather and landscape variables.

Figure 4.26 (p. 97) displays the mean number of cases reported at the VDC level for 30 null model runs. Results indicate a similar spatial distribution as the original model output. However, this was expected due to higher population densities in these regions. A greater population leads to a higher probability that individuals will be randomly selected to be set to

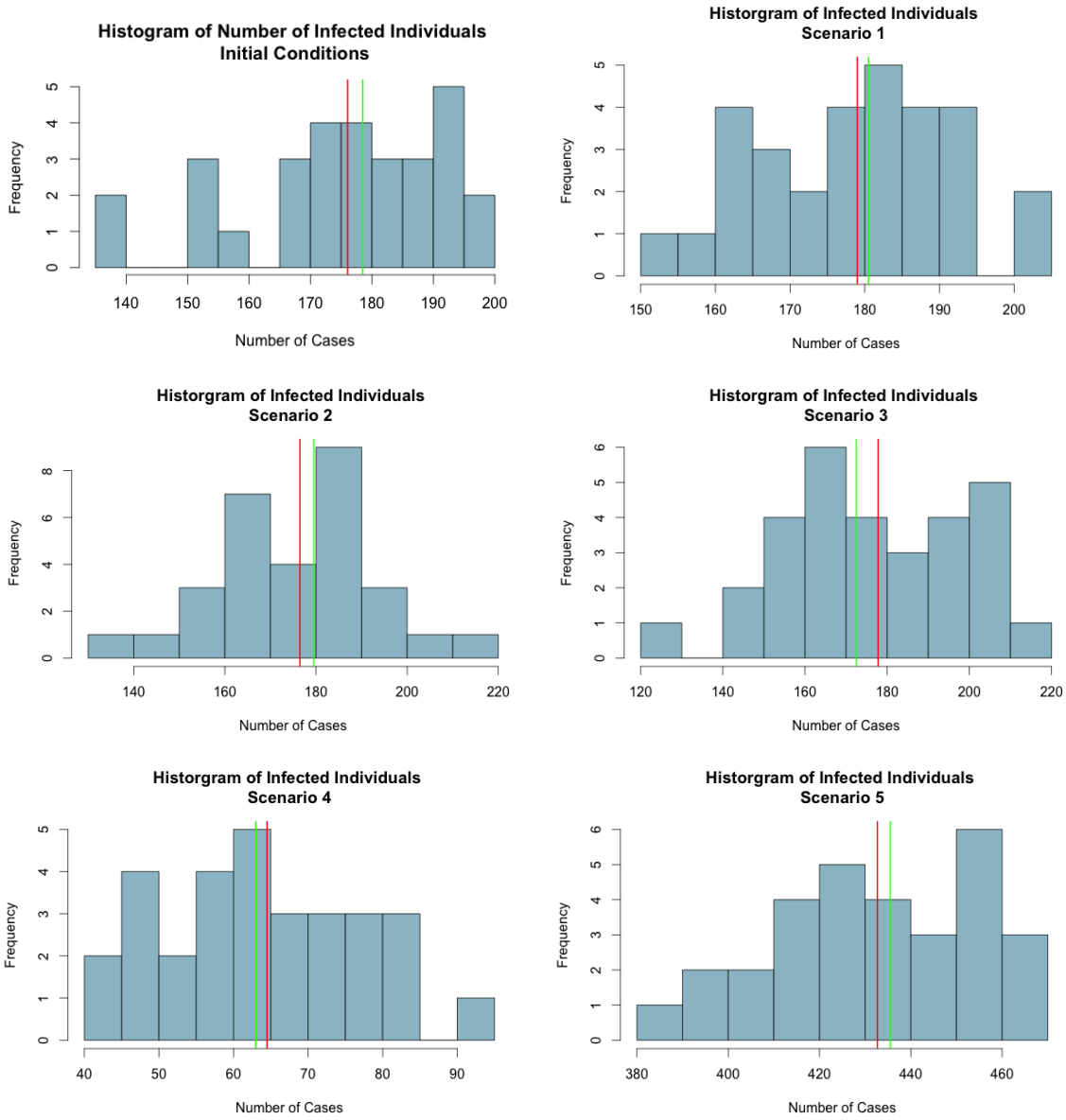


Figure 4.25: Histograms for total infected individuals per sensitivity scenarios. The mean and median for the number of infected individuals per model run are shown as red and green vertical lines, respectively.

infected.

Figure 4.27 (p. 98) shows the temporal dynamics of the mean number of cases at the VDC level for 30 null model runs. It is apparent that the seasonal dynamics of infections is not present as shown in the initial model output (Figure 4.19 on page 85). Results from the null model analysis indicate that the temporal dynamics in the initial model, which are influenced by seasonal variation in weather variables, are important components of this research.

4.11 Discussion

This research presents an individual-based modelling perspective to analyzing the epidemiology of Japanese encephalitis in the Kathmandu Valley. By employing environmental data and a few basic rules that govern how landscape factors can lead to and individual's increased probability of contracting JEV, we are able to reproduce temporal and spatial characteristics similar to JE and AES surveillance data. This model can be used to analyze various scenarios, such as the impacts of climate change or immunization programs on the human incidence of JE in the Kathmandu Valley. It is hoped that this model can be slightly modified to examine the dynamics of JEV in other regions and also applied to similar vector-borne diseases, such as the West Nile virus.

One of the primary findings of this research is that the spatial distribution of the model output was more dispersed than the surveillance data. As shown in Figure 4.18 (p. 84), Figure 4.21 (p. 86), and Figure 4.23 (p. 88) more cases were reported in rural VDCs in the model output than in the surveillance data. Peri-urban VDCs that did not show any cases of JE in the surveillance data, where landscape and population characteristics were roughly similar to regions that did report cases, were more likely to produce JE cases in the model. This was a hypothesized outcome of the model as it potentially represents the issue of non-reporting in rural regions. Non-reporting in rural regions may be caused by different medicinal treatment preferences (e.g. preference for traditional healers rather than hospital treatment) and/or limited access to healthcare facilities.

4.11.1 Limitations

One of the primary concerns with model validation lies with inherent issues of AES and JE surveillance data. AES surveillance data, although possessing a greater number of reported cases and thus providing better reliability than the JE surveillance data, is only reported at the district level (Figure 4.15, p. 76). Spatial analysis and accuracy are thus sacrificed. For these reasons, the use of JE surveillance data is valuable due to it's greater precision. However, there are also issues in reporting JE cases. Many cases of JE go unreported due to lack of healthcare facilities in rural areas and/or cultural beliefs (preference for traditional healers over hospital care). For example, Kakkar et al. (2013) showed that JE surveillance data was highly unreliable for decision-making and prevention strategies. Because of these issues with both sets of surveillance data, and due to previous research by Robertson et al. (2013), this research makes the assumption that the number of JE cases should lie between JE and AES surveillance numbers.

Another limitation with the present model is that many assumptions are made regarding population mobility within the Kathmandu Valley. Infectious disease transmission is highly dependent on human mobility, involving both internal/external migration as well as daily mobility patterns. Interaction between virus amplifying host and other landscape factors are the primary influence on the presence/absence of JE. The current model would be considerably improved if micro- and macro-scale human mobility patterns were incorporated.

The spatial distribution of JEV's amplifying hosts is a crucial component in determining landscape risk of many vector-borne diseases. In the current model, we implement pig population data obtained from Nepal's Ministry of Agricultural Development, which reports the number of pigs at the VDC level. The precise actual location of pig farms, however, was unknown. We based our initial placement of pig farms on knowledge obtained by JE experts in Nepal from the National Zoonoses and Food Hygiene Research Centre in Kathmandu, Nepal. JE experts stated that pig farms are more likely to be located near river banks, where pig farmers can access river water for various farm activities. Thus, in the presented model, pigs were stationary entities located, at a random location within 50 m of a river.

Agents that represented JEV's amplifying host – pig agents – did not exhibit dynamic

behavioural rules. Pigs remained stationary throughout the duration of the model. Although the model period was only one year, potential mobility of pigs was not included. Joshi (2006) outlined the pig farm and pig meat market system. This work described how pigs are transported from the farm and distributed throughout the market system. An infected pig is thus able to transmit the JEV along this route, adding complexity and increasing risk to urban areas. This subsystem has not been included in this model. Although it was omitted due to lack of accurate data on this subsystem and thus to reduce uncertainty, such a process can play an important role in JEV transmission.

Section 4.8.2 describes the spatial distribution of the model's output, which was similar to the spatial pattern of the JE surveillance data. Although this was a positive outcome of the simulation, there may be other factors contributing to this similarity. One interpretation of this result is that the population density in VDCs where most cases were reported in the model output is much higher, affecting the ratio of non-infected to infected agents. Although this is inherent in incidence estimations (incidence being a measure of probability), and it is quite possible that the higher rates in the number of cases in high population-dense regions in the model output is the main contribution to the similar spatial pattern to actual JE surveillance data, this could be the reason why the spatial pattern of JE surveillance data is the way it is.

The current model does not incorporate human behaviours and their influence on micro-scale land-use practices. Agent-based models have been used to investigate the effects of human behaviours and land cover/land-use change (Deadman et al., 2004; Rindfuss et al., 2008; Brown and Robinson, 2011; Robinson et al., 2012). Micro-scale processes may be influencing land-use patterns that are not represented in the current model. We attempt to account for this by incorporating the primary land-use factor that leads to mosquito-borne disease risk (i.e. the planting of rice-paddy fields, which occurs during the monsoon season (Hecker et al., 2013)).

4.11.2 Future Research

Scenario analysis regarding issues such as climate change, land-use and landcover change, and the evaluation of vaccination programs can be implemented using the model presented in this thesis. JE is highly dependent on the presence/absence of it's mosquito vector. An

increase in vector abundance is probably the most important factor in determining the number of JE cases, both in humans and in pigs (Hecker et al., 2013). Increases in temperature will undoubtedly affect *Culex* numbers, increasing their feeding/biting behaviours and speeding up life-cycle development. Increases in temperature can also increase the at-risk human population, influencing the amount of time spent outdoors. Our model currently uses climate data that can easily be modified to influence vector abundance. The following three paragraphs discuss possible scenario analyses that can be implemented with the presented model.

Increasing temperatures and seasonal variation of rainfall, will undoubtedly have an effect on human JE incidence. Hecker et al. (2013) showed how Nepali experts in the field of JE epidemiology anticipate increases in temperature and precipitation will positively influence the number of human JE cases in Nepal. Also, a thorough review of the literature conducted by Dhimal et al. (2015) concluded that climate change will intensify vector-borne disease epidemics in the mountain regions of Nepal. The model has the potential to analyze climatic scenarios under varying conditions. By increasing temperatures up to 2°C, we will show how the spatial distribution of human JE cases may change; most importantly, showing how JE cases may increase in areas of higher elevation.

The spatial configuration of pig farms will influence the rate of JEV transmission. In our model, we assumed that non-infected pigs that are in proximity to infected pigs are at higher risk of becoming infected themselves. Thus, a higher density of pigs, especially near rice paddy fields, where mosquito density is higher, will increase the risk of acquiring the JEV.

Vaccination programs for livestock are especially costly for resource poor countries such as Nepal. Thus they are rarely implemented, undoubtedly influencing the rate of transmission within the system. Considering that the number of infected pigs contributes to landscape risk, a useful implementation of this model would be to examine pig vaccination programs. Questions regarding the spread of JEV and the influence of infected amplification hosts could be examined using such a model.

The implementation of human behaviours into the present model would provide valuable information for analyzing JE education programs. Individual-scale behaviours, such as the use of insect repellent and/or mosquito nets, reduce an individual's likelihood of contracting the

JEV. Educational programs, such as those implemented by the National Zoonoses and Food Hygiene Research Centre in Nepal, could be assessed. Educational programs, which involve dedicated teachers and financial assistance, could target regions of greater risk as identified in the current model's spatial output.

A recent project called 'FlowMinder' (http://www.flowminder.org/) was initiated for the purposes of disaster response and disease outbreak management. FlowMinder combines billions of data points from mobile operator, satellite imagery, and household survey data to assess at-risk human populations. FlowMinder recently used mobile data from Ncell, Nepal's primary mobile provider, to track displaced populations after the 7.8 magnitude earthquake that occurred on 25 April 2015. FlowMinder estimated that approximately 1.8 million people above normal levels were displaced from their home districts, causing dependence on sources such as Government of Nepal (2014) to be obsolete. The implementation of FlowMinder data would considerably improve the validity of the current model's output.

Landscape characteristics remain static in the presented model. Satellite imagery retrieved from the RapidEye constellation was used to classify landcover of the study area. This provided high spatial resolution of the landscape. However, high temporal resolution was sacrificed for high spatial resolution. Future research, or adjustments to the current model should implement Landsat imagery, which would allow for a changing landscape throughout the study period. Another limitation relating to landcover is that satellite imagery for many regions in Nepal is not available during the monsoon season due to constant cloud cover. It is essentially impossible to obtain a satellite image of this region during the rainy season. Beneath-cloud cover, or drone obtained imagery, would be too costly and time-intensive for the purposes of this study.

As stated in Section 4.11.1, market systems for the transportation of pigs and pig meat could potentially spread JEV. This system was omitted from the current model due to lack of data on this system. Future work should assess the impact of this system on the transmission of JEV. The inclusion of an estimated system could be beneficial for analyzing potential scenarios, however the omission of this system reduces added complexity and introduced uncertainty due to this complexity.

4.11.3 Conclusions

The agent-based model presented in this research, although needing refinement, closely reproduces the temporal and spatial characteristics of reported JE surveillance data. The initial hypothesis presented in Section 1 stated that we expected output results to be less similar to JE surveillance. It was expected that model results would show a more dispersed spatial pattern of cases, representing the variation in disease reporting behaviours between rural and urban populations (e.g., rural poor seeking traditional healers regarding an illness rather than making a hospital visit, or issues of healthcare accessibility). This is a positive outcome of the model, as there is potential for it to identify high-risk regions.

The presented model in this thesis is a relatively simplistic representation of the complete JEV transmission system. The primary factors that dictate JEV risk to humans in this model were landscape factors (proximity to agricultural land and pig farms) and climatic variables (maximum and minimum temperatures, as well as weekly precipitation values). Although it is necessary to mode such a complex system with only the most important components, other factors that influence JE risk in humans were omitted. Factors such as occupation and mosquito avoidance behaviours, are not included in this model. The necessary data needed to inform individual agent behaviour was not available.

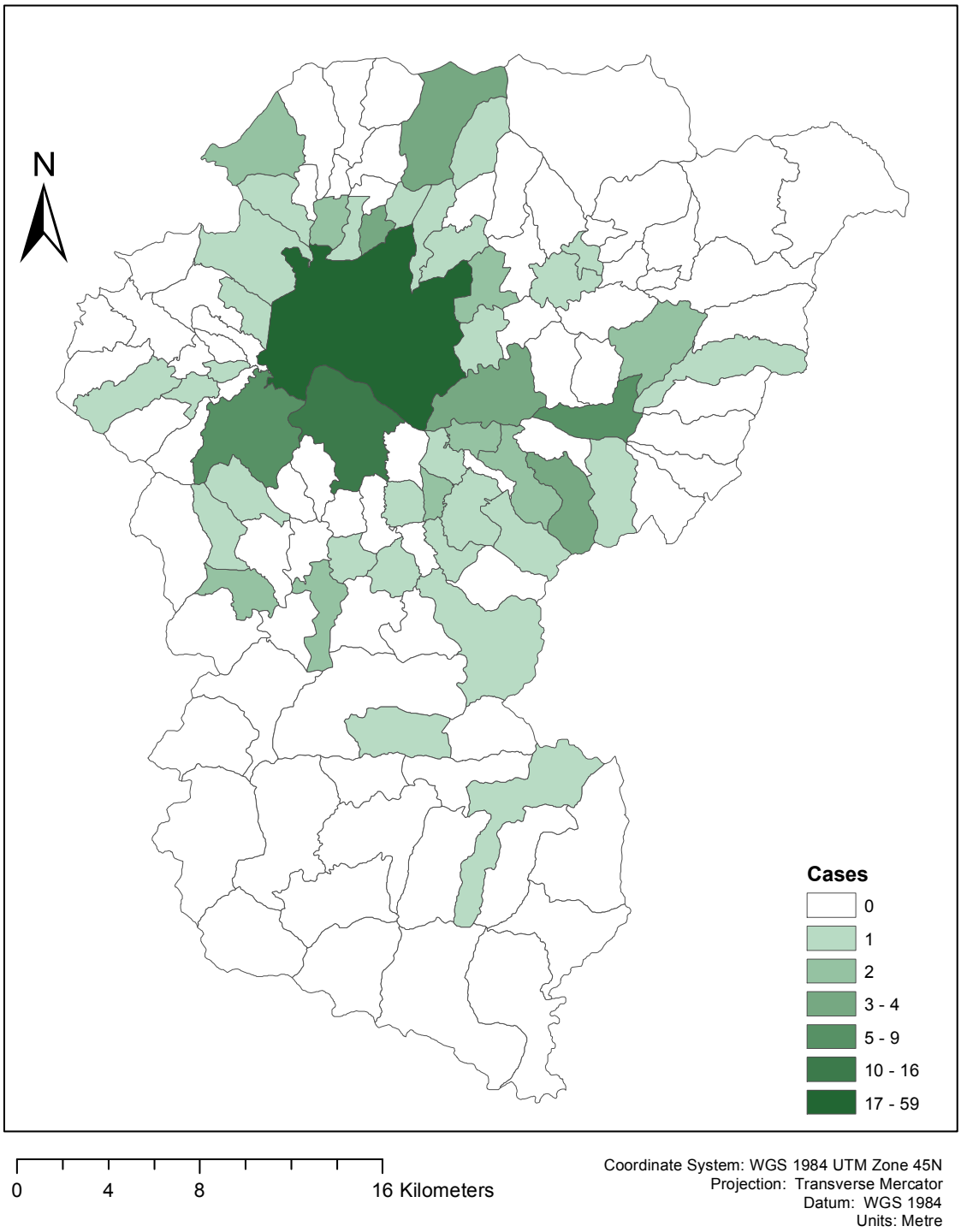


Figure 4.26: The spatial distribution of the mean number of cases from 30 null model runs. *risk-probability-values* of patches were all set to 1.6205×10^{-4} and remained static throughout the simulation.

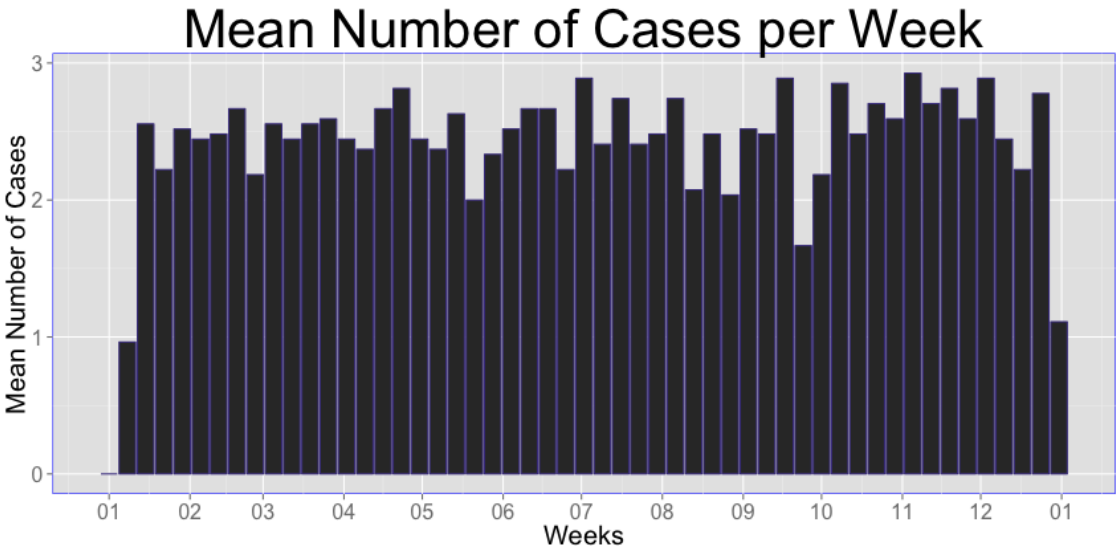


Figure 4.27: Temporal dynamics of the mean number of cases per week for 30 null model runs. *risk-probability-values* of patches were all set to 1.6205×10^{-4} and remained static throughout the simulation.

Chapter 5

Conclusion

Japanese Encephalitis is a relatively rare disease to contract, even for those living in high-risk regions. It is estimated that about 67,900 cases occur annually (Campbell et al., 2011; LaBeaud et al., 2011). To put this into perspective, approximately 198 million cases of malaria were reported for the year 2013, which caused approximately 584,000 deaths (World Health Organization, 2015a). Access to JE and AES surveillance data, as well as an established research connection with the National Zoonoses and Food Hygiene Research Centre in Kathmandu, Nepal, directed the research focus of this thesis. The study and the development of disease risk mapping and modelling techniques play a large role for guiding prevention and response strategies, and the allocation of resources for developing, or resource-poor regions. It is hoped that the methods presented in this thesis can be used to supplement current methods that guide intervention strategies to reduce the burden of disease caused, not only by JEV, but other vector-borne diseases as well.

The transmission cycle of the JEV is similar to that of other arboviruses and vector-borne diseases. For example, the West Nile Virus, which recently expanded into North America, is amplified by bird hosts and transmitted by *Culex pipens* mosquitoes. Landscape characteristics that contribute to promote mosquito populations would certainly influence the incidence of the disease within the human population. Other vector-borne diseases that would require a similar risk landscape as the one proposed in this thesis are yellow fever, Dengue, chikungunya, Rift Valley fever, and the zika virus, among others. The methods presented in this thesis, which

examine JE risk at various geographic scales, in theory, could be applied to o

Chapter 3 proposed a method for mapping potential future disease risk at the population level in relation to JEV's amplification hosts, pigs. The purpose of this method was to identify regions that will have high population densities and potential high virus host populations. The major assumption supporting this research was that a higher number of cases of JE will be reported if the at-risk human population is high in conjunction with high host populations. This generalization is one approach to accounting for the uncertainty or the probability of an individual contracting the disease.

Chapter 4 presented a method for examining the risk of contracting JEV at the individual level, using the Kathmandu Valley as a case example. The primary research goal of this paper was to develop a method for analyzing disease risk in association to landscape characteristics and weather patterns. It is hoped that this method can be used to examine various scenarios, such as the effects of precipitation or temperature shifts on disease incidence. The individual-based approach to disease risk modelling will also be useful for analyzing various interventions, such as the effects of vaccination programs. However, the focus of this thesis was to develop a model that is able to replicate the spatial and temporal distribution of surveillance data, so that future scenarios can be tested.

A finding of Chapter 4 from the null analysis (Section 4.10) showed that, even without a changing risk surface, more cases of JE occurred in more densely populated VDCs. This is most likely due to the higher probability of a case occurring within a greater population, even though the rate may be similar to VDC with lower case counts. This suggests that the motivation of Chapter 3, regarding the at-risk human population, is reasonable to assume.

Future work regarding Chapter 4 should implement human behaviours. For example, mosquito avoidance behaviours learned through Japanese Encephalitis education programs implemented by the National Zoonoses and Food Hygiene Research Centre, can be programed at the individual level. The presented model could then map the emergent spatial patterns of disease.

The ideal disease surveillance scenario for Japanese Encephalitis would involve the accurate and timely reporting of all laboratory-confirmed cases. Accurate and timely surveillance data

promote and inform appropriate and effective response strategies to reduce, or even eliminate, disease incidence. However, this ideal scenario is a costly endeavor and it is often the case that resource-poor countries do not prioritize surveillance efforts for a disease that, although serious and often fatal, portrays a relatively low incidence rate, considering other life-threatening diseases. Many countries have improved surveillance efforts in recent years. For example, Nepal now collects JE case data from 125 sentinel sites, up from 64 sites when their surveillance program was started in 2004 (Shrivastava, 2014). Methods that supplement disease surveillance efforts are a valuable tool for low-resource countries. Modelling disease risk and transmission according to dynamic environmental variables (temperature and rainfall) allow for better predictions of potential outcomes.

5.1 Research Contributions

Previous research has delved into various methods for vector-borne disease risk mapping. For the case of Japanese Encephalitis, and the transmission cycle of its virus, the primary risk mapping technique has been that of ecological niche modelling (Masuoka et al., 2010; Miller et al., 2012). This method for estimating risk, although valuable, can potentially omit important inter- and intra-species behaviour that influences the transmission cycle. Another approach to disease mapping, is the Susceptible-Infected-Removed model and it's many potential variations. These types of models are shown to be most valuable when examining human-to-human transmission. However, SIR models inherently do not account for population heterogeneity, inter-species interaction (transmission between vector, host, and human), and, most importantly when studying the geographic distribution of disease transmission, space.

This work presents two methods that can be used to supplement disease surveillance data by mapping disease risk for Japanese Encephalitis. The first chapter of this thesis details a generalized method for mapping JE risk at a very small scale (large area) using population projections and virus host populations to infer the regions of potentially high risk in the near future. This generalized method may be useful for estimating trends over a large area. This method does not consider individual behaviours, such as exposure to landscape risks that

contribute to one's probability to contracting the JEV.

The second method of this thesis presents an agent-, or individual-based model that takes into account the spatial characteristics of the environment, including climatic and landcover variables, as well as a realistic representation of the human population in the Kathmandu Valley of Nepal. A rule-based Susceptible-Exposed-Infect-Removed (SEIR) model was implemented in this model to capture the incubation period of the virus. Agents transition between SEIR compartment according to a few simple rules, which dictate the overall incidence rate produced by the model. Methodology presented in this thesis was able to reproduce the temporal and spatial dynamics of JE cases reported in the Kathmandu Valley, Nepal.

In the model presented in Chapter 4, an individual's (agent's) likelihood of transitioning to a new SEIR compartment is influenced by their mobility and encounters with dynamic landcover, temperature, and precipitation variables. Thus, the transition parameters are dynamic and vary from individual to individual. The primary research contribution of this paper is the approach used to analyze risk.

Appendices

Appendix A

The Overview, Design Concepts, and Details Protocol

Grimm et al. (2006) devised a standard protocol for developing agent-based models. This protocol is termed 'Overview, Design Concepts, and Details' (ODD) and consists of seven elements that are intended to guide modellers for describing their ABM processes. The purpose of this standard is to organize and present information in a consistent form (Grimm et al., 2006). The protocol also facilitates reimplementation of models and replication of results (Grimm et al., 2006). In their recent textbook titled 'Agent-Based and Individual-Based Modelling: A Practical Introduction', Railsback and Grimm (2012) describe the seven elements of the ODD protocol as follows:

Purpose: The purpose of this element is to clearly define the research question that will be addressed by the ABM. Although this is a seemingly unchallenging step, it identifies what system will be modelled and what knowledge will be gained from the model. This step outlines the *What, Where*, and *Why* portions of a research question.

Entities, State Variables, & Scales: In this element of the ODD protocol, the modeller defines what types of *things* will be represented in the model. The landscape and the agents are the more common types of entities. This section defines the agents in the ABM and what types of state variables these agents possess. There may be multiple types of agents within the

model and each type may possess different state variables that define them. State variables are used to characterize the agents within the model. For example, age, sex, income, etc. are such attributes that an individual agent might have. State variables may be static, such as sex/gender, or they may change through time, such as income and occupation.

The scale of the model is also defined in this element of the ODD protocol. Temporal and spatial scales are important components of an ABM and must be defined by the modeller. ABMs often for a specified period of time, and this must defined in a way that make sense with the research question. In this element of the ODD protocol, the modeller decides how time will be simulated in the model and what each time step represents (i.e. days, weeks, or years) and how long the model will run for.

Process Overview & Scheduling: In this element of the ODD protocol, the dynamics of the ABM are defined. Processes within the model describe how the state variables of the model's agents change through time (excluding static variables, such as sex). Processes are how the modeller defines the *behaviours* of the individual agents and agent types. More specifically, processes of an ABM define what the agents are doing within the model.

The schedule of an ABM is the order in which the processes are executed. By clearly outlining the schedule of the model, a clear description of the model is obtained. Within this element of the ODD protocol, the sequence of actions that the entities undertake are clearly outlined.

Design Concepts: The purpose of this ODD element is to asked key questions on various concepts of the model. A unique quality of ABMs is that they produce multiple outcomes. Therefore, the modeller may want to ask various questions related to these outcomes. Railsback and Grimm (2012) list eleven key concepts of an ABM, including: basic principles, emergence, adaptation, objectives, learning, prediction, sensing, interaction, stochasticity, collectives, observations. Each concept relates to the various agents' processes (behaviours) and the other elements of the ODD protocol. Various questions Questions relating to each of these concepts are explained in more detail in Railsback and Grimm (2012, p. 41).

Initialization: This element of the ODD protocol describes how the modeller will set up the environment of the model at the beginning of the simulation. It also describes the initial values of the entities' state variables. The modeller can define *where* agents are located, what initial attributes the agents posses, and what the initial conditions of the environment are. It is common that the results of ABMs are dependent on the initial conditions of a model.

Input Data: Input data often consists of data on conditions that change over time, but are not necessarily influenced by the agents within the model. This may include data on temperature, rainfall, or government policy. Such information may influence the individual behaviour of an agent and affect the overall outcome of the ABM.

Submodels: All major processes within the ABM are submodels. Submodels are what specifically determine an agent's behaviour. A submodel consists of several if/then statements that depend on the current state of an agent. If the a certain set of conditions are met, the agent returns a specific behaviour.

Appendix B

Interpolation Cross Validation Results

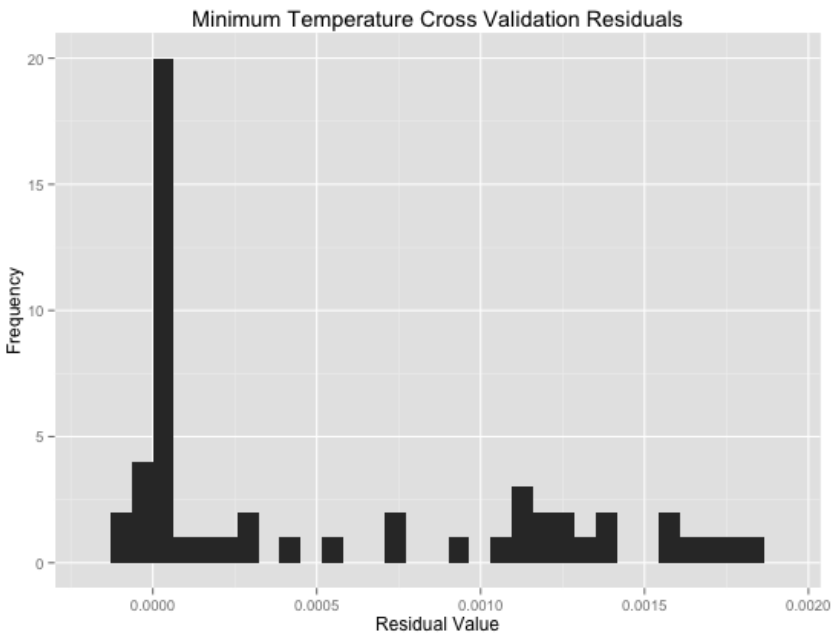


Figure B.1: The mean of the residuals obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated minimum temperatures surfaces.

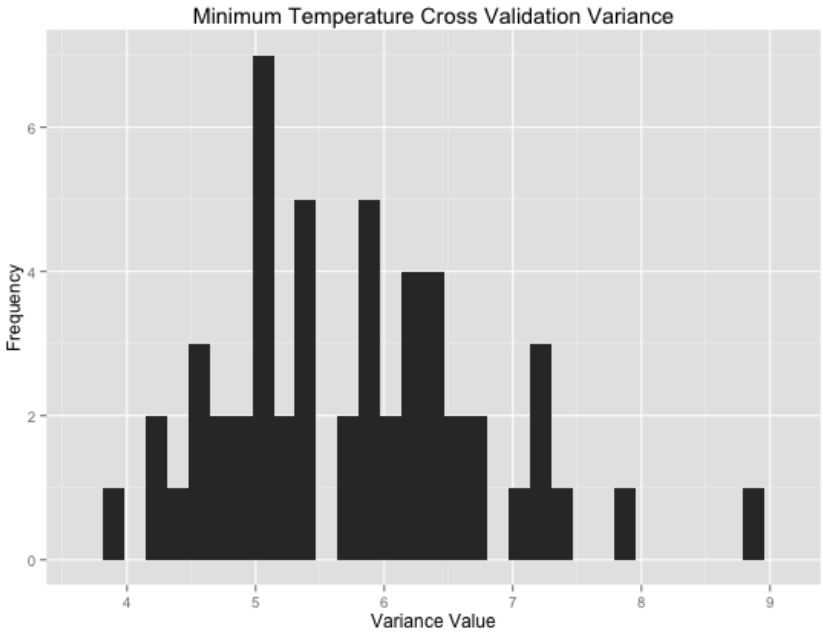


Figure B.2: The mean variances obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated minimum temperature surfaces.

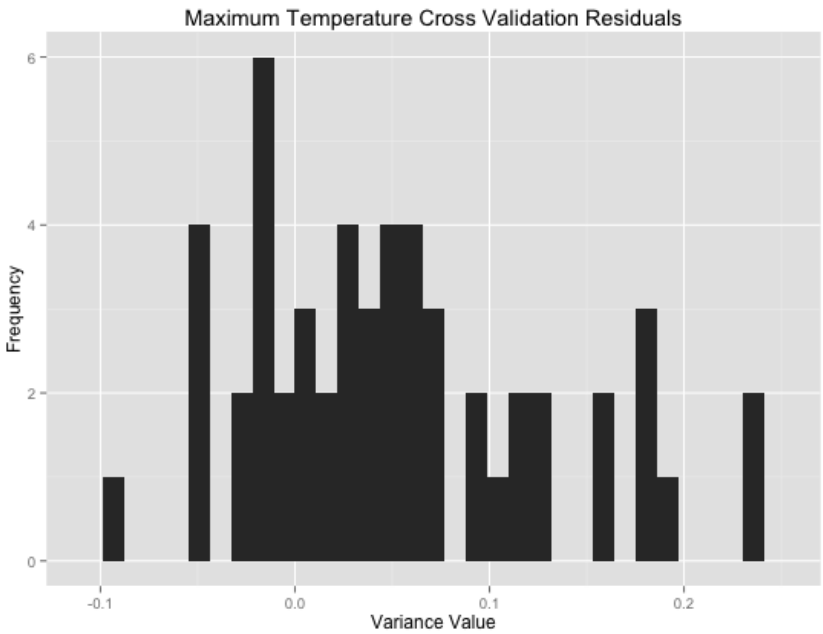


Figure B.3: The mean of the residuals obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated maximum temperatures surfaces.

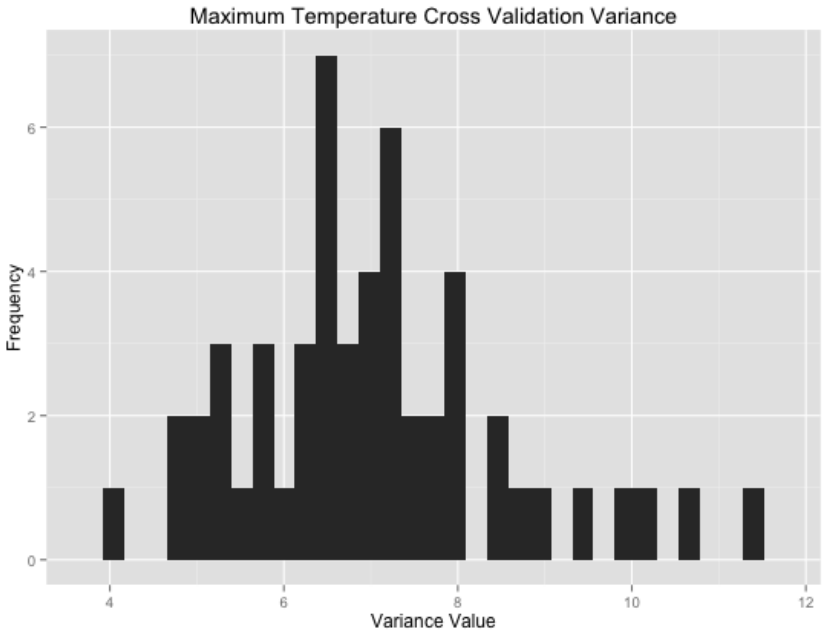


Figure B.4: The mean variances obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated maximum temperature surfaces.

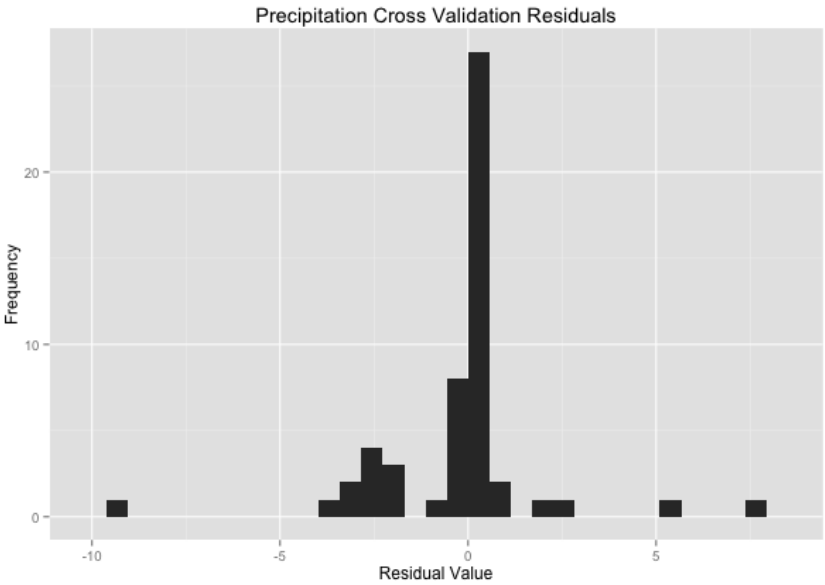


Figure B.5: The mean of the residuals obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated precipitation surfaces.

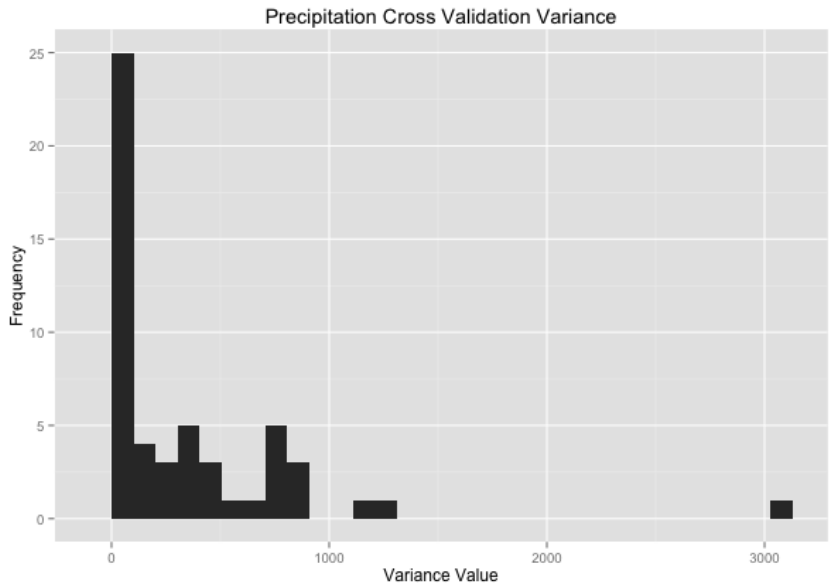


Figure B.6: The mean variances of the residuals obtained by a leave-one-out cross validation between predicted and observed values for weekly interpolated precipitation surfaces.

Bibliography

Akiba, T., K. Osaka, S. Tang, M. Nakayama, A. Yamamoto, I. Kurane, N. Okabe, and T. Umenai (2001). Analysis of japanese encephalitis epidemic in western nepal in 1997. *Epidemiology and infection 126*(01), 81–88.

Alexander, K. A., B. L. Lewis, M. Marathe, S. Eubank, and J. K. Blackburn (2012). Modeling of wildlife-associated zoonoses: applications and caveats. *Vector-Borne and Zoonotic Diseases 12*(12), 1005–1018.

Animal Production and Health: Food and Agriculture Organization of the United Nations (2007). Gridded livestock of the world (glw). http://www.fao.org/ag/againfo/resources/en/glw/home.html. [Online; accessed 25 August 2014].

Arifin, S. N., Y. Zhou, G. J. Davis, J. E. Gentile, G. R. Madey, and F. H. Collins (2014). An agent-based model of the population dynamics of anopheles gambiae. *Malaria journal 13*(1), 1–20.

Balcan, D., V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences 106*(51), 21484–21489.

Beretta, E. and Y. Takeuchi (1995). Global stability of an sir epidemic model with time delays. *Journal of mathematical biology 33*(3), 250–260.

Bhattachan, A., S. Amatya, T. R. Sedai, S. R. Upreti, and J. Partridge (2009). Japanese Encephalitis in Hill and Mountain Districts, Nepal. *Emerging Infectious Diseases 15*(10), 1691.

Bjørnstad, O. N., B. F. Finkenstädt, and B. T. Grenfell (2002). Dynamics of measles epidemics: estimating scaling of transmission rates using a time series sir model. *Ecological Monographs 72*(2), 169–184.

Brown, D. G. and D. T. Robinson (2011). Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl. *Ecology and Society*. [Online; accessed 20 February 2014].

Campbell, G. L., S. L. Hills, M. Fischer, J. A. Jacobson, C. H. Hoke, J. M. Hombach, A. A. Marfin, T. Solomon, T. F. Tsai, V. D. Tsu, et al. (2011). Estimated Global Incidence of Japanese Encephalitis: A Systematic Review. *Bulletin of the World Health Organization 89*(10), 766–774.

CDC (2013). Japanese Encephalitis Surveillance and Immunization – Asia and the Western Pacific, 2012. Morbidity and Mortality Weekly Report. *Centers for Disease Control and Prevention 62*(33), 658–662.

Centers for Disease Control and Prevention (2013). Japanese encephalitis surveillance and immunization–asia and the western pacific, 2012. *MMWR. Morbidity and mortality weekly report 62*(33), 658.

Channan, S., K. Collins, and W. R. Emanuel (2014). Global mosaics of the standard modis land cover type data.

Coker, R. J., B. M. Hunter, J. W. Rudge, M. Liverani, and P. Hanvoravongchai (2011). Emerging infectious diseases in southeast asia: Regional challenges to control. *The Lancet 377*(9765), 599–609.

Conrad, C., S. Fritsch, S. Lex, F. Löw, G. Rücker, G. Schorcht, M. Sultanov, and J. Lamers (2012). Potenziale des fired edgefikanals von rapideye zur unterscheidung und zum monitoring landwirtschaftlicher anbaufrüchte am beispiel des usbekischen bewässerungssystems khorezm. *RapidEye Science Archive (RESA)-Vom Algorithmus zum Produkt, E. Borg, H. Daedelow, and AR Johnson, Eds., Springer, Berlin: GITO*, 201–217.

Dapeng, L., Y. Huijun, Y. Rengou, S. Jindou, and W. Ze (1995). Socio-Economic Status and Micro-Environmental Factors in Relation to the Risk of Japanese Encephalitis: A Case-Control Study. *Southeast Asian journal of tropical medicine and public health 26*(2), 276–279.

Darsie Jr, R., S. Pradhan, et al. (1990). The mosquitoes of nepal: their identification, distribution and biology. *Mosquito Systematics 22*(2), 69–130.

de Almeida, S. J., R. P. M. Ferreira, Á. E. Eiras, R. P. Obermayr, and M. Geier (2010). Multi-agent modeling and simulation of an aedes aegypti mosquito population. *Environmental Modelling & Software 25*(12), 1490–1507.

Deadman, P., D. Robinson, E. Moran, and E. Brondizio (2004). Colonist household decision making and land-use change in the amazon rainforest: an agent-based simulation. *Environment and Planning B 31*, 693–710.

Deguen, S., G. Thomas, and N. P. Chau (2000). Estimation of the contact rate in a seasonal seir model: application to chickenpox incidence in france. *Statistics in medicine 19*(9), 1207–1216.

Department of Economic and Social Affiars: Population Division of the United Nations (2014). The world population situation in 2014. pp. 1–38.

Desai, A., S. Shankar, P. Jayakumar, A. Chandramuki, M. Gourie-Devi, B. Ravikumar, and V. Ravi (1997). Co-existence of cerebral cysticercosis with japanese encephalitis: a prognostic modulator. *Epidemiology and infection 118*(02), 165–171.

Dhakal, S., D. D. Joshi, A. Ale, M. Sharma, M. Dahal, Y. Shah, D. K. Pant, and C. Stephen (2014). Regional Variation in Pig Farmer Awareness and Actions Regarding Japanese Encephalitis in Nepal: Implications for Public Health Education. *PloS one 9*(1), e85399.

Dhakal, S., C. Stephen, A. Ale, and D. Joshi (2012). Knowledge and practices of pig farmers regarding japanese encephalitis in kathmandu, nepal. *Zoonoses and Public Health 59*(8), 568–574.

Dhimal, M., B. Ahrens, and U. Kuch (2015). Climate change and spatiotemporal distributions of vector-borne diseases in nepal–a systematic synthesis of literature. *PloS one 10*(6).

DIVA-GIS (2013, October). Country level data. http://www.diva-gis.org/Data. [Online; accessed 25 October 2013].

Dommar, C. J., R. Lowe, M. Robinson, and X. Rodó (2014). An agent-based model driven by tropical rainfall to understand the spatio-temporal heterogeneity of a chikungunya outbreak. *Acta tropica 129*, 61–73.

Erlanger, T. E., S. Weiss, J. Keiser, J. Utzinger, and K. Wiedenmayer (2009). Past, Present, and Future of Japanese Encephalitis. *Emerging infectious diseases 15*(1), 1.

Esteva, L. and C. Vargas (1999). A model for dengue disease with variable human population. *Journal of Mathematical Biology 38*(3), 220–240.

Eubank, S., H. Guclu, V. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang (2004). Modelling disease outbreaks in realistic urban social networks. *Nature 429*(6988), 180–184.

Food and Agriculture Organization of the United Nations (2014). Production / live animals - number of pigs in nepal. http://faostat3.fao.org/faostat-gateway/go/to/browse/Q/QA/E. [Online; accessed 25 August 2014].

Fraser, C., S. Riley, R. M. Anderson, and N. M. Ferguson (2004). Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences of the United States of America 101*(16), 6146–6151.

Fuks, H. and A. T. Lawniczak (2001). Individual-based lattice model for spatial spread of epidemics. *Discrete Dynamics in Nature and Society 6*(3), 191–200.

GeoHive (2013). GeoHive Historic, Current, and Future Population of Asia. http://www.geohive.com/earth/his_proj_asia.aspx. [Online; accessed 10 October 2013].

Germann, T. C., K. Kadau, I. M. Longini, and C. A. Macken (2006). Mitigation strategies for pandemic influenza in the united states. *Proceedings of the National Academy of Sciences 103*(15), 5935–5940.

Getis, A. and J. K. Ord (1992). The analysis of spatial association by use of distance statistics. *Geographical analysis 24*(3), 189–206.

Gottschalk, M. and M. Segura (2000). The pathogenesis of the meningitis caused by streptococcus suis: the unresolved questions. *Veterinary microbiology 76*(3), 259–272.

Government of Nepal (2014). National Planning Commission Secretariat; Central Bureau of Statistics. http://cbs.gov.np/?p=2021. [Online; accessed 19 January 2014].

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, et al. (2006). A Standard Protocol for Describing Individual-Based and Agent-Based Models. *Ecological modelling 198*(1), 115–126.

Gu, W. and R. J. Novak (2009). Agent-based modelling of mosquito foraging behaviour for malaria control. *Transactions of the Royal Society of Tropical Medicine and Hygiene 103*(11), 1105–1112.

Hailegiorgis, A. and A. T. Crooks (2012). Agent-based modeling for humanitarian issues: Disease and refugee camps.

Hanna, J. N., S. A. Ritchie, D. A. Phillips, J. Shield, M. C. Bailey, J. S. Mackenzie, M. Poidinger, B. J. McCall, and P. J. Mills (1996). An outbreak of japanese encephalitis in the torres strait, australia, 1995. *The Medical Journal of Australia 165*(5), 256–260.

Hart, F. C. (2010). *The Effect of Environmental Variables on Local West Nile Virus Infection Rates in Culex Mosquitoes Using an fiEcological NichefiModel.* Ph. D. thesis, Ohio University.

Hawelka, B., I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti (2014). Geo-located twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science 41*(3), 260–271.

Hecker, K., S. El Kurdi, D. Joshi, and C. Stephen (2013). Using network analysis to explore if professional opinions on japanese encephalitis risk factors in nepal reflect a socio-ecological system perspective. *EcoHealth 10*(4), 415–422.

Ichikawa, D., K. Wakamori, and M. Suzuki (2014). Identification of paddy fields in northern japan using rapideye images. In *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, pp. 2090–2093. IEEE.

Impoinvil, D. E., T. Solomon, W. W. Schluter, A. Rayamajhi, R. P. Bichha, G. Shakya, C. Caminade, and M. Baylis (2011). The spatial heterogeneity between japanese encephalitis incidence distribution and environmental variables in nepal. *PloS one 6*(7), e22192.

International Centre for Integrated Mountain Development (2013). Basic information: Nepal. http://www.icimod.org/?q=1152. [Online; accessed 21 January 2014].

Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and P. Daszak (2008). Global Trends in Emerging Infectious Diseases. *Nature 451*(7181), 990–993.

Joshi, D., P. Bista, and H. Joshi (1995). Japanese encephalitis: a public health problem in nepal (report from 1978–1994).

Joshi, D. and S. Gaidamovich (1982). Serological surveillance of virus encephalitis in nepal. *Bull Vet Soc Anim Hlth, Nepal*, 8–12.

Joshi, D. D. (2006). Pig farming systems in nepal: How can research and development enhance benefits to the poor farmers of nepal. *National Zoonoses and Food Hygiene Research Centre. Kathmandu, Nepal.*

Joshi, D. D. and J. S. Marg (2013). Review on japanese encephalitis outbreak cases in nepal during the year 2011. *INTECH*.

Joshi, D. D. and C. Stephen (2013, August). Year 2 interim technical report. Technical report, National Zoonoses and Food Hygiene Research Centre, G.P.O. Box no: 1885 Chagal, Kathmandu, Nepal.

Kakkar, M., E. T. Rogawski, S. S. Abbas, S. Chaturvedi, T. N. Dhole, S. S. Hossain, and S. K. Krishnan (2013). Acute encephalitis syndrome surveillance, kushinagar district, uttar pradesh, india, 2011–2012. *Emerg Infect Dis 19*(9), 1361–1369.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B: Biological Sciences 266*(1421), 859–867.

Keiser, J., M. F. Maltese, T. E. Erlanger, R. Bos, M. Tanner, B. H. Singer, and J. Utzinger (2005). Effect of Irrigated Rice Agriculture on Japanese Encephalitis, Including Challenges and Opportunities for Integrated Vector Management. *Acta Tropica 95*(1), 40–57.

Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Volume 115, pp. 700–721. The Royal Society.

Konishi, E. and T. Suzuki (2002). Ratios of subclinical to clinical japanese encephalitis (je) virus infections in vaccinated populations: evaluation of an inactivated je vaccine by comparing the ratios with those in unvaccinated populations. *Vaccine 21*(1), 98–107.

Kumari, R., K. Kumar, A. Rawat, G. Singh, N. K. Yadav, and L. S. Chauhan (2013). First indigenous transmission of japanese encephalitis in urban areas of national capital territory of delhi, india. *Tropical Medicine & International Health 18*(6), 743–749.

LaBeaud, A. D., F. Bashir, and C. H. King (2011). Measuring the burden of arboviral diseases: the spectrum of morbidity and mortality from four prevalent infections. *Population health metrics 9*(1), 1.

Lambin, E. F., A. Tran, S. O. Vanwambeke, C. Linard, and V. Soti (2010). Pathogenic Landscapes: Interactions Between Land, People, Disease Vectors, and Their Animal Hosts. *International journal of health geographics 9*(54), 1–13.

Le Flohic, G., V. Porphyre, P. Barbazan, and J.-P. Gonzalez (2013). Review of Climate, Landscape, and Viral Genetics as Drivers of the Japanese Encephalitis Virus Ecology. *PLoS neglected tropical diseases 7*(9), e2208.

Lekone, P. E. and B. F. Finkenstädt (2006). Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study. *Biometrics 62*(4), 1170–1177.

Li, S., N. Hartemink, N. Speybroeck, and S. O. Vanwambeke (2012). Consequences of landscape fragmentation on lyme disease risk: a cellular automata approach. *PloS one 7*(6), e39612.

Linard, C., N. Ponçon, D. Fontenille, and E. F. Lambin (2009). A Multi-Agent Simulation to Assess the Risk of Malaria Re-Emergence in Southern France. *Ecological Modelling 220*(2), 160–174.

Macal, C. M. and M. J. North (2008). Agent-Based Modeling and Simulation: ABMS Examples. In *Proceedings of the 40th Conference on Winter Simulation*, pp. 101–112. Winter Simulation Conference.

Mackenzie, J. S., D. J. Gubler, and L. R. Petersen (2004). Emerging Flaviviruses: The Spread and Resurgence of Japanese Encephalitis, West Nile and Dengue Viruses. *Nature Medicine 10*, S98–S109.

Maltsoglou, I. and K. Taniguchi (2004). Poverty, livestock and household typologies in nepal. *Pro-Poor Livestock Policy Initiative (PPLP). Rome: Food and Agriculture Organization of the United Nation.*

Masuoka, P., T. A. Klein, H.-C. Kim, D. M. Claborn, N. Achee, R. Andre, J. Chamberlin, J. Small, A. Anyamba, D.-K. Lee, et al. (2010). Modeling the distribution of culex tritaeniorhynchus to predict japanese encephalitis distribution in the republic of korea. *Geospatial health 5*(1), 45–57.

Miller, R. H., P. Masuoka, T. A. Klein, H.-C. Kim, T. Somer, and J. Grieco (2012). Ecological Niche Modeling to Estimate the Distribution of Japanese Encephalitis Virus in Asia. *PLoS Neglected Tropical Diseases 6*(6), e1678.

Mills, C. E., J. M. Robins, and M. Lipsitch (2004). Transmissibility of 1918 pandemic influenza. *Nature 432*(7019), 904–906.

Misra, U. K. and J. Kalita (2010). Overview: Japanese Encephalitis. *Progress in neurobiology 91*(2), 108–120.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 17–23.

Mountain GeoPortal (2012). Kathmandu Valley GIS Datasets. http://geoportal.icimod.org/downloads/Download.aspx?ID=18. [Online; accessed 18 January 2014].

Murray-Rust, D., D. T. Robinson, E. Guillem, E. Karali, and M. Rounsevell (2014). An open framework for agent based modelling of agricultural land use change. *Environmental Modelling & Software 61*, 19–38.

Murty, U. S., M. S. Rao, N. Arunachalam, et al. (2010). The effects of climatic factors on the distribution and abundance of japanese encephalitis vectors in kurnool district of andhra pradesh, india. *J Vector Borne Dis 47*(1), 26–32.

Nemeth, N. M., A. M. Bosco-Lauth, R. H. Sciulli, R. B. Gose, M. T. Nagata, and R. A. Bowen (2010). Serosurveillance for Japanese Encephalitis and West Nile Viruses in Resident Birds in Hawai'i. *Journal of Wildlife Diseases 46*(2), 659–664.

Noulas, A., S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo (2012). A tale of many cities: universal patterns in human urban mobility. *PloS one 7*(5), e37027.

Pant, G. R. (2006). A serological survey of pigs, horses, and ducks in nepal for evidence of infection with japanese encephalitis virus. *Annals of the New York Academy of Sciences 1081*(1), 124–129.

Partridge, J., P. Ghimire, T. Sedai, M. B. Bista, and M. Banerjee (2007). Endemic Japanese Encephalitis in the Kathmandu Valley, Nepal. *The American Journal of Tropical Medicine and Hygiene 77*(6), 1146–1149.

Pesavento, U. (1995). An implementation of von neumann's self-reproducing machine. *Artificial Life 2*(4), 337–354.

Peterson, A. T. (2001). Predicting species'geographic distributions based on ecological niche modeling. *The Condor 103*(3), 599–605.

Peterson, A. T. (2015). Mapping disease transmission risk: Enriching models using biogeography and ecology. *Emerging infectious diseases 21*(8), 1489.

Railsback, S. F. and V. Grimm (2012). *Agent-Based and Individual-Based Modeling: A Practical Introduction.* Princeton University Press.

Rayamajhi, A., R. Singh, R. Prasad, B. Khanal, and S. Singhi (2007). Study of japanese encephalitis and other viral encephalitis in nepali children. *Pediatrics International 49*(6), 978–984.

Recio, J., P. Helmholz, and S. Müllerb (2011). Potential evaluation of different types of images and their combination for the classification of gis objects cropland and grassland. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38*, 4.

Reeves, W., B. Brookman, and W. M. Hammon (1948). Studies on the flight range of certain culex mosquitoes, using a fluorescent-dye marker, with notes on culiseta and anopheles. *Mosq News 8*, 61–69.

Rindfuss, R. R., B. Entwisle, S. J. Walsh, L. An, N. Badenoch, D. G. Brown, P. Deadman, T. P. Evans, J. Fox, J. Geoghegan, et al. (2008). Land use change: complexity and comparisons. *Journal of Land Use Science 3*(1), 1–10.

Robertson, C., D. K. Pant, D. D. Joshi, M. Sharma, M. Dahal, and C. Stephen (2013). Comparative Spatial Dynamics of Japanese Encephalitis and Acute Encephalitis Syndrome in Nepal. *PloS one 8*(7), e66168.

Robinson, D., D. Murray-Rust, V. Rieser, V. Milicic, and M. Rounsevell (2012). Modelling the impacts of land system dynamics on human well-being: Using an agent-based approach to cope with data limitations in koper, slovenia. *Computers, Environment and Urban Systems 36*(2), 164–176.

Rodriguez-Meza, M. A. (2012). Spatial and temporal dynamics of infected populations: the mexican epidemic. *arXiv preprint arXiv:1204.2761.*

Rozendaal, J. (1997). Mosquitoes and other biting diptera, in vector control. *World Health Organization, Geneva, Switzerland*, 5–177.

Sallam, M. F., A. M. Al Ahmed, M. S. Abdel-Dayem, and M. A. Abdullah (2013). Ecological

niche modeling and land cover risk areas for rift valley fever vector, culex tritaeniorhynchus giles in jazan, saudi arabia.

Shrivastava, A. (2014). Acute encephalitis syndrome surveillance: Challenges and the way forward. *Journal of Communicable Diseases 46*(1), 73–85.

Sibley, D. A. (2000). *The Sibley Guide to Birds.* Knopf, New York.

Singh, A., S. K. Saxena, A. K. Srivastava, and A. Mathur (2012). Japanese encephalitis: A persistent threat. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences 82*(1), 55–68.

Sohn, Y. M. (2000). Japanese encephalitis immunization in south korea: past, present, and future. *Emerging infectious diseases 6*(1), 17.

Solomon, T. (2004). Flavivirus encephalitis. *New England Journal of Medicine 351*(4), 370–378.

Solomon, T., H. Ni, D. W. Beasley, M. Ekkelenkamp, M. J. Cardosa, and A. D. Barrett (2003). Origin and Evolution of Japanese Encephalitis Virus in Southeast Asia. *Journal of Virology 77*(5), 3091–3098.

Solomon, T. and P. Winter (2004). *Neurovirulence and host factors in flavivirus encephalitis evidence from clinical epidemiology.* Springer.

Tacoli, C. (1998). Rural-urban interactions; a guide to the literature. *Environment and Urbanization 10*, 147–166.

Thapa, R. B. and Y. Murayama (2008). Spatial structure of land use dynamics in kathmandu valley. *37*, 11–16.

Thapa, R. B. and Y. Murayama (2009). Examining spatiotemporal urbanization patterns in kathmandu valley, nepal: Remote sensing and spatial metrics approaches. *Remote Sensing 1*(3), 534–556.

Thapa, R. B. and Y. Murayama (2010). Drivers of urban growth in the kathmandu valley, nepal: examining the efficacy of the analytic hierarchy process. *Applied Geography 30*(1), 70–83.

The World Bank (2013a). Agriculture & rural development: Rural population (% of total population). http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS. [Online; accessed 01 November 2013].

The World Bank (2013b, March). The spatial transition.

The World Health Organization (2008). *WHO - Recommended Standards for Surveillance of Selected Vaccine-Preventable Diseases.* Department of Vaccines and Biologicals, Health Technology and Pharmaceuticals, World Health Organization.

Thorpe, W. and T. Jemaneh (2008). *Pig systems in Asia and the Pacific: how can research and development enhance benefits to the poor.* ILRI (aka ILCA and ILRAD).

Turtle, L., M. Griffiths, and T. Solomon (2012). Encephalitis caused by flaviviruses. *Qjm 105*(3), 219–223.

Ulam, S. (1950). Statistical mechanics of cellular automata, 1952. In *Proceedings of the International Congress on Mathematics*, Volume 2, pp. 264–275.

UNdata (2015). Nepal: Country Profile. http://data.un.org/CountryProfile.aspx?crName=Nepal#Social. [Online; Accessed 25 Auguest 2015].

United Nations (2002). *World Urbanization Prospects: The 2001 Revision*, Volume 216. United Nations Publications.

Van den Hurk, A. F., S. A. Ritchie, and J. S. Mackenzie (2009). Ecology and geographical expansion of japanese encephalitis virus. *Annual review of entomology 54*, 17–35.

Verdonschot, P. F. and A. A. Besse-Lototskaya (2014). Flight distance of mosquitoes (culicidae): a metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologica-Ecology and Management of Inland Waters 45*, 69–79.

Wang, L., W. Hu, R. J. S. Magalhaes, P. Bi, F. Ding, H. Sun, S. Li, W. Yin, L. Wei, Q. Liu, et al. (2014). The role of environmental factors in the spatial distribution of japanese encephalitis in mainland china. *Environment international 73*, 1–9.

Weaver, S. C. and A. D. Barrett (2004). Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature Reviews Microbiology 2*(10), 789–801.

Weeks, J. R. (2010). Defining urban areas. In *Remote Sensing of Urban and Suburban Areas*, pp. 33–45. Springer.

Whitley, R. J. (1990). Viral encephalitis. *New England Journal of Medicine 323*(4), 242–250.

Wilensky, U. (1999). Netlogo. *Center for Connected Learning and Computer-Based Modeling*. [Online; accessed 22 February 2014].

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature 311*(5985), 419–424.

World Health Organization (2014a). Immunization, vaccines, and biologicals: National passive surveillance. http://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/passive/en. [Online; accessed 13 October 2014].

World Health Organization (2014b). Japanese encephalitis: Media centre. [Online: Accessed 17 November 2014].

World Health Organization (2014c). Media centre: Japanese encephalitis. http://http://www.who.int/mediacentre/factsheets/fs386/en/. [Online; accessed 13 May 2015].

World Health Organization (2015a). Global health observatory (gho) data. http://www.who.int/gho/malaria/en/. [Online; Accessed 16 December 2015].

World Health Organization (2015b). Water sanitation health: Water-related diseases. http://www.who.int/water_sanitation_health/diseases/encephalitis/en/. [Online; Accessed 05 July 2015].

World Health Organization (2015c). Zoonoses; Diseases. http://www.who.int/zoonoses/diseases/en. [Online; Accessed 23 October 2015].

World Health Organization: Country Office for Nepal (2011). Vaccine Preventable Diseases Surveillance. http://www.nep.searo.who.int/EN/Section4/Section29/Section88.htm. [Online; Accessed 23 March 2014].

Wu, J., Z. Gao, and H. Sun (2004). Simulation of traffic congestion with sir model. *Modern Physics Letters B 18*(30), 1537–1542.

Wu, Y.-C., Y.-S. Huang, L.-J. Chien, T.-L. Lin, Y.-Y. Yueh, W.-L. Tseng, K.-J. Chang, and G.-R. Wang (1999). The epidemiology of japanese encephalitis on taiwan during 1966-1997. *The American journal of tropical medicine and hygiene 61*(1), 78–84.

Yeasmin, K. and R. Hyndman (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software 27*.

Yu, J., D. Jiang, and N. Shi (2009). Global stability of two-group sir model with random perturbation. *Journal of Mathematical Analysis and Applications 360*(1), 235–244.

Zhao, L., H. Cui, X. Qiu, X. Wang, and J. Wang (2013). Sir rumor spreading model in the new media age. *Physica A: Statistical Mechanics and its Applications 392*(4), 995–1003.