# EFFECTS OF VECTOR MIGRATION ON SYLVATIC *TRYPANOSOMA CRUZI* TRANSMISSION

by

BRITNEE A CRAWFORD

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2012

I dedicate this to my son, Easton, with the goal that he be inspired to believe that anything is possible with a dream, belief in that dream, and the dedication to pursue it.

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Christopher Kribs Zaleta for his constant support and encouragement. He has spent countless hours teaching, advising, and mentoring me, and for that I am truly grateful. He has taught me many things over the years, and I truly believe I have been blessed to know him. I would also like to thank my graduate committee members Dr. Gaik Ambartsoumian, Dr. James Grover, Dr. Hristo Kojouharov, and Dr. Anuj Mubayi for their interest in my research and for their suggestions and advice throughout this project.

Finally, I would like to thank my family. I could not have finished this work without their constant support and help. I am so blessed to have a loving family who is willing to sacrifice their time to help me pursue my goals. Words cannot express what it means to me to have so much love and support surrounding me daily. I am so grateful for my mother and her constant support and sacrifice since I was young. She has always made me a top priority, and I only hope that I can be as good a mother to my children as she has been to me. I am extremely grateful for my husband who has been steadfast and selfless as I pursued my studies. You are my partner, and best friend. I also wish to thank my aunt who has always been there for me, but most importantly has sacrificed her time every week this past year to help take care of our son while I studied. For that I am truly grateful. There are countless other people who have been integral parts of my life, and been there for me to support and encourage.

This project was partially supported by a 2008 Norman Hackerman Advanced Research Program grant, and by the National Science Foundation under Grant DMS-1020880.

"I give you thanks, Creator and God, that you have given me this joy in thy creation, and I rejoice in the works of your hands."–Johannes Kepler

April 18, 2012

ABSTRACT

EFFECTS OF VECTOR MIGRATION ON SYLVATIC
*TRYPANOSOMA CRUZI* TRANSMISSION

BRITNEE A CRAWFORD, Ph.D.

The University of Texas at Arlington, 2012

Supervising Professor: Christopher Kribs-Zaleta

Vector-borne diseases have had a major impact on global health concerns since their discovery in the 1800's. A vector-borne disease is one transmitted to human or animal host via an invertebrate vector (usually an insect). Chagas' disease, caused by the parasite *Trypanosoma cruzi*, is transmitted via insect vectors from the *Triatoma* family. Although human infection with Chagas' is of importance, the disease is maintained in sylvatic (in the wild) transmission cycles. This study examines the effects of vector migration on the spread of *T. cruzi* in certain sylvatic cycles in the southeastern U.S. and Mexico from several angles. First, a framework for deriving estimates for migration rates is developed. This framework translates local, small-scale vector dispersal information (extracted from biological studies) into global, large-scale migration rates, which can then be used in metapopulation "patch" models or spatially explicit (discrete in space) models. To apply the migration rate derivation framework, a metapopulation compartmental model describing the transmission cycles is presented as a system of ordinary differential equations coupled via migration of the vectors. Standard threshold analysis techniques provide insight into the role of

migration through the basic reproductive number, $R_0$, while numerical investigations illustrate migration's effects on prevalence of *T. cruzi*. Finally, in order to gain a fuller understanding of the spread of *T. cruzi* in these regions, a spatially explicit model (cellular automaton) is developed with the main goal of measuring invasion speed of the epidemic. Effects on invasion speed and direction of invasion are examined based on certain vector migration characteristics, such as preference for dispersal in a particular direction. When compared with no preference for direction, results show the increase in speed is greater if the preferred direction is more in alignment with the natural geography of the region than the decrease in speed if the preferred direction is not aligned with the natural geography.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS . . . . . iv

ABSTRACT . . . . . vi

LIST OF ILLUSTRATIONS . . . . . xi

LIST OF TABLES . . . . . xiv

Chapter Page

1. INTRODUCTION . . . . . 1

1.1 Mathematical modeling . . . . . 4

1.2 Problem statement . . . . . 6

2. VECTOR MIGRATION AND DISPERSAL RATE FOR SYLVATIC *T. CRUZI* TRANSMISSION . . . . . 8

2.1 Introduction . . . . . 8

2.2 Patches . . . . . 12

2.2.1 Patch 1 . . . . . 13

2.2.2 Patch 2 . . . . . 15

2.2.3 Patch 3 . . . . . 17

2.3 Dispersal and migration rates . . . . . 18

2.3.1 Dispersal rates . . . . . 19

2.4 Preferred direction . . . . . 20

2.4.1 Local migration framework . . . . . 22

2.4.2 Global migration framework . . . . . 26

2.5 No preferred direction . . . . . 28

2.5.1 Estimation of dispersal kernel . . . . . 29

2.5.2 Calculation of migration rates for no preferred direction . . . . 32

2.6 Numerical calculations for preferred direction . . . . . 35

2.7 Conclusions . . . . . 38

3. A METAPOPULATION MODEL FOR SYLVATIC *T. CRUZI* TRANSMISSION WITH VECTOR MIGRATION . . . . . 41

3.1 Introduction . . . . . 41

3.2 Problem formulation and model . . . . . 44

3.2.1 Problem statement and assumptions . . . . . 44

3.2.2 The Model . . . . . 47

3.3 Analysis . . . . . 51

3.3.1 One patch, one host, one vector, no vertical transmission . . . 51

3.3.2 Patch 2, no migration, no vertical transmission . . . . . 53

3.3.3 Patches 1 and 2, 1 host 1 vector with vertical transmission and unidirectional migration of infected vectors . . . . . 55

3.3.4 Patches 1 and 2, 1 host 1 vector with vertical transmission and unidirectional migration of all vectors . . . . . 56

3.3.5 Patches 1 and 2, 1 host 1 vector, no vertical transmission, bidirectional migration of infected vectors . . . . . 58

3.3.6 Patches 1 and 2, 2 hosts 2 vectors, no vertical transmission, unidirectional migration of infected vectors . . . 60

3.3.7 Patch 1, 2, and 3, 2 hosts 2 vectors, no vertical transmission, unidirectional migration of infected vectors . . . 61

3.3.8 Synthesis . . . . . 63

3.4 Numerical Results . . . . . 64

3.4.1 General demographic parameters . . . . . 64

3.4.2 Estimation of infection rate parameters . . . . . 71

3.4.3 Numerical solutions . . . . . 75

3.5 Conclusions . . . . . 83

4. INVASION SPEED IN CELLULAR AUTOMATON MODELS FOR *T. CRUZI* VECTOR MIGRATION . . . . . 87

4.1 Introduction . . . 87
4.2 Literature Review . . . 89
4.3 Methods . . . 93
4.4 Migration rates for cellular automata . . . 97
4.5 Analysis . . . 100
4.5.1 Determining threshold presence of epidemic . . . 100
4.5.2 Invasion speed . . . 103
4.6 Results . . . 107
4.6.1 Basic trends in speed and direction . . . 108
4.6.2 Role of preferred direction . . . 114
4.7 Conclusions . . . 118
5. CONCLUSION . . . 122
Appendix
A. ERROR CORRECTIONS FOR MIGRATION RATE CALCLATION . . . 129
B. R0 AND ENDEMIC PREVALENCE CALCULATIONS . . . 147
C. MODEL EQUATIONS AND PARAMETERS . . . 160
REFERENCES . . . 164
BIOGRAPHICAL STATEMENT . . . 175

LIST OF ILLUSTRATIONS


| Figure | | Page |
|---|---|---|
| 2.1 | Patches | 13 |
| 2.2 | Interaction between global and local coordinate systems | 20 |
| 2.3 | Varying direction with bold points of origin ($e = 0.7$) | 21 |
| 2.4 | Varying eccentricity with bold points of origin ($\theta_0 = -\pi/4$) | 22 |
| 2.5 | Preferred direction framework. Shaded region represents the area for which dispersals have crossed patch boundary originating from a distance $d$ away from the patch boundary. | 24 |
| 2.6 | Global coordinate system | 27 |
| 2.7 | Patch boundaries linear segments | 34 |
| 2.8 | Migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line) ($e = 0.5$) | 37 |
| 2.9 | Migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line)($e = 0.1$) | 37 |
| 3.1 | Model. The migration rates represent outgoing rates which must be adjusted by the patch area ratios for incoming rates | 47 |
| 3.2 | Correspondence of $q_S$ and $q_W$ with $\beta_{WS}$, with units 1/yr, and $\beta_{SW}$, with units infected hosts/vector/yr | 73 |
| 3.3 | Infected vectors migrating north only; dark curve represents vectors, light curve represents hosts | 76 |
| 3.4 | Infected vectors migrating north only | 76 |
| 3.5 | All vectors migrating north only; dark curve represents vectors, light curve represents hosts | 77 |
| 3.6 | Infected vectors moving south only; dark curve represents vectors, light curve represents hosts | 78 |
| 3.7 | Infected vectors migrating south only | 78 |

3.8 All vectors migrating south only; dark curve represents vectors, light curve represents hosts . . . . . 78

3.9 Infected vectors bidirectional migration; dark curve represents vectors, light curve represents hosts . . . . . 79

3.10 Infected vectors bidirectional migration; dark curve represents vectors, light curve represents hosts . . . . . 80

3.11 All vectors bidirectional migration; dark curve represents hosts, light curve represents vectors . . . . . 80

4.1 Grid framework with cells 26.5 km by 26.5 km . . . . . 95

4.2 Sample regions of integration for a generic vector dispersal ellipse with preferred direction northwest. . . . . 98

4.3 Sample cell showing corner angles and equations of boundary segments for a given point of dispersal, $(x, y)$. . . . . 98

4.4 Northward migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line) ($e = 0.5$) . . . . . 100

4.5 Host infection prevalence in a single (sample) cell, expressed as a proportion of final endemic prevalence, with threshold coordinates based on superimposed 3-piecewise linear regression . . . 102

4.6 Speed vs. time with results by patch (top) and species (bottom)) . . . 110

4.7 Direction vs. time for component method. Direction is measured in degree measures counterclockwise from due east with 90° representing north . . . . . 111

4.8 Direction field for no preference of direction ($e = 0$) (shading indicates magnitude) . . . . . 112

4.9 Power law regression fit for speed vs. migration rate scaling factor $k$ with $0 < k \leq 15$. The equation for the fit is $6.17k^{0.44}$. . . . . 113

4.10 Speed vs. time (comparison of methods) . . . . . 114

4.11 Direction vs. time (both methods represented) with direction represented as degree measured counterclockwise from due east, with 90° representing north . . . . . 115

4.12 Direction field plot for northeast preferred direction of migration with $e = 0.5$. . . . . 117

4.13 Bar graph of average invasion speed vs. degree of preference (eccentricity) with preferred direction northeast . . . . . 118

4.14 A given front for northeastern preferred direction of migration. . . . . 119

5.1 Comparison of prevalence vs. time by patch for 3-cell model in Chapter 2 . . . . . 126

5.2 Comparison of prevalence vs. time by patch for cellular automaton model in Chapter 3 . . . . . 127

5.3 Comparison of prevalence vs. time results in Chapter 2 and 3. Results are presented by patch. . . . . 127

LIST OF TABLES


| Table | | Page |
|---|---|---|
| 2.1 | Dispersal data from [74, 75] for a single flight | 30 |
| 2.2 | Averages of proportions of vectors dispersing within 0-100m for a single flight | 30 |
| 2.3 | Estimated 35-day vector dispersal distribution, extrapolated from single-flight data in Tables 2.1 and 2.2 | 33 |
| 2.4 | Geographical parameter values | 33 |
| 2.5 | $\theta_j$ values for outward normals for patch boundary segments | 36 |
| 2.6 | Migration rates for $e = 0.5$, Patches 2/3 | 38 |
| 2.7 | Migration rates for $e = 0.5$, Patches 1/2 | 38 |
| 3.1 | Demographic parameters | 65 |
| 3.2 | Patch 1 density estimates | 66 |
| 3.3 | *Neotoma micropus* population density estimates | 68 |
| 3.4 | Patch 2 density estimates | 69 |
| 3.5 | Patch 3 density estimates | 71 |
| 3.6 | *T. cruzi* prevalence estimates from [44] | 72 |
| 3.7 | Stercorarian infection rate parameters | 74 |
| 3.8 | Migration rates for no preferred direction (units in 1/year) | 81 |
| 3.9 | Adjusted migration rates for no preferred direction (units in 1/year) | 81 |
| 3.10 | Equilibrium prevalence levels for species based on migration (northward preference for direction) | 82 |
| 4.1 | Specific rates for varying preferred directions, $e = 0.5$, rates in units 1/yr | 100 |
| 4.2 | Threshold times | 104 |

4.3 Eastward velocities (cell diameters/$yr$ . . . . . 104

4.4 Northward velocities (cell diameters/$yr$) . . . . . 104

4.5 Statistical measures for speed (no preferred direction of migration) using cellwise threshold prevalence, units km/yr . . . . . 109

4.6 Statistical measures for speed (no preferred direction of migration) using average threshold prevalence of 0.07, units km/yr . . . . . 109

4.7 Statistical measures for speed (no preferred direction) using front method, units km/yr . . . . . 109

4.8 Average speeds (in units km/yr) and directions (in degree measures north of due east) by patch and overall. When a preferred direction is given, $e = 0.5$. . . . . 116

4.9 Average speed (units km/yr) and direction (degrees north of due east) overall and by patch for varying levels of eccentricity. . . . . 119

5.1 Epidemic growth phases for Chapter 2 and Chapter 3 models. Units are in years. . . . . 126

# CHAPTER 1

# INTRODUCTION

Among the types of infectious diseases, vector-borne diseases are at the top of the list of global health concern. Vector-borne diseases are transmitted to hosts (human or animal) via insects (or arachnids), known as the "vector". Some of the most commonly known vector-borne pathogens are transmitted by mosquitoes (e.g., West Nile virus and malaria), flies (e.g. leishmaniasis), and ticks (e.g. Lyme disease). A vector-borne disease of major concern in the Americas, transmitted via insect vectors from the subfamily Triatominae, is Chagas' disease. Chagas' disease, discovered in 1909, is widespread in Mexico, Central America, and South America. An estimated 8 to 11 million people are currently infected, with many who are unaware of their infection. [92]

*Trypanosoma cruzi* is a protozoan parasite primarily known for causing Chagas' disease. The parasite is contained in the feces of the insect vectors, and primarily transmitted through the bite wound (or mucous membranes) of the host. Upon entering through the skin cells of the host, the parasite transforms to amastigotes and multiplies, transforming into blood trypomastigotes, entering the blood stream of the host. In humans, infection with the parasite is lifelong. Once infected with the parasite, a person will enter the acute phase of the infection, which may be asymptomatic. If symptoms are present, the patient may display fever, nausea, vomiting, and/or swelling around the site of the bite. After the acute phase, the parasites usually travel to the tissue of the cardiac and gastrointestinal systems where they usually go undetected. If left untreated, a person can develop serious complications involving

the cardiac and gastrointestinal systems (characteristics of chronic Chagas' disease) [15].

There are many species which may be considered a vector for Chagas' disease, all belonging to the family of Reduviidae insects. This family includes over 7000 species, with approximately 130 species belonging to the subfamily, Triatominae. Members of this subfamily may also be referred to as kissing bugs or conenose bugs, each name an identifying characteristic of the primary vectors of Chagas' disease (from the genera *Triatoma*). *Triatoma* vectors are easily identified by their "cone-shaped" nose and flattened wide body (usually with red or dark brown markings on the edges of the wings). *Triatomas* are "heat-seeking" insects [33], often times known to bite their human hosts near the mouth (during the night while sleeping), hence the name kissing bugs.

Since its discovery, Chagas' disease research has been primarily focused on human infection and transmission. Control measures have been implemented in certain parts of South America. The Southern Cone Initiative, initiated in 1991, was designed to interrupt domestic transmission via eliminating the main vector species, *Triatoma infestans*, in certain South American countries. Uruguay, Chile, and Brazil are now certified free of *T. cruzi* transmission via *T. infestans*. The Southern Cone Initiative has induced awareness of *T. cruzi* infection in other countries, such as Mexico, where the disease remains endemic. However, not near enough attention has been given since it has been estimated that approximately 1.8 million people in Mexico are infected, with much less than that being reported to the ministry of health (approximately 3,500 cases from 2000-2010) [8]. At present, direct transmission via the *Triatoma* vector to humans is not of major concern to the United States. Fewer than 10 cases of autochthonous transmission have been reported [7]. Although more at-

tention is being given to Chagas', incidents of the disease still remain underreported, and Chagas' is classified as a neglected parasitic infection in the United States [15].

Although there have been few human cases in the U.S., the disease remains endemic in sylvatic cycles throughout Mexico and the United States. Sylvatic transmission cycles are vector-host cycles that occur naturally in the wild. In the United States, triatomine vectors are found in 26 states [46], involving approximately 11 triatomine vector species (with 8 of the 11 in Texas) and over 100 mammalian species. In the United States, some of the most common sylvatic hosts include opossums (*Didelphis virginiana*) and raccoons (*Procyon lotor*) in the southeastern parts of the country and woodrats (*Neotoma micropus*) in Texas (extending also into northern parts of Mexico). Other species, such as dogs, armadillos, skunks, and chickens have also been noted as relevant species in sylvatic settings (with canines part of some domestic and peridomestic cycles [13, 94]). Of the 11 vector species and mammalian species listed here, we identify two primary vector species in the southeastern U.S., *Triatoma sanguisuga*, found all along the southeastern Atlantic coast from Florida into central Texas, and *Triatoma gerstaeckeri*, found mostly from central Texas south into states in northern Mexico [41]. In addition to the complex vector-host cycles, there are also multiple strains of *T. cruzi* circulating in these populations. There are 6 known strain types of *T. cruzi*, types I-VI, of which types I and IV are circulating in the United States. There are distinct differences between the strains, from host specificity to virulence. *T. cruzi* I, associated with Chagas' disease, is the primary strain circulating in Mexico (also found in hosts in the U.S.), while type IV is almost exclusively found in the United States [69].

Infection with *T. cruzi* can be transmitted in these sylvatic settings in several different ways. Once a vector takes a bloodmeal from an infected host, the parasite enters the insect's gut. Then, the most common route of infection to the host is

stercorarian transmission, in which an infected vector defecates in or near the bite wound and the parasite enters the host through the broken skin (or possibly mucous membranes). Other types of transmission, only recently being researched, involve hosts becoming infected through predation on infected hosts (oral transmission) and transplacental transmission from mother to baby (vertical transmission). Although these types of transmission may be classified as alternative transmission modes, it has been shown that in some cycles, these modes are in fact enough to sustain sylvatic transmission with low stercorarian transmission rates [43].

## 1.1 Mathematical modeling

The field of mathematical biology has a broad spectrum of applications. Population dynamics is one of the most well-established fields of mathematical biology. There are many types of models that can be used to describe population dynamics, all of which fall under one of two categories: deterministic models or stochastic models. In deterministic models, the future state of the system is uniquely determined by the values of the parameters and initial states of the variables in the system. On the other hand, stochastic models operate under the assumption that randomness is present, and thus the solutions to the system are not uniquely determined, but rather represented by probability distributions. The main elements used in modeling population dynamics are time and space. Models may incorporate time and space in many ways. For example, deterministic systems may be continuous in time (ordinary differential equations), continuous in time and space (partial differential equations), or discrete in time (difference equations). Each of these systems operates under basic assumptions regarding the biological process being modeled.

Specific models used to describe population growth begin with the exponential model presented by Malthus (1798) and the logistic model developed by Verhulst

(1838). Each of the models alone may not be adequate to describe more complicated demographic processes, but has been an invaluable step in the evolution of mathematical modeling of population dynamics. Overlapping with the field of population dynamics is the field of mathematical epidemiology. The field of mathematical epidemiology has been progressing steadily since the 18th century with Daniel Bernoulli's modeling of smallpox. In his work, he set out to show that inoculation would reduce the death rate due to smallpox and increase the population of France. Most of the main developments in mathematical epidemiology were made in the early 20th century. Ross published his work on malaria (1911) in which he showed that malaria could be controlled by reducing the mosquito population in a certain region. This work is particularly useful in the sense that it was one of the first mathematical epidemic models of a vector-borne disease and one of the first to address the idea of threshold behavior. Shortly after Ross' studies on malaria, Kermack and McKendrick developed a deterministic epidemic model (1927) in which the population under study is divided into distinct groups or compartments, referred to as a compartmental model. The population is divided into each compartment depending on their status with respect to the infection. The basic structure (for the Kermack and McKendrick model) is an SIR model in which individuals of the population are classified as either susceptible, infectious, or recovered and transition through the compartments based on assumed characteristics of the infection process. There could be many other model structures, such as SI, in which members are either susceptible or permanently infected, such as is the case with *T. cruzi*, or an SEIR model, in which the "E" represents the class of individuals exposed to the infection, but not yet infectious.

In a compartmental model, results include determining under what conditions the disease will be endemic, and alternatively under what conditions the disease will

eventually "die out". To determine these conditions, one computes a well-known threshold parameter called $R_0$ (pronounced "R-nought"). This *basic reproductive number*, although unitless, can be viewed as the number of new infections caused when an infected individual is introduced into an entirely susceptible population. If, for a particular disease, $R_0 > 1$, the disease can be expected to persist, and if $R_0 < 1$, the disease is expected to eventually die out. Standard analysis of a system of ODEs includes finding *equilibria* or fixed points of the system. In the field of epidemiology, equilibria have very specific meaning. In general, there are 2 types of equilibrium values, the disease-free equilibrium (DFE) and endemic equilibrium (EE) (there may be more than one of these). The most general method to determine under what conditions the equilibria will be locally asymptotically stable is to use the Jacobian matrix, but in epidemiological compartmental models, one of the most widely used methods is to use the next-generation matrix to compute $R_0$ [24]. Then, in terms of equilibria, if $R_0 < 1$, then the DFE is locally asymptotically stable, while if $R_0 > 1$, then the DFE is unstable. We note that these criterion do not always hold, and there are cases in which a model might have unstable endemic equilibria and the DFE is stable, yet the disease may still persist [85].

## 1.2 Problem statement

Because of the complex nature of *T. cruzi* transmission cycles, the mechanisms through which the disease is spread among the sylvatic populations require further study. Migration of disease vectors has been given little attention, yet may play an important role in *T. cruzi* transmission. Insect populations are particularly sensitive to climate variations and changes, which in turn affects the distribution of the populations infected with vector-borne diseases [21]. This indicates a need to examine effects of migration of *Triatoma* vectors on the spread of *T. cruzi*. Although few hu-

man cases have been reported in the U.S., migration of disease vectors from southern regions may spread the more virulent strains of *T. cruzi* northward if environmental conditions permit the vector (and host) to live and reproduce under the climatic conditions.

In this study, we wish to examine several important aspects related to migration of *Triatoma* vectors. First, because there are so few studies on dispersal and migration capabilities of *Triatoma* vectors, we derive a theoretical framework designed to estimate migration rates if certain vector dispersal parameters are known. This framework will be used to estimate migration rates for the primary *T. cruzi* vectors in parts of the southeastern U.S. and northern Mexico. Because the movement of vectors may be the link among the various sylvatic vector-host cycles, we wish to determine to what extent migration affects the prevalence of *T. cruzi* infection among the sylvatic settings in the aforementioned regions of North America. This effect will be studied via an SI compartmental model of ordinary differential equations, for which standard threshold quantities such as $R_0$ are calculated. The model depicts distinct transmission cycles linked by vector migration. This model will be analyzed numerically to determine how prevalence of *T. cruzi* is affected by varying migration rates. Finally, a more spatially explicit model will be introduced, in order to answer questions related to speed (and direction) of invasion of the epidemic. This model uses the simpler SI model as a basis to develop a high resolution cellular automaton (a grid based model in which the status of each cell in the grid depends on certain rules based on the status of neighboring cells). The vector migration rates will be derived from the local migration framework derived in Chapter 1.

# CHAPTER 2

# VECTOR MIGRATION AND DISPERSAL RATE FOR SYLVATIC *T. CRUZI* TRANSMISSION

## 2.1 Introduction

The parasite *T. cruzi*, endemic to the Americas, is maintained primarily in sylvatic cycles, although the majority of research on the transmission and control of the parasite has been devoted to domestic and peridomestic cycles in which the parasite can be transmitted to humans causing Chagas' disease. Direct transmission from human to human is not possible (except through blood transfusion and vertical transmission) and human to vector transmission is too inefficient for the parasite to maintained in domestic cycles alone [32]; thus a thorough understanding of sylvatic *T. cruzi* transmission cycles is critical to any long-term disease control plans. Because the parasite is maintained in sylvatic cycles and communication among the cycles occurs mainly as a result of vector movement among different sylvatic host populations, the need to study the spatial spread of *T. cruzi* across sylvatic regions is of importance.

The sylvatic transmission of *Trypanosoma cruzi* is complex due its presence and circulation in multiple hosts and vector species as well as large variation in strain types. The hosts, vectors, and strains of *T. cruzi* vary throughout North and South America. In sylvatic settings in North America, the parasite is transmitted between *Triatoma* vectors and reservoir hosts, mainly raccoons (*Procyon lotor*), opossums (*Didelphis virginiana*), and several species of woodrat, namely the southern plains woodrat (*Neotoma micropus*) [69]. In the U.S., hosts are associated with two strains of *T. cruzi*, types I and IIa (recently reclassified as strain IV [96]). In this study, we are

particularly interested in the sylvatic transmission of *T. cruzi* IV in cycles that range from northern Mexico to throughout the south-southeastern United States because of the overlap of distinct transmission cycles. In the wild, the infection cycles with type IV are maintained by the *Triatoma* species, *Triatoma gerstaeckeri* and *Triatoma sanguisuga*, and hosts, raccoons and woodrats. The raccoon is found throughout most of North America, but primarily inhabits wooded areas, and is commonly found near water. The woodrat habitat is characterized by areas supporting cactus growth and thorny desert shrubs. Little information is given in the literature regarding the habitats of *T. gerstaeckeri* and *T. sanguisuga*. However, one study [41] gives an account of the biogeography of these species in Texas. Based on their information and other studies, *T. gerstaeckeri* is found in dryer regions with a dense but scrubby vegetation [89], while the habitat of *T. sanguisuga* covers a broader range of vegetation. *T. sanguisuga* has been found in Texas and other southeastern U.S. states.

In this region, there are several known distinct transmission cycles. In northern Mexico and the southernmost part of Texas, the main *T. cruzi* vector is *Triatoma gerstaeckeri*, associated in sylvatic settings almost exclusively with the southern plains woodrat [41, 65, 27]. In the eastern portion of Texas into the southeast United States, the predominant vector is *Triatoma sanguisuga*, part of the *lecticularia* complex [72, 39] and is commonly associated with raccoons and opossums [66]. A recent study done in the U.S. analyzed 107 isolates of *T. cruzi*, and determined that raccoons (and other hosts) are primarily infected with type IV, while opossums were only infected with type I [69], and in fact are immune to type IV [70]. Since we are concerned with the transmission of type IV, we do not include opossums as a host in our study. In Texas and Mexico, the two cycles overlap due to *T. sanguisuga* also being found in association with woodrats [27].

In order to describe the distinct transmission cycles geographically, we define three distinct regions (or patches). We define two outer patches based on the *T. gerstaeckeri*-woodrat cycle and the *T. sanguisuga*-raccoon cycle, with an overlap patch containing both vector species and both hosts. The patch boundaries are based on the Omernik ecoregion system, political boundaries, and host and vector distribution maps found in literature. The Omernik system was derived in 1987 in collaboration with the U.S. Environmental Protection Agency. The Omernik system is hierarchical in structure and consists of 4 levels. Ecoregions are areas with similar ecosystems, and the boundaries are determined by patterns of vegetation, climate, geology, wildlife, water quality, soils, and human land use [62]. To determine the patches used in this model, we use the level 3 ecoregion system, consisting of 194 regions describing North America.

Since the geographical scale of Chagas disease spread ranges from northern Mexico through southeast United States, we assume communication between the patches is through the migration of the *T. cruzi* vectors, rather than migration of hosts. Although hosts may move, since the geographical scale is so large and covers a wide range of terrain, we do not assume hosts are as likely to cross into unsuitable habitat, thereby crossing patch boundaries. Thus we do not assume migration of hosts plays a significant role in spreading the parasite across regions. The term migration used here defines movement of vectors that cross patch boundaries. There is little to no information regarding dispersal capabilities of *T. sanguisuga* and *T. gerstaeckeri*. However, there have been several experimental studies on the dispersal capabilities of the South American species, *T. infestans* and *T. sordida*. The movement specific experimental studies are primarily light trap collections [86, 87] and mark-release experiments [74, 75]. The light trap collections are designed to determine seasonal variations of *triatomine* dispersal rather than to determine actual flight range, while

the mark-release experiments are designed to determine flight range under natural conditions. It is assumed upon maturation vectors initiate dispersal in search for a host. Vectors will also disperse due to death of host or if a host fails to return to the nest. Results of these experiments suggest that flight initiation is usually associated with low nutritional status and high temperatures [86], which is evidence for the fact that dispersal mainly occurs in search for a host.

Describing dispersal and migration is important for understanding ecological and epidemiological contexts and mechanisms. There have been several mathematical frameworks modeling movement of species, but many focus on modeling species movement rather than mechanistic ways to derive dispersal or migration rates. Reaction-diffusion equations have been used to model dispersal (in continuous time and space) due to the extensive amount of study and results [57, 14, and references therin]. Other types of dispersal models are based in discrete time, such as dispersal modeled with integrodifference equation by Kot et al. [42]. The focus of this work is to incorporate dispersal into a population model using a redistribution (distribution) kernel. Results show that broad-tailed dispersal distributions that can exhibit accelerating invasion speeds rather than constant-speed traveling waves. Mathematical models specific to studying dispersal of *Triatoma* vectors include models focusing on *Triatoma* invasion of domestic areas [79, 4]. In [79], Slimi et al. develop a spatio-temporal model based on cellular automata at a small geographical scale (a village), where dispersal is described by parameter defined as the average number of vectors entering and leaving the village per unit time. Results of this model suggest alternative control strategies to prevent yearly infestation of villages. Barbu et al. [4], also interested in the invasion of domestic areas, consider grid-based models. Dispersal is measured by the number of adult vectors dispersing in each cell inside the village. This study compared several different spatially explicit models and determine that vectors that invade vil-

lages are not only immigrating from peri-domestic habitats, but sylvatic habitats as well. Thus control measures should not only be focused on the peri-domestic areas, but also on the forests surrounding villages. In each of these studies, dispersal on a small geographical scale was studied. In this study, we use small-scale continuous space information to generate large-scale, discrete-space migration rates, compatible with classical metapopulation models.

We first describe the patches—the transmission cycle defining each patch as well as the geographical boundaries and landscape. We then estimate dispersal rates for vectors and apply those rates to translate local dispersal to global migration across a large geographical scale. We first consider whether vectors have a preference for direction of dispersal, and then the simplest case if there is no preference of direction of dispersal. Finally, we use the framework to calculate migration rates for the vector species mentioned here and compare rates for varying preferred directions and degree of preference for a particular direction. The framework derived here may be applicable to dispersal and migration of other species.

## 2.2 Patches

Patch 1 extends from south Texas to northern Mexico (including portions of Coahuila, Tamaulipas, and Nuevo León). Patch 1 is the region in which the primary *T. cruzi* hosts and vectors are woodrats and *T. gerstaeckeri* [40, 13]. Patch 3 extends from central Texas to the eastern half of the United States, including Louisiana, Mississippi, Alabama, Georgia, the Carolinas, and the Florida panhandle. The primary vector is *T. sanguisuga* associated with raccoons [66]. Patch 2 is the region in which the two transmission cycles overlap due to *T. sanguisuga*'s association with woodrats as well as raccoons in this region. Patch 2 includes the southwest portion of Texas and part of Coahuila, Mexico. The patches can be seen in figure 2.1.

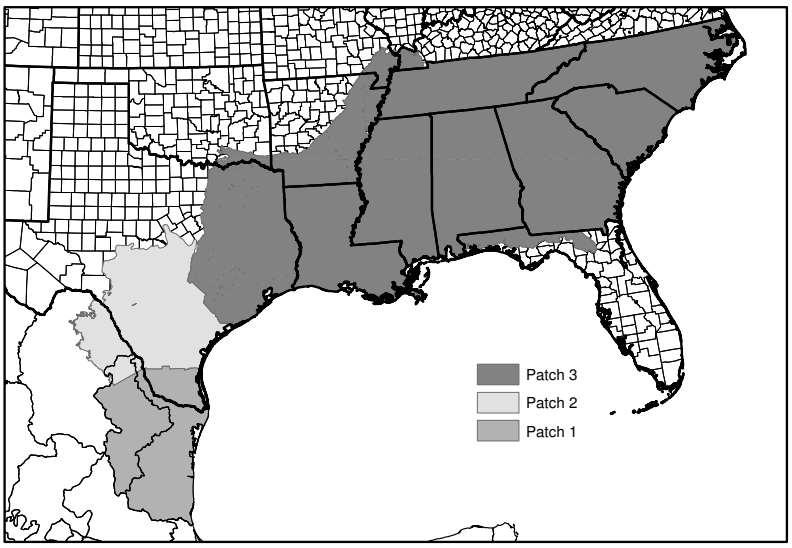


Figure 2.1. Patches.

### 2.2.1 Patch 1

Patch 1 is defined by the *T. gerstaeckeri*-woodrat transmission cycle and includes portions of south Texas and certain northern states of Mexico. The primary ecoregion in patch 1 is the Southern Texas Plains which has a generally semi-arid climate dominated by scrub, thorny brush vegetation. The thorny brushland is dominated by mesquite and certain species of cactus, among other desert scrub vegetation [61]. Curto de Casas et al. determine a geographic distribution for *T. gerstaeckeri* based on other scientific papers recording findings of this species in Mexico [20], which includes the states of Coahuila, Tamaulipas and Nuevo León. Recently Cruz-Reyes and Pickering-López conducted a review of literature regarding *T. cruzi* in Mexico, and determined *T. gerstaeckeri* to be found in the same states mentioned previously [19]. In their study, they do include a small northeast portion of San Luis de Potosí

as potential habitat for *T. gerstaeckeri*. Recently, in a study on the risk assessment of Chagas disease in Texas, Sarkar et al. [72], a risk assessment predicts high suitable habitat for *T. gersataeckeri* in northeast Mexico to central and east Texas, which aligns closely with patches 1 and 2. Since our focus is on *T. gerstaeckeri* and *T. sanguisuga*, we do not include this state in patch 1, due to the fact that this region coincides with the native area of the dominant vector in Mexico, *Triatoma dimidiata*. Because the patch 1 climate and vegetation are consistent with the southern plains woodrat preferred habitat, we are certain that this species is found through patch 1. According to the IUCN Redlist of Threatened Species [50], the southern plains woodrat has also been found in northeast Coahuila, Tamaulipas, and Nuevo León.

Because patch 1 does not include the *T. sanguisuga*-raccoon transmission cycle, the northern boundary of patch 1 is primarily based on the geographic extent of *T. sanguisuga*. Kjos et al. [41] collected samples of *T. sanguisuga*, which was not found south of the southern borders of Webb, Duvall, Jim Wells, and Kleberg counties in Texas. Furthermore, Ibarra-Cerdeña et al. [39] determined a very similar boundary in Texas for the southern range of the complex *Triatoma lecticularia*, which includes the species *T. sanguisuga*. We mention here that raccoons may be found statewide in Texas. However, based on known county records, the raccoon geographic distribution diminishes in the south to southwest portion of Texas. Beasley et al. [6] show that raccoons prefer forested areas over all other landcover types (grassland, shrubland, agriculture areas), with the least preferred habitat being shrubland. Henner et al. [37] found that raccoons prefer trees as their den site 91% more than ground den sites. They also conclude that raccoons consistently selected den sites that allowed access to free water. Thus, due to the lack of abundant tree cover and free water in patch 1, we do not include raccoons as a host of *T. cruzi* in this region. We further mention that no studies to date have associated the vector *T. gerstaeckeri* with the

raccoon, thus agreeing with the conclusions that raccoons would not be a significant host for *T. cruzi* in this region.

### 2.2.2 Patch 2

Patch 2 contains the most diverse climate and vegetation due to the overlap of host and vector species in the model. This patch contains the entire Edwards Plateau, portions of the Southern Texas Plains, and the southernmost parts of the Texas Blackland Prairies, East Central Texas Plains, and Western Gulf Coastal Plain. The vegetation is dominated in the south by the shrubland of the Southern Texas Plains. The Edwards Plateau has a similar vegetation dominated by scrub forest. Mesquite occurs throughout the entire region with ash, juniper, and Texas oak dominant in the southern and eastern regions of the plateau which extend into the Texas Blackland Prairies and East Central Texas Plains. The Blackland Prairies and East Central Texas Plains (also called the Post Oak Savanna) are predominantly covered by prairie grasses and savannas (a transition between grassland and forest). Trees in this region include Post oak and blackjack oak, with mesquite invading the southernmost parts of these regions [61]. This region also includes a portion of the Western Gulf Coastal Plain with a primary grassland vegetation; however, much of the region has been invaded by mesquite trees, oaks, and prickly pear cactus [61].

Patch 2 is defined as the region in which all four species can be found with *T. sanguisuga* feeding on both host species, woodrats and raccoons. In patch 2, woodrats are predominantly found in the regions dominated by mesquite and cactus. Studies have found that *N. micropus* densities are directly linked to availability of cactus [83, 12]. In fact, Thies et al. [83] found that cacti was the most abundant material in construction of the woodrat nests (over 50% of the material), as well as a main source of food. Thus, in general, the southern plains woodrat can be found in most of

patch 2, albeit with lower population density than in regions completely dominated by thorny brushland (i.e., patch 1). In patch 2, raccoons are mostly found in regions with trees and near water. This corresponds to raccoons most likely preferring habitat in the southern part of the Edwards Plateau into the Texas Blackland Prairies and East Central Texas Plains. In this region, both vectors *T. gerstaeckeri* and *T. sanguisuga* can be found. We further mention here that in this region, *T. sanguisuga* feeds on both woodrats and raccoons [65, 13, 66, 40]. We assume that due to local vector dispersal *T. sanguisuga* sometimes move between raccoon dens and woodrat nests so that they will be considered a single population.

The southern portion of patch 2 is determined by the geographic extent of *T. sanguisuga* (as mentioned previously). The eastern boundary of patch 2 is primarily based on the preferred habitat of the woodrat, since the woodrat prefers to makes its home in regions with cactus growth. The regions east of patch 2 are dominated by forest and have extensive tree cover, less favorable for cactus growth. According to county records, the southern plains woodrat has a distinct eastern boundary in Texas [22], which we use as the eastern boundary of patch 2. The northern boundary of patch 2 is the northern boundary of the ecoregion, Edwards Plateau. Although the southern plains woodrat habitat extends further north than this boundary, according to the work done by Kjos et al. [41], the vector *T. gerstaeckeri* has not been located any further east or north than the boundary of patch 2. The western boundary is determined by the Southern Texas Plains ecoregion. The regions west of patch 2 are predominantly desert (as part of the Chihuahuan Desert ecoregion). Although the woodrat may be found in this region, we do not have any information regarding *T. gerstaeckeri* or *T. sanguisuga* being significant vectors in this region. We note that in the risk assessment by Sarkar et al. [72], data collection of *T. sanguisuga* from two museum collections in the western United States predicts suitable habitat in areas

west of this boundary, although they mention the need for more data collection to test this prediction. Other species such as *Triatoma rubida* and *Triatoma protracta* may be the dominant vectors in this region [63].

### 2.2.3 Patch 3

Patch 3 contains the *T. sanguisuga*-raccoon *T. cruzi* transmission cycle and covers a great area of land. The vegetation in this region is dominated by forest and woodland, including the South Central Plains and Southeastern Plains as the main ecoregions. We mention here that there are other hosts of *T. cruzi* in this region, most notably opossums. However, we only consider raccoons as the primary host since we are concerned with the transmission of *T. cruzi* Type IV, to which opossums are immune [70].

The southern plains woodrat is not native to this region. The western boundary of patch 3 is determined by the eastern range of the southern plains woodrat [22] and the western boundary of *T. sanguisuga* [41]. The southern boundary of patch 3 is determined by the northern boundary of the woodrat as well as *T. gerstaeckeri* in the eastern part of Texas [41, 22]. *T. sanguisuga* and the raccoon are found in many parts of the eastern half of the U.S. covering a broad expanse of climate and vegetation. We note, however, that the southern boundary of patch 3 extends not further than the Florida panhandle. The ecoregion south of patch 3 is the Southern Coastal Plain, with a broad range of vegetation, but mostly consists of coastal lagoons, marshes, and swampy lowlands. Although raccoons may inhabit this type of land, research suggests that the complex of the triatomine species, *T. lecticularia*, which includes *T. sanguisuga*, is not found in this region. There are reports of *T. sanguisuga* found north of the northern boundary of patch 3. They are isolated reports, and *T. sanguisuga* is not believed to be as dense north of this boundary as they are in the the southern

part of the U.S.; thus we do not include the regions to the north of Tennessee and North Carolina.

## 2.3 Dispersal and migration rates

To estimate the parameters that describe the movement of vectors across patch boundaries, we will consider two types of rates, dispersal and migration rates. Dispersal refers to any movement of vectors, while migration will be used to denote movement only across patch boundaries. We consider the patches in continuous 2-D space so that at every point exists some density of vectors. We assume that there are two main reasons a vector may migrate in search for a host: maturation and loss of host. We assume that upon maturation, vectors migrate in search for a host. Furthermore, we also assume that some proportion of adult vectors may disperse to search for a new host when the previous host dies or fails to return to the nest. We note here that not all vectors that mature may need to disperse if there is a host present in their location upon maturation. Thus, the vector dispersal rate is a sum of two rates, $m_1$ and $m_2$. We define $m_1$ as the dispersal rate due to maturation, and $m_2$ as the dispersal rate due to loss of host. Due to difficulty estimating $m_2$, we will estimate $m_1$ using the assumption that 100% of vectors that mature will migrate so as to include the dispersal rate due to loss of host. We will then calculate the cross-patch migration rate as the product of the dispersal rate, $m$ and the migration proportion (the proportion of dispersals that cross patch boundaries.) We will denote the migration proportion as $M$ and the cross-patch migration rates will be denoted as $\bar{m}$.

### 2.3.1 Dispersal rates

We first separately calculate the rate $m$ at which vectors in any location will disperse. Since we assume that vectors begin to travel at maturity, we consider the 2-stage model with constant local population including per capita birth, death and maturation rates, where $J$ represents the local population of juveniles (nymphs) at time $t$, and $A$ represents the local population of adults, and $J + A = N$.

$$J' = rA\left(1 - \frac{A}{K}\right) - \gamma J - \mu_J J,$$
$$A' = \gamma J - \mu_A A,$$

where $r$ represents the intrinsic growth rate, $K$ is the carrying capacity of the vector population, $\mu_J$ and $\mu_A$ are the per capita natural mortality rates for juveniles and adults, respectively, and $\gamma$ represents the rate at which juveniles become adults (or maturation rate, units 1/year).

We can now determine the dispersal rate $m$ for each species due to maturation as the per capita maturation rate $\gamma$ times the proportion of juveniles at equilibrium. We define the proportion of juveniles at equilibrium as $x^* = \frac{J^*}{N^*}$, so then $(1-x^*) = \frac{A^*}{N^*}$. To determine $x^*$, we use the proportionalized equilibrium condition ($J' + A' = 0$),

$$\gamma x^* N^* - \mu_A(1 - x^*)N^* = 0,$$

and determine for $N^* \neq 0$, $x^* = \frac{\mu_A}{\mu_A+\gamma}$. Thus, for each vector species,

$$m_1 = \gamma x^* = \gamma\left(\frac{\mu_A}{\mu_A + \gamma}\right).$$

Using the parameter estimates from [44], for *T. sanguisuga*($\gamma = 1/(2.25\text{year})$, $\mu_A = 1/(1.44\text{year})$, we obtain $m_{1S} = 0.271/\text{year}$. For *T. gerstaeckeri* ($\gamma = 1/\text{year}$, $\mu_A = 1/0.78\text{year})$, $m_{1G} = 0.562/\text{year}$.

## 2.4 Preferred direction

To derive the migration proportion (the proportion of dispersals that cross the patch boundary), we wish to sum up all of the local vector dispersals originating in one patch that cross the patch boundary into the neighboring patch. As shown in Figure 2.2, we wish to sum up all of the dispersals originating from each point $(x, y)$ that end up on the other side of the patch boundary. In order to derive the migration rate, we consider two different coordinate systems. The local coordinate system, used to sum all of the patch-crossing vectors originating from a given point, will be polar, while the global coordinate system summing over all points of origin will be rectangular. To simplify calculations, we will assume that the patch boundary is piecewise linear.

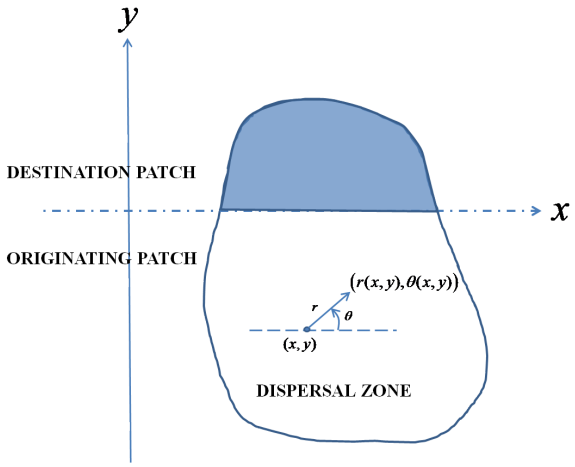


Figure 2.2. Interaction between global and local coordinate systems.

We consider that vector dispersal is described by three properties: dispersal distance, preferred direction of dispersal, and degree of preference for a particular direction. Each of these properties is inherent to vector dispersal and independent of the coordinate systems. We model vector dispersal as a piecewise constant distribu-

tion in two-dimensional continuous space. To account for the preferred direction, we define this distribution using a sequence of nested ellipses, where each ellipse in the sequence represents a particular level of dispersal. The dispersal is constant on rings bounded by the ellipses, with the appropriate proportion of vectors dispersing within each ring.

Dispersal properties are reflected in the distribution parameters. We define the sequence $\{b_i\}$ as the threshold dispersal distances perpendicular to the preferred direction. The values $\theta_0$ and $e$ represent the preferred direction and degree of preference, respectively. These values are global parameters and do not depend on any coordinate systems. The value $\theta_0$ is defined counterclockwise relative to the outward normal to the patch boundary. The degree of preference parameter $e$ gives the eccentricity of each ellipse. We note that $0 < e < 1$, so that for values of $e$ very close to 1, the degree of preference is very strong, while if $e = 0$, we assume the vectors have no preferred direction of dispersal, and the model reduces to concentric circles, which will be described in detail in the following section. In Figures 2.3 and 2.4, we illustrate vector dispersal with sets of nested ellipses for varying values of $\theta_0$ and $e$.

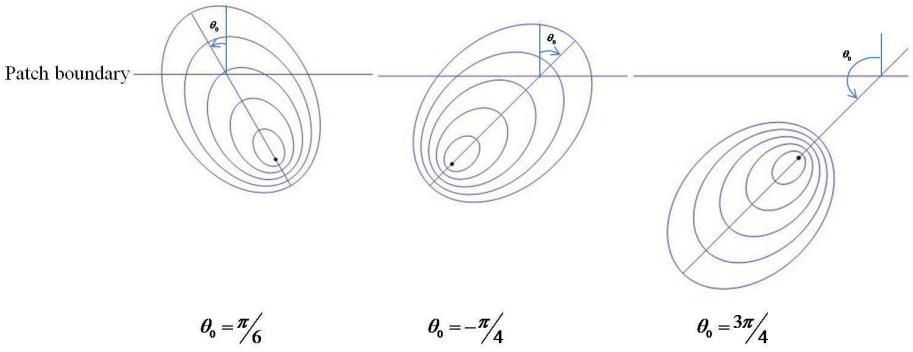


Figure 2.3. Varying direction with bold points of origin ($e = 0.7$).

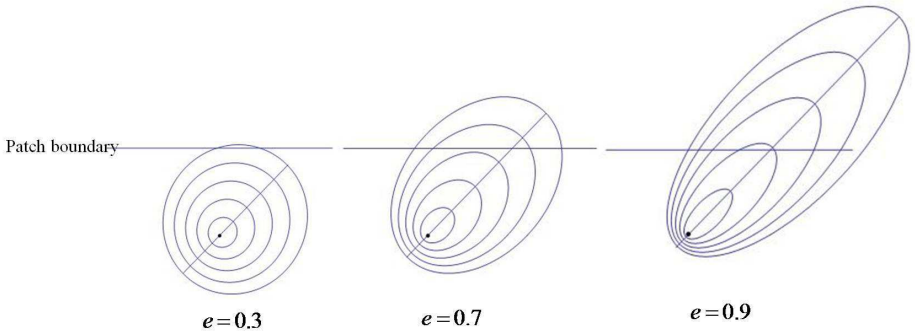


Figure 2.4. Varying eccentricity with bold points of origin ($\theta_0 = -\pi/4$).

Next, we will derive a framework for a preferred direction of dispersal. We first define (local) dispersal distribution(s) using nested ellipses in which dispersal originates from a given point. Then we derive the global migration proportion from each point of dispersal in the originating patch.

### 2.4.1 Local migration framework

To determine the total number of patch crossing dispersals originating from a given point, we will consider a local coordinate system defined by polar coordinates $(r, \theta)$. We assume that vector dispersal originates at the pole. The variable $r$ defines the dispersal distance from the pole and $\theta$ is measured counterclockwise. The reference value $\theta = 0$ is measured $\pi/2$ clockwise from the outward normal to the patch boundary (see Figure 2.2).

The preferred direction of migration will be modeled using a sequence of nested ellipses with a common focus at the pole. Then each ellipse $r_i$ is given by the equation

$$r_i(\theta) = \frac{b_i\sqrt{1-e^2}}{1 - e\sin(\theta - \theta_0)}.$$

As stated previously, $\theta_0$ and $e$ are global parameters depending only on the nature of vector dispersal and independent of the coordinate systems used. The ellipse $r_i(\theta)$ is defined so that the major axis points toward the direction of $\theta_0$. As

mentioned previously, the sequence $\{b_i\}$ defines the dispersal distances perpendicular to the preferred direction, $\theta_0$. Based on dispersal data, we define a function, $\Psi(r,\theta)$, that measures proportion of vectors per square kilometer. Because *Triatoma* dispersal data is scarce we will define $\Psi(r,\theta)$ to be piecewise constant on elliptical rings $R_i(\theta) = \{(r,\theta) : r_{i-1} < r \leq r_i\}$, where all $r_i$ have the same $e$ and $\theta_0$, but differing widths, given by $b_i$. Then

$$\Psi(r,\theta) = \begin{cases} f_1, & 0 < r \leq r_1(\theta) \\ f_2, & r_1(\theta) < r \leq r_2(\theta) \\ f_3, & r_2(\theta) < r \leq r_3(\theta) \\ . & \\ . & \\ . & \\ f_n, & r_{n-1}(\theta) < r \leq r_n(\theta) \end{cases}$$

where each constant function $f_i$ is the proportion of vectors per square kilometer dispersing from the focus at the pole to the $i^{th}$ ring $R_i(\theta)$ (measured in kilometers). We further assume no vectors disperse further than $r_n(\theta)$.

To construct $f_i$, we use the proportionalized population density for each elliptical ring, where $c_i$ is the proportion of vectors dispersing from the given point to $R_i$, so then $\sum_i c_i = 1$. Then the proportion of vectors per square kilometer is

$$f_i = c_i / A(R_i),$$

where $A(R_i)$ is the area of the elliptical ring bounded by $r_i$ and $r_{i-1}$. Then $A(R_i) = \pi(a_i b_i - a_{i-1} b_{i-1})$, where $a_i$ is the length of the semi-major axis. Since $a_i$ can be given in terms of the length of the semi-minor axis $b_i$ as $a_i = \frac{b_i}{\sqrt{1-e^2}}$, we have

$$A(R_i) = \frac{\pi(b_i^2 - b_{i-1}^2)}{\sqrt{1-e^2}},$$

and consequently

$$f_i = \frac{c_i\sqrt{1-e^2}}{\pi(b_i^2 - b_{i-1}^2)}. \tag{2.1}$$

Here $b_i$ and $b_{i-1}$ are half the length of the minor axis for the outer and inner ellipses, respectively. Furthermore, we note that $\int\int \Psi(r,\theta)\, dA = \sum_i f_i A(R_i) = \sum_i c_i = 1$.

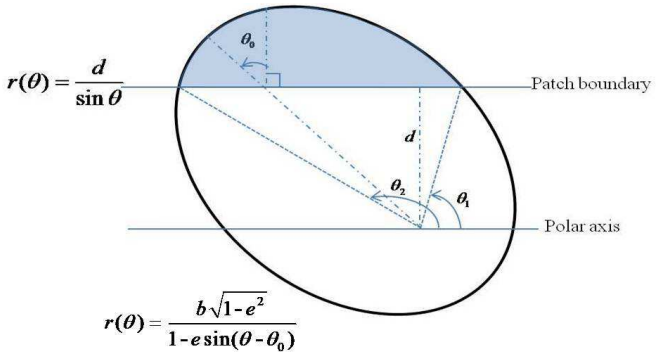


Figure 2.5. Preferred direction framework. Shaded region represents the area for which dispersals have crossed patch boundary originating from a distance $d$ away from the patch boundary..

We assume that vector dispersal originates at a focus of the ellipse located at the pole at a distance, $d$, away from the patch boundary. To compute the vector dispersals crossing the patch boundary, we wish to determine the two $\theta$ values, $\theta_1$ and $\theta_2$ (see Figure 2.5), such that the outermost ellipse $r_n(\theta)$ intersects the patch boundary. The parameters $d$, $\theta_1$ and $\theta_2$ are constant over $r, \theta$, but variable over $x, y$. Thus, for the point of dispersal, $d$ $km$ from the patch boundary, we solve

$$\frac{d}{\sin(\theta)} = \frac{b_i\sqrt{1-e^2}}{1 - e\sin(\theta - \theta_0)} \tag{2.2}$$

for $\theta$, obtaining two $\theta$ values, $\theta_1(d)$ and $\theta_2(d)$, which become the lower and upper limits, respectively, of the $d\theta$ integral. We note that $\theta_1$ and $\theta_2$ depend on $e$ and $\theta_0$

as well. But, since these are global parameters, we do not write $\theta_1$ and $\theta_2$ as explicit functions of $e$ and $\theta_0$.

By solving (2.2), we determine by using the identity $\sin(A-B) = \sin A \cos B - \cos A \sin B$,

$$\begin{aligned} \frac{d}{\sin\theta} &= \frac{b_n\sqrt{1-e^2}}{1-e\sin(\theta-\theta_0)} \\ d - de\sin(\theta-\theta_0) &= b_n\sqrt{1-e^2}\sin\theta \\ d - de\sin\theta\cos\theta_0 + de\cos\theta\sin\theta_0 &= b_n\sqrt{1-e^2}\sin\theta \\ (de\cos\theta_0 + b_n\sqrt{1-e^2})\sin\theta - de\sin\theta_0\cos\theta &= d \end{aligned} \tag{2.3}$$

Since we can write a sum of two sinusoidal functions as a single sinusoidal function, then $a_1\sin\theta + a_2\cos\theta$ can be written as $B\sin(\theta+\phi)$, where $B = \sqrt{a_1^2+a_2^2}$ and $\phi = \arctan(a_2/a_1)$. Then in equation (2.3),

$$B = \sqrt{(de\cos\theta_0 + b_n\sqrt{1-e^2})^2 + (de\sin\theta_0)^2}$$

and

$$\phi = \arctan(-de\sin\theta_0/(de\cos\theta_0 + b_n\sqrt{1-e^2})).$$

It then follows that $\theta_1 = \arcsin(d/B) - \phi$ and $\theta_2 = \pi - \arcsin(d/B) - \phi$.

We define $\hat{M}(-d)$ (for a negative argument indicating location relative to the patch boundary) as a unitless quantity representing the proportion of vectors crossing the patch boundary originating from the pole. Thus,

$$\hat{M}(-d) = \int_{\theta_1(-d)}^{\theta_2(-d)} \int_{-d/\sin\theta}^{r_n(\theta)} \Psi(r,\theta)\, r\, dr\, d\theta,$$

where we define $\hat{M}(-d) = 0$ for $d$ such that $r_n(\theta)$ does not intersect $d/\sin(\theta)$. A visual representation can be seen in Figure 2.5. Furthermore, we define the maximum distance a vector can be from the patch boundary and still disperse to the patch boundary, as the point on the outermost ellipse such that there is only one intersection

with the patch boundary. This maximum distance, denoted $r_{max}$, is calculated by finding the unique solution to $\theta_1(-d) = \theta_2(-d)$. If $\theta_1(-d) = \theta_2(-d)$, it follows that $-d = B$, so then

$$r_{max} = b_n \left( \frac{e}{\sqrt{1-e^2}} \cos\theta_0 + \sqrt{1 + \frac{e}{\sqrt{1-e^2}} \cos\theta_0} \right). \tag{2.4}$$

### 2.4.2 Global migration framework

To sum patch-crossing vector dispersals over all points of origin in the given patch, we define now a rectangular global coordinate system for each boundary segment. The origin of this system is defined to be the counterclockwisemost point (with regard to the originating patch) on the piecewise linear boundary segment between the two patches (see Figure 2.6 for an illustration) and the angle between consecutive segments is no greater than 90° in either direction. Each point of dispersal is then described by the rectangular coordinates, $(x, y)$, with $y$ defined by the vertical distance to the patch boundary, where $y < 0$ represents the originating patch and $y > 0$ represents the destination patch. Thus the patch boundary is given for each segment by the line $y = 0$. We further note that $r_{\max}$ differs for each segment because $\theta_0$ differs for each segment.

We note that as mentioned in section 2.3.1, only a proportion of vectors are capable of dispersal. We wish to calculate the proportion of vectors in the entire patch that cross the patch boundary. Because $\hat{M}(d)$ is the proportion of vectors originating from the pole that cross the patch boundary, we integrate $\hat{M}(y)$ (since in the global coordinate system, $d = -y$) for each point $x, y$ within the maximum dispersal range. This integral will give us the area times the proportion of all vectors dispersing. Thus, to determine the cross-patch migration rate, we divide by the area

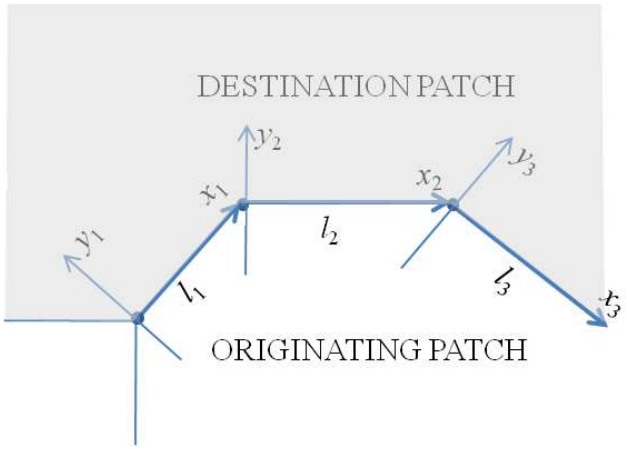


Figure 2.6. Global coordinate system.

$A$ of the patch from which the dispersal originates and multiply by the dispersal rate, $m$, to obtain

$$\bar{m}_j = \frac{m}{A} \int_0^{l_j} \int_{-r_{max}}^{0} \hat{M}(y)\, dy\, dx,$$

where $l_j$ represents the length of the $j^{th}$ linear boundary segment.

Furthermore, since the integral does not depend on $x$, $\bar{m}_j$ becomes

$$\bar{m}_j = \frac{m \cdot l_j}{A} \int_{-r_{max_j}}^{0} \hat{M}(y)\, dy \tag{2.5}$$

for each linear segment of boundary. Then $\bar{m} = \sum_j \bar{m}_j$.

This method does have some limitations and constraints. First, in order for the rectangular coordinate systems described above to remain entirely within the patches of origin, the angle turn between any two consecutive segments be no greater than 90°. Second, the method considers dispersals originating from a given point to cross a single linear segment of a patch boundary. In practice, dispersals which originate sufficiently close to the corner where two such segments meet may cross boundary segment lines beyond the segments' endpoints, in which case the method

miscounts them (overcounts or undercounts). To address the issue, we also require that $\min_j l_j > \bar{a}_{\max}$,i.e., the shortest boundary segment must still be longer than the longest dispersal distance, in order for the dispersal ellipse generated from any point to cross no more than two boundary segments. Then we introduce corrective calculations to resolve the errors induced near corers.

For each type of corner (concave or convex), these computations cover regions in which the originating dispersals are overcounted (counted as crossing a patch boundary, when they should not be) or undercounted (failure to count dispersals that cross patch boundaries). The appendix provides detailed definitions of each error type. The numerical approximations for the migration rates in sections 2.5.2 and 2.6 have been adjusted for the error, which, for the patches in this study, sum to no more than 3%.

## 2.5 No preferred direction

In the simplest case, we assume vectors have no preferred direction of dispersal. We then consider the eccentricity $e$ is 0, and thus the nested ellipses reduce to concentric circles, each with radius $b_i$. Since when $e = 0$, each $r_i$ becomes $b_i$, $\Psi$ depends only on the sequence $\{b_i\}$. In this case, vector dispersal originates from the center of the circles. Here,

$$\Psi(r,\theta) = \begin{cases} f_1, & 0 < r \leq b_1 \\ f_2, & b_1 < r \leq b_2 \\ f_3, & b_2 < r \leq b_3 \\ . & \\ . & \\ . & \\ f_n, & b_{n-1} < r \leq b_n \end{cases} \tag{2.6}$$

and

$$f_i = \frac{c_i}{\pi(b_i^2 - b_{i-1}^2)}. \tag{2.7}$$

### 2.5.1 Estimation of dispersal kernel

In order to derive numerical values for $\Psi$ with no preferred direction, we utilize two studies in which *Triatoma* vector dispersal is the focus. Both studies were mark-release experiments performed in the vector's natural climatic conditions in which a preferred direction of dispersal could not be determined. Due to the lack of dispersal information regarding *T. gerstaeckeri* and *T. sanguisuga*, we use two studies on two South American *Triatoma* species, *T. infestans* and *T. sordida* [74, 75]. Although these species are not naturally located in the patches in this model, we assume that the dispersal capabilities of each of these species are at least an upper bound for the dispersal capabilities of *T. gerstaeckeri* and *T. sanguisuga*.

For each study, vectors were released from a central location and captured during the night. To determine distances flown, circles were drawn out with fluorescent paint at regular intervals (in meters). For each study, the proportion of vectors traveling within a specific range of distances was reported. We give the raw data and an average of those proportions in Tables 2.1 and 2.2. We note that in [74], specific values for 60, 75, and 90 m were not given, but rather values for <100 m. In order to use this data with the same intervals provided in [75], we decomposed the data in [74] for values <100 m by using the same proportions in [75]. These values are marked with an * in Table 2.1. In each study, there appears to be a distinction between trivial flights (0-5 m) and longer distance flights (> 50 m). Thus, for vector dispersal we assume that all flights were either trivial or long range, i.e., we assume no flights occur for 5-50 m.

Table 2.1. Dispersal data from [74, 75] for a single flight

| [74] | [75] | Range |
|---|---|---|
| 0.552 | 0.469 | 0-5m |
| 0.034* | 0.015 | 60m |
| 0.018* | 0.008 | 75m |
| 0.018* | 0.008 | 90m |
| 0.378 | 0.504 | >100m |

Table 2.2. Averages of proportions of vectors dispersing within 0-100m for a single flight

| Proportion | Interval |
|---|---|
| $c_1 = 0.5105$ | 0-5m |
| $c_2 = 0$ | 5-50m |
| $c_3 = 0.0375$ | 50-75m |
| $c_4 = 0.0125$ | 75-100m |
| $c_5 = 0.4395$ | >100m |

In each study, only a proportion of the vectors were actually recovered, and it is assumed that the remainder of the bugs flew farther than 100m. After taking the averages of each study, we determine that 44% of the bugs flew farther than 100m. Since we do not have a maximum dispersal distance, we will derive the numerical values for $\Psi$ using three possible values (low-mid-high estimates) for the maximum dispersal distance. We further note each value of $b_i$ represents a range of distances traveled for one flight.

Because we have 5 intervals determined by the experiments, to determine $b_5$ assuming a minimum dispersal capability we first note that in equation (2.7), $c_i$ represents the proportion of bugs traveling between $b_{i-1}$ and $b_i$ meters for one flight. Thus, $c_5 = 0.44$ represents the proportion of bugs traveling between 100 and $b_5$ meters

and $f_5 = \frac{0.44}{\pi(b_5^2 - 100^2)}$. If $b_5$ is to be a minimum and the distribution is to be unimodal, then $f_5 = f_4$. The remainder of the $c_i$ values are given in Table 2.2. Thus,

$$\begin{gathered} f_4 = f_5 \\ \frac{c_4}{\pi((b_4)^2 - (b_3)^2)} = \frac{c_5}{\pi((b_5)^2 - (b_4)^2)} \\ \frac{0.125}{\pi(100^2 - 75^2)} = \frac{0.44}{\pi(b_5^2 - 100^2)} \end{gathered} \tag{2.8}$$

Solving equation (2.8) for $b_5$, we determine $b_5 = 405$m.

To determine a maximum value for $b_5$, we consider results published by Schofield et al. [48]. In this paper it is determined that *Triatoma* vectors are capable of flights up to 1350 meters. Thus, we choose a maximum value for $b_5$ to be 1350m. We note that because of limited data, in all future calculations, we will use the high dispersal range of 1350m.

Based on experimental results by Pippin [65] regarding the vectors in the model, we estimate that on average a vector will travel to find a host within 2 weeks after feeding, but will starve (or no longer have the energy to fly) within 3 weeks of flying. Then, the maximum number of flying days is estimated to be 35. We note that because of the limited amount of data, this number is only an estimate determined by one study. Then, in order to model dispersal over more than one flight, we need a distribution of the sum of the 35-day flight distances. As we increase the number of days flown to 35, by the Central Limit Theorem, the 35-day distribution will become approximately normally distributed with mean, 35 times the mean for the original 1-day distribution and standard deviation $\sqrt{35}$ times the standard deviation of the original distribution.

To develop the 35-day distribution, we determine the proportion of vectors flying per meter for each of the ranges in Table 2.2. We use these values as estimates for the heights of a 1D piecewise constant function, (call $\bar{P}$) with intervals of length 1 meter.

To obtain the 1D distribution for the sum of the flight distances for 35 days of flying, we will repeatedly apply the convolution to $\bar{P}$ 35 times. For computational simplicity, we approximate this normal distribution with a piecewise constant function, $P(r)$, with units proportion of vectors per meter, where

$$P(r) = h_j \text{ for } b_{(j-1)} < r \leq b_j \text{ with } b_0 = 0 \tag{2.9}$$

We chose $P$ to have 9 pieces for optimal processing time. The interval for the final piece is chosen to be the entire right tail of the normal distribution. The remaining breakpoints and values for the function are then determined by a least-squares optimization routine. Finally, we multiply each value of $P$ by its appropriate interval to obtain the proportion of vectors traveling within the given distance interval using the equation

$$c_j = h_j \cdot (b_j - b_{(j-1)}). \tag{2.10}$$

Recall, the distribution function $\Psi$ is a function of 2 variables, $r$ and $\theta$, where

$$\Psi(r, \theta) = f_j \text{ for } b_{(j-1)} < r \leq b_j \tag{2.11}$$

In order to derive estimates for $\Psi$, we use the $c_j$ values (as the proportion of vectors dispersing to the $j^{th}$ elliptical ring) in equation (2.1).

### 2.5.2 Calculation of migration rates for no preferred direction

In the case of no preferred direction, the intersection of the patch boundary with the outermost circle, is found by solving

$$\frac{d}{\sin(\theta)} = b_9,$$

where $b_9$ is the radius of the outermost circle. Then,

Table 2.3. Estimated 35-day vector dispersal distribution, extrapolated from single-flight data in Tables 2.1 and 2.2

| $j$ | $b_j$ | $h_j$ |
|---|---|---|
| 1 | 2485.05 | $9.96 \times 10^{-8}$ |
| 2 | 4970.11 | $2.40 \times 10^{-6}$ |
| 3 | 7455.16 | $2.34 \times 10^{-5}$ |
| 4 | 9940.21 | $9.26 \times 10^{-5}$ |
| 5 | 12425.26 | $1.51 \times 10^{-4}$ |
| 6 | 14910.32 | $1.02 \times 10^{-4}$ |
| 7 | 17395.37 | $2.81 \times 10^{-5}$ |
| 8 | 19880.42 | $3.18 \times 10^{-6}$ |
| 9 | 47216 | $3.58 \times 10^{-10}$ |

Table 2.4. Geographical parameter values

| Patch area | Bound. lengths |
|---|---|
| $A_1 = 1.594 \times 10^{11}\text{m}^2$ | $l_{AB} = 193,600\text{m}$ |
| $A_2 = 1.783 \times 10^{11}\text{m}^2$ | $l_{BC} = 137,707\text{m}$ |
| $A_3 = 1.133 \times 10^{12}\text{m}^2$ | $l_{DE} = 99,691\text{m}$ |
| | $l_{EF} = 103,427\text{m}$ |

$$\theta_1(d) = \arcsin(d/b_9) \text{ and } \theta_2(d) = \pi - \arcsin(d/b_9). \tag{2.12}$$

To manage the complexity of the calculations we will break the patch boundaries in Figure 2.1 into four linear segments (two for patch 1/2 boundary and two for patch 2/3 boundary) as seen in Figure 2.7. We will calculate separate migration rates for each (appropriate) vector species between patches 1 and 2 and between patches 2 and 3. We define $l_{12}$ as the sum of the length of the two linear segments making up the boundary between patches 1 and 2, while $l_{23}$ is the sum of the lengths of the linear segments making up the boundary between patches 2 and 3. But, we note that for $e = 0$, we do not need to consider multiple boundary segments, since there is no preference for direction. Thus, rather than using separate terms for each boundary

segment, we use one term with the sum of the lengths of each boundary segment for each patch boundary.

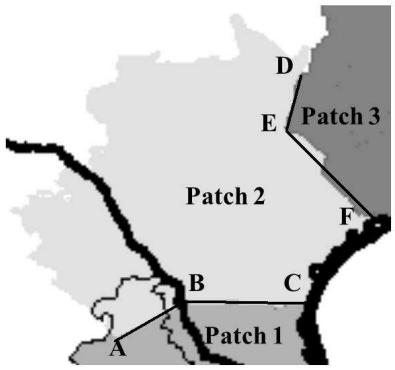


Figure 2.7. Patch boundaries linear segments.

To calculate the migration rate of *T. gerstaeckeri* from patch 1 to patch 2,

$$\bar{m}_{12} = \frac{m_G}{A_1}(l_{12}) \int_{-r_{max}}^{0} \int_{\theta_1(y)}^{\theta_2(y)} \int_{\frac{y}{\sin\theta}}^{r_9} \Psi(r,\theta)\, r\, dr\, d\theta\, dy, \tag{2.13}$$

where the $\theta$ limits are given in equation (2.12), $r_{max} = b_9$, $m_G = 0.552$/year. The geographic parameters are given in Table 2.4.

The migration rate of *T. gerstaeckeri* from patch 2 to patch 1 is

$$\bar{m}_{21} = \frac{m_G}{A_2}(l_{12}) \int_{-r_{max}}^{0} \int_{\theta_1(y)}^{\theta_2(y)} \int_{\frac{y}{\sin\theta}}^{r_9} \Psi(r,\theta)\, r\, dr\, d\theta\, dy. \tag{2.14}$$

Then, $m_{12} = 0.0042$/year and $m_{21} = 0.0038$/year.

The migration rate of *T. sanguisuga* from patch 2 to 3 and 3 to 2 can be calculated similarly. The geographical parameter values for no preferred direction are given in Table 2.4.

## 2.6 Numerical calculations for preferred direction

We now consider how preferred direction affects the migration rate. In the case of preferred direction,

$$\Psi(r,\theta) = f_j \text{ for } b_{(j-1)} < r < b_j \tag{2.15}$$

where

$$f_i = \frac{c_i\sqrt{1-e^2}}{\pi(b_i)^2 - (b_{i-1})^2)}$$

and

$$r_i(\theta) = \frac{b_i\sqrt{1-e^2}}{1-\sin(\theta-\theta_0)}.$$

The specific values for $c_i$ are determined by the the values in Table 2.3 and equation (2.10).

Based on the description of calculating the migration rates for a preferred direction in section 2.5.2, we calculate the migration rates for varying preferred directions. The migration rate from patch 1 to 2 is then defined to be

$$\begin{aligned}\bar{m}_{12} = &\frac{m_G}{A_1}\left(l_{AB}\int_{-r_{\max}}^{0}\int_{\theta_1(y)}^{\theta_2(y)}\int_{\frac{y}{\sin\theta}}^{r_9(\theta)}\Psi(r,\theta)\,r\,dr\,d\theta\,dy\right)\\ &+\frac{m_G}{A_1}\left(l_{BC}\int_{-r_{\max}}^{0}\int_{\theta_1(y)}^{\theta_2(y)}\int_{\frac{y}{\sin\theta}}^{r_9(\theta)}\Psi(r,\theta)\,r\,dr\,d\theta\,dy\right).\end{aligned} \tag{2.16}$$

We note that when $r_{\max}$ appears in each integral, it is specific to each patch boundary (since it depends on $\theta_0$ (see equation (2.4))). Furthermore, when $\theta_0$ appears in each integral, it is defined to be the angle counterclockwise from the outward normal of the specific linear boundary segment to the major axis of the ellipse. One way to determine $\theta_0$ is to first define the angle made by the major axis of the ellipse, $\bar{\theta}$ (in the counterclockwise direction), where $\bar{\theta}_0 = 0$ represents an ellipse with a northward preferred direction. We then define the angle of the outward normal to the $j^{th}$ boundary segment, $\theta_j$ (in the counterclockwise direction), where $\theta_j = 0$ is the angle for the

Table 2.5. $\theta_j$ values for outward normals for patch boundary segments

| Segment | $\theta_j$ |
|---|---|
| AB | 215° |
| BC | 0° |
| DE | 80° |
| EF | 108° |

outward normal for a horizontal boundary segment (running east to west). Then,we define $\theta_0$ for segment $j$ relative to $\theta_j$ by using $\theta_0 = \bar{\theta}_0 - \theta_j$. For example, if the preferred direction is west, then for segment AB,$\bar{\theta}_0 = 90°$ and $\theta_0 = 90° - 35° = 55°$ (when the migration rate is calculated from patch 1 to 2). We note that for the migration rate from patch 2 to 1, the outward normal would be $\theta_j + 180°$. Table 2.5 gives the values for $\theta_j$ for each boundary segment (for rates from patches 1 to 2 and 2 to 3).

The other migration rates are calculated similarly. We note here that we will use the same calculation for $b_9$ as in section 2.5.1 (assuming the highest range of dispersal).

To determine how preferred direction affects migration, we will calculate the varying cross-patch migration rates as $\theta_0$ ranges from 0 to $2\pi$ and compare with the migration rates for no preference for direction. If we consider the migration rate as a function of $\theta_0$, the function is even, with a maximum value when the preferred direction is out of the patch ($\theta_0 = 0$) and a minimum value when the preferred direction is toward the inside of the patch from which migration originates ($\theta_0 = \pi$). An example of this result can be seen in Figure 2.8. Figure 2.8 shows the migration rate from patch 1 to 2 for $e = 0.5$ and as $\theta_0$ varies from 0 to $2\pi$, compared with the migration rate from patch 1 to 2 assuming no preference for direction.

We also compare the rates for $e = 0.1$ and $e = 0.5$ as seen in Figure 2.8 and Figure 2.9. In Figure 2.8, we observe that if the preferred direction is closer to being

directly out of the origination patch, the increase in migration rate (compared with no preference of direction) is greater than the decrease if the preferred direction of migration is directed into the origination patch. However, we notice that for lower values of $e$, as represented in Figure 2.9, this effect is not as apparent. In fact, for $e = 0.1$, the amplitude between the highest migration rate (at $\theta_0 = 0$) and the rate with no preferred direction is only 1.2 times the amplitude between the lowest migration rate (at $\theta_0 = \pi$) and the rate with no preferred direction, while for $e = 0.5$, the amplitude between the highest rate and rate with no preferred direction is 2.5 times higher than the amplitude between the lowest rate and the rate with no preferred direction. Thus, for weaker degree of preference, the amplitude of the effect of the preferred direction becomes closer to being symmetric with respect to direction out of the origination patch versus direction into the origination patch.

A summary of migration rates for varying preferred directions with $e = 0.5$ is given in Tables 2.6 and 2.7. As expected, the migration rates are highest when the direction of migration is directly into the destination patch. For example, the

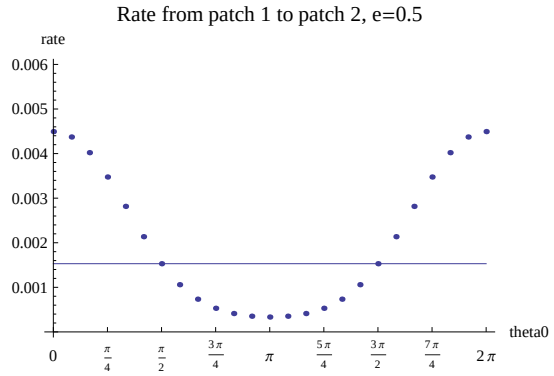


Figure 2.8. Migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line) ($e = 0.5$).

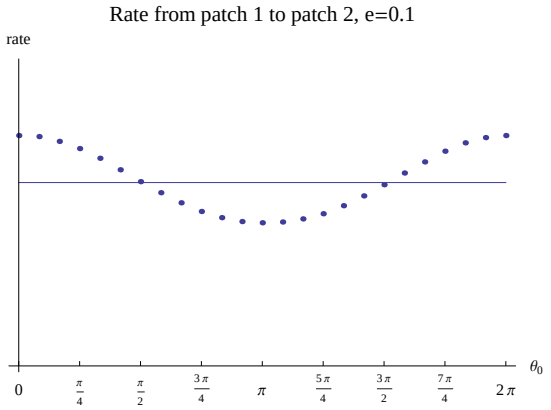


Figure 2.9. Migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line)($e = 0.1$).

highest migration rate from patch 2 to 3 is when the preferred direction is east, and the highest migration rate from patch 3 to 2 is when the preferred direction is west.

Table 2.6. Migration rates for $e = 0.5$, Patches 2/3

| Pref. direc. | $\bar{m}_{23}$ | $\bar{m}_{32}$ |
|---|---|---|
| No preference | 0.00101 | 0.000155 |
| (N) | 0.00116 | 0.000165 |
| (NW) | 0.000415 | 0.000373 |
| (W) | 0.000240 | 0.000497 |
| (SW) | 0.000344 | 0.000408 |
| (S) | 0.000935 | 0.000190 |
| (SE) | 0.00228 | 0.0000721 |
| (E) | 0.00324 | 0.0000405 |
| (NE) | 0.00260 | 0.0000599 |

Table 2.7. Migration rates for $e = 0.5$, Patches 1/2

| Pref. direc. | $\bar{m}_{12}$ | $\bar{m}_{21}$ |
|---|---|---|
| No preference | 0.00427 | 0.00385 |
| (N) | 0.0111 | 0.000941 |
| (NW) | 0.0107 | 0.000976 |
| (W) | 0.00685 | 0.00220 |
| (SW) | 0.00267 | 0.00622 |
| (S) | 0.00114 | 0.0102 |
| (SE) | 0.00121 | 0.00997 |
| (E) | 0.00304 | 0.000547 |
| (NE) | 0.00745 | 0.00204 |

## 2.7 Conclusions

The calculations in this chapter rely heavily on experimental and field studies on *Triatoma* vectors. Because there are virtually no studies on the dispersal capabilities on the North American species focused on in this study, *T. gerstaeckeri* and *T. sanguisuga*, we utilize studies on the more well known vectors from South America, *T. infestans* and *T. sordida*. Since *Triatoma* vectors are sensitive to temperature and humidity for flight initiation [21], we acknowledge the difference in climate between the South American regions and the patches considered here. Because of the lack of data on the North American vectors, we note the need for experimental studies regarding distance and frequency of vector dispersal.

We assume vector dispersal is based on three inherent properties: dispersal distance, preferred direction of dispersal, and degree of preference for direction. The

parameters describing these properties have different effects on the migration rate. Increasing dispersal distance will increase the migration rate, while varying the preferred direction may increase or decrease the rate depending on the actual geographical location of the origination and destination patch and the actual preferred direction (into or out of the origination patch). We observe asymmetry in the migration rates depending on whether the preferred direction is out of or in toward the origination patch. As the degree of preference increases, the increase in migration rate if preferred direction is out of the patch is greater than the decrease in migration rate if the preferred direction is into the patch, when compared with the rate with no preferred direction. This means that preference in direction tends to increase migration rates overall, since preference increases migration (patch-crossing) on the preferred side more than it decreases it on the opposite, non-preferred side of a patch.

One of the primary goals of population modeling is to describe how individual events on a local scale build into collective effects on a larger scale, with much study having been made of the resulting emergent properties of systems; dispersal of infectious individuals has often been studied in continuous space in order to capture spatially small-scale movement, but here we develop a framework for incorporating such small-scale dispersal into simpler, spatially discrete models that take a more global perspective in defining populations by common epidemiological and ecological characteristics. The framework established in this study allows us to translate local dispersal to global migration across regions, which may be applicable to other types of migrating species if dispersal distance capabilities are known. Migration rates may be calculated under the assumption of no preferred direction of dispersal or an assumed (or known) preference for a particular direction. The methodology provides a way to generate migration rates to be used in other epidemiological or ecological models. More specifically, there are applications to vector-borne diseases, for which

climate change has made the invasion of vector into new territories a major public health concern.

Other mathematical studies focusing on migration of *T. cruzi* vectors [79, 4] concentrate on migration over a small scale (a village) into domestic areas, whereas here we derive a method to calculate the rate of migration by connecting local dispersal to migration over a large scale. Large scale migration should be considered because of the spread of *T. cruzi* across the Americas through sylvatic settings. Generalizations to the model could include variation in habitat suitability, where different migration rates could be calculated for species traveling from a region of higher habitat preference (suitability) to a region of lower habitat suitability or vice versa.

In the next chapter, we apply these rates to study the effects of migration of *T. cruzi* vectors across patch boundaries model of three distinct transmission cycles of *T. cruzi* in northern Mexico and southeastern United States.

# CHAPTER 3

# A METAPOPULATION MODEL FOR SYLVATIC *T. CRUZI* TRANSMISSION WITH VECTOR MIGRATION

## 3.1 Introduction

Based on Chapter 2, we have an understanding of how local vector dispersal can be described in terms of global effects, so that we may now consider a model that describes sylvatic cycles of *T. cruzi* over a large geographic area. Because *T. cruzi* is maintained in sylvatic cycles, we recognize the need to study the spatial spread of the disease, especially in North America, where risk of Chagas has only recently been discussed. In the United States, sylvatic hosts are commonly associated with *T. cruzi* strain type I and IIa [69], with type IIa recently reclassified as type IV [96]. Here we investigate several models of *T. cruzi*, incorporating multiple modes of transmission and multiple patches. More specifically, we wish to focus our efforts on the effects of vector migration on sylvatic *T. cruzi* strain type IV transmission in two different North American host-vector cycles.

A vector-borne disease is most commonly spread between hosts indirectly through the bite of an insect vector. Many vector borne diseases have been studied using mathematical models, including the familiar Ross model for malaria [71], later adapted by Macdonald in 1957 [51]. The Ross-Macdonald model is an SI model using differential equations including susceptible and infected humans and mosquitoes. Since the discovery of Chagas disease in 1909, mathematicians have been modeling its spread in humans, animals, and vectors. To date, the majority of mathematical models for Chagas disease have been studied in humans and vectors, rather than the animal

hosts. Velasco-Hernández [88] modeled Chagas in humans using a model structure similar to the Ross-Macdonald malaria model, but included another infectious compartment for chronically ill humans. Since infection with *T. cruzi* is maintained in reservoir (sylvatic) hosts and human transmission cycles cannot be sustained without them [32], recently more attention has been given to the spread of the *T. cruzi* parasite in *Triatoma* vectors and associated animal hosts [43, 45]. In each model, Kribs uses a deterministic SI model with one host and one vector to study the effects of alternative transmission modes for *T. cruzi*, namely vector consumption by animal hosts and vertical transmission in hosts. Results show that vertical transmission is not enough to maintain the infection cycle alone, but vertical transmission along with even an inefficient host-vector transmission cycle can sustain the *T. cruzi* infection cycle. Due to the nature of transmission of vector-borne disease, in which vectors and hosts (especially in sylvatic settings) may be easily affected by weather (mainly temperature and humidity) and landscape, spatial spread is a key element in studying a vector-borne disease.

Spatial spread of a disease can be modeled using continuous or discrete space. The majority of mathematical models involving the spatial spread of infectious diseases in continuous time and space are modeled using reaction-diffusion systems taking the form of a system of partial differential equations. Some studies incorporating spatial spread are the spread of rabies in the fox population [60], and the vector borne diseases dengue [52] and West Nile virus [49]. The results of such systems are generally described using traveling waves which describe the process of the spread of the disease, most often over a homogeneous landscape. In each model, movement of either hosts or vectors (or both) is considered, with the underlying assumption that the movement is random.

Other types of models incorporating spatial spread include multi-patch metapopulation models in which movement occurs between $n$ patches. Several models have been studied, including a multi-species model by Arino et al. [3], in which analytical results are given for several multi-species, multi-patch models. In the study, a formula is derived for the basic reproductive number $R_0$ for multiple species and multiple patches and global stability for the disease-free equilibrium is established for $R_0 < 1$. Allen et al. [2] gives a 3-patch model of hantavirus spread in reservoir and spillover species in which the outer patches represent the preferred habitat of the reservoir and spillover species and the middle patch represents the boundary region in which the species overlap. We note here that the overlap region was temporally- and spatially-dependent. The movement here is described in terms of number of visits per year to the boundary region and length of time spent there. Reproductive numbers for each patch were calculated, and it was determined that the greater number of interactions among species caused the reproductive number of the overlap patch to exceed the reproductive numbers of the patches representing preferred habitat, thereby causing a greater possibility of disease persistence.

The vectors in this study are two different species of triatomine bugs commonly found in southern and south-eastern U.S. and Mexico, *Triatoma gerstaeckeri*, from the *phyllosoma* complex and *Triatoma sanguisuga* of the *lecticularia* complex [72, 65, 35] . The primary hosts associated with these vectors are raccoons (*Procyon lotor*) in the southeastern United States and woodrats (*Neotoma micropus*) in southern Texas and Mexico. In this study we consider three different geographical areas we refer to as patches. Each patch is described by a distinct transmission cycle between the hosts and vectors in the model. We consider that one species of vector and host is in patch 1 and 2, while the other species of vector and host is in patches 2 and 3. Patch 1 is defined by the *T. gerstaeckeri*-woodrat infection cycle in northern Mexico

and southern Texas. Patch 3 is the south-eastern United States, including parts of Texas, Louisiana, Mississippi, Alabama, Georgia, and portions of the Carolinas and Florida panhandle and is defined by the *T. sanguisuga*-raccoon infection cycle, while patch 2 includes the south to southwest parts of Texas and a part of Coahuila, Mexico where we consider both species of vector and host to overlap. We assume the communication between patches and between cycles in patch 2 occurs through the movement of *Triatoma* vectors. *T. sanguisuga* migrates between patches 2 and 3, and between raccoons and woodrats in patch 2, while *T. gerstaeckeri* migrates between patches 1 and 2. We wish to investigate how increased vector migration affects prevalence in the overlap patch compared to the single-cycle patches.

In this study we develop and analyze an *S-I* metapopulation model with the aforementioned hosts and vectors, in which hosts may exhibit vertical transmission and vectors migrate between patches. We carry out standard analysis techniques, such as calculating $R_0$ for various sub-models of the larger metapopulation model as well as determining existence of endemic equilibria analytically (when tractable) in order to see effects of vector migration on $R_0$. Finally, we perform numerical analysis on the full model to determine effects of migration on prevalence of *T. cruzi*.

## 3.2 Problem formulation and model

### 3.2.1 Problem statement and assumptions

To establish stable large-scale demographics for hosts and vectors, we assume that the growth for each species will be logistic, and neither the hosts nor vectors identified in the model exhibit disease-induced mortality. Studies have shown that the *T. cruzi* infecting raccoons and opossums in areas in the south-southeastern U.S. are not pathogenic, and do not appear to cause any symptoms of Chagas [64, 93].

Although triatomine bugs feed on many hosts in the wild, we consider woodrats and raccoons based on data that correlates their geographic location very closely with the vectors [65, 13, 66, 40]. In literature reviewed here, we have found the only host associated with *T. gerstaeckeri* is the southern plains woodrat [41, 27]. Therefore, we consider that the southern plains woodrat is the preferred host for the vector *T. gerstaeckeri*.

In this model there are several infection rates to be considered. These infection rates may differ from vector to host and host to vector as well as by patch (geographical region). Biologically, the rate of infection between the hosts and vectors in each patch should be different. However, as seen in section 3.4 when calculating numerical estimates for the infection rate parameters (using a procedure called back-calculation), it is mathematically necessary to keep some of the rates the same. Thus, we assume that the rate of infection from raccoons to *T. sanguisuga* is the same in patches 2 and 3 and the rate of infection from *T. gerstaeckeri* to woodrats is the same in patches 1 and 2.

Literature suggests that *T. sanguisuga* will feed on other hosts besides raccoons, namely woodrats [27]. Thus, in patch 2, *T. sanguisuga* feeds on raccoons and woodrats. Some proportion of vector-woodrat contacts are made with *T. sanguisuga*, thus we define $q_W$ as the proportion of vector-woodrat contacts made with *T. sanguisuga*, while $1-q_W$ is the proportion made with *T. gerstaeckeri*. Furthermore, since *T. sanguisuga* feeds on both hosts, we define $q_S$ as the proportion of *T. sanguisuga* contacts made with raccoons, while $1-q_S$ is the proportion made with woodrats. We denote the per vector infection rate from *T. sanguisuga* to woodrats as $\beta_{SW}$ and the per vector infection rate from woodrats to *T. sanguisuga* to be $\beta_{WS}$.

*T. cruzi* has been confirmed to be transmitted vertically among mice in laboratory conditions [59]. However, there is limited data on vertical transmission of *T.*

*cruzi* in sylvatic hosts. We will assume that *T. cruzi* can be transmitted vertically in raccoons and woodrats.

Infection contact rates could be limited by the host or vector population. For our model, we assume that the limiting factor for infection from host to vector will be the vector population. The hosts are plentiful enough for vectors to feed as much as desired. Therefore, the vector population density will be the driving force in determining the infection rate, and the contact process saturates more quickly in the vector population than in the hosts. Thus, we consider the infection term from host to vector in the model to be based on several factors. Using similar derivation as in [44], we define the per-vector biting rate as $z$ (in units of contacts per vector per time), and thus the total vector-feeding contact rate as $z \cdot N_v$ (with units of bites/time). Thus, to calculate the rate of new vector infections, we multiply the total vector-feeding contact rate by the proportion of contacts that involve uninfected vectors and infected hosts, multiplied by the proportion of contacts that result in an infection ($\pi_v$) (units of infected vectors/bite) obtaining

$$zN_v \cdot \frac{S_v}{N_v} \cdot \frac{I_h}{N_h} \cdot \pi_v = (\pi_v z)\frac{I_h}{N_h} S_v.$$

To simplify, we write $\beta_h = \pi_v z$ (in units of 1/time).

We also need to describe the rate at which vectors infect hosts. Again, assuming that the vector to host infection will be limited by the vector population (vectors feeding as frequently as desired), we multiply the total vector-feeding contact rate by the proportion of contacts involving uninfected hosts and infected vectors, multiplied by the proportion of contacts resulting in an infection (with units infected hosts/bite),

$$zN_v \cdot \frac{S_h}{N_h} \cdot \frac{I_v}{N_v} \cdot \pi_h = (\pi_h z)\frac{I_v}{N_h} S_h.$$

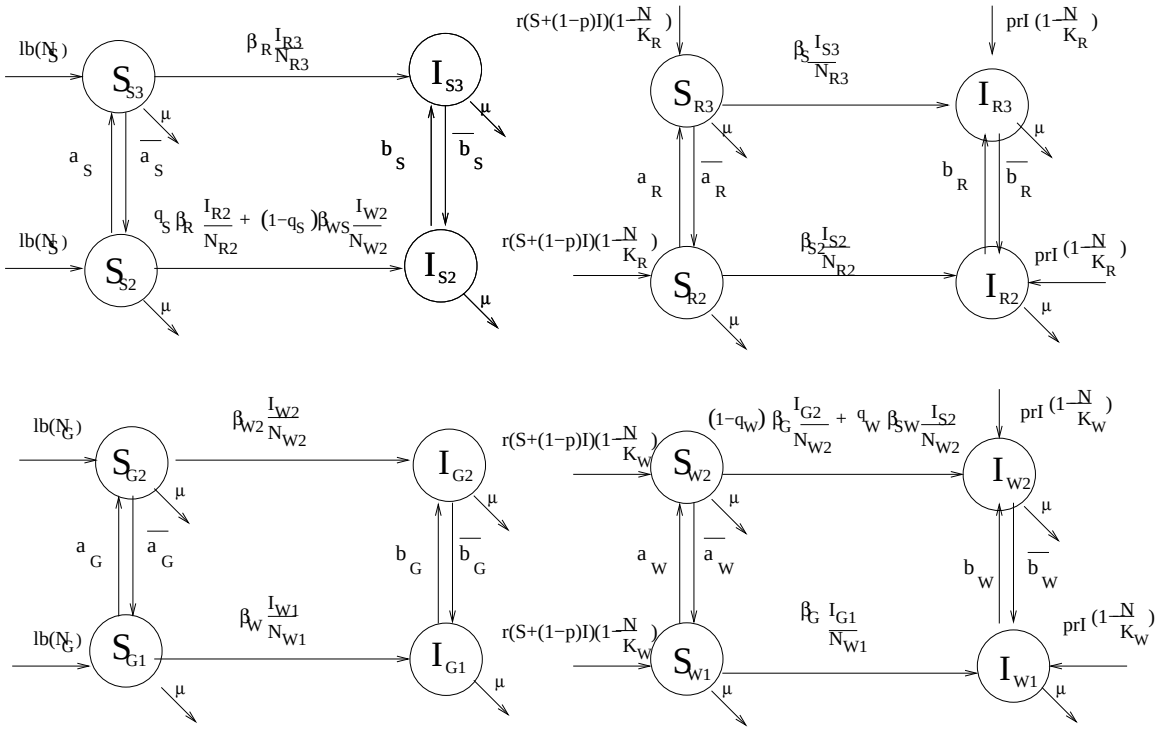


Figure 3.1. Model. The migration rates represent outgoing rates which must be adjusted by the patch area ratios for incoming rates (see system (3.1)).

In a similar manner, we define $\beta_v = \pi_h z$. But, we note here that $\beta_v$ is not in units of 1/time, but rather infected hosts per vector per time. We will apply this assumption to the infection terms in the model using the appropriate vector and host subscripts.

### 3.2.2 The Model

The model presented here is an *S-I* model incorporating migration and vertical transmission.

For each species we will use the general logistic birth rate $b(N) = rN\left(1 - \frac{N}{K}\right)$, where $r$ represents the intrinsic population growth rate, $N$ represents the total population density, and $K$ is the carrying capacity of the population density. We will use this term for each vector and host population using the appropriate subscripts. We further

mention that the hosts in the model exhibit vertical transmission. Therefore, we incorporate this into the model by defining the following functions: $f(N, I)$ and $g(N, I)$ represent the birth rates for the hosts exhibiting vertical transmission. If only a proportion $p$ $(0 < p < 1)$ of infected hosts transmit vertically, then $g(N, I) = pI \cdot r\left(1 - \frac{N}{K}\right)$, where $p$ is the proportion of the offspring of infected hosts born infected with *T. cruzi*. Thus $f(N, I) = (S + (1 - p)I) \cdot r\left(1 - \frac{N}{K}\right)$. We will apply this assumption to both hosts in the model, applying the appropriate subscripts. The natural per host and per vector mortality rates are denoted by $\mu_R, \mu_W, \mu_S$, and $\mu_G$. In considering the behavior of the total vector and host populations, it is true that the intrinsic population growth rate is greater than the mortality rate $(r > \mu)$. If $r < \mu$, the vector and host populations would go extinct. Furthermore, for this model we will assume linear migration based on the idea of local dispersal. We note that there are several different migration rates considered. We assume that hosts and vectors move at different rates, and those rates differ by species, by infection status, and by direction of migration. The migration parameters in the model are denoted by $b_i$ or $\bar{b}_i$ for vectors and $a_i$ or $\bar{a}_i$ for hosts, designating difference in direction of migration. Each subscript, $i$, is used to designate the migrating species, R, S, G, or, W. For example, $\bar{b}_G$ represents the migration rate of *T. gerstaeckeri* from patch 1 to 2, while $b_G$ is the rate of *T. gerstaeckeri* from patch 2 to 1. Since migration is generally small compared to demographic processes, we further assume that $r > \mu + b_i$ and that $\mu > b_i$, and apply this assumption uniformly among all species.

Furthermore, we note here that each patch has a different area, and the migration rates derived in Chapter 2 are affected by the size of the patch from which migration originates. In the metapopulation model described here, each differential equation represents the change in population density over time, thus the size of patch must also affect the population density of the vectors and hosts in each patch. Since

we are considering migration of vectors between patches, the differences in population densities must be accounted for. For, example in order to account for the differing patch sizes, we consider the equation for the absolute number of infected *T. sanguisuga* vectors in patch 3, where $A_3$ denotes the area (in $m^2$) of patch 3 and the state variables have units of density. Then, the absolute number of vectors in patch 3 is given by

$$(I_{S3}A_3)' = \beta_R \frac{I_{R3}}{N_{R3}} S_{S3}A_3 - (\mu_S + b_S)I_{S3}A_3 + \bar{b}_S I_{S2}A_2.$$

It follows that

$$I'_{S3} = \beta_R \frac{I_{R3}}{N_{R3}} S_{S3} - (\mu_S + b_S)I_{S3} + \bar{b}_S \frac{A_2}{A_3} I_{S2}.$$

In the model, we write $\bar{b}_S \frac{A_2}{A_3} = \tilde{b}_S$ and $b_S \frac{A_3}{A_2} = \hat{b}_S$. This notation will be used more generally for northward and southward migration rates, respectively. We make an implicit assumption that the migration rates are low enough to not cause equilibrium population densities in any patch to exceed the region's carrying capacity. If this assumption is violated, and the carrying capacities are exceeded, the equations would need to be adjusted to distribute the resulting negative logistic term (as a result of additional density dependent deaths) proportionally between the susceptibles and infectives.

We therefore derive the model given below and seen in Figure 3.1.

$$
\begin{aligned}
S'_{S3} &= r_S\left(1-\frac{N_{S3}}{K_{S3}}\right)N_{S3} - \beta_R\frac{I_{R3}}{N_{R3}}S_{S3} - (\mu_S+a_S)S_{S3} + \tilde{a}_S S_{S2}\\
I'_{S3} &= \beta_R\frac{I_{R3}}{N_{R3}}S_{S3} - (\mu_S+b_S)I_{S3} + \tilde{b}_S I_{S2}\\
S'_{R3} &= r_R\left(S_{R3}+(1-p_R)I_{R3}\right)\left(1-\frac{N_{R3}}{K_{R3}}\right) - \beta_S\frac{I_{S3}}{N_{R3}}S_{R3} - (\mu_R+a_R)S_{R3} + \tilde{a}_R S_{R2}\\
I'_{R3} &= p_R r_R I_{R3}\left(1-\frac{N_{R3}}{K_{R3}}\right) + \beta_S\frac{I_{S3}}{N_{R3}}S_{R3} - (\mu_R+b_R)I_{R3} + \tilde{b}_R I_{R2}\\
S'_{S2} &= r_S\left(1-\frac{N_{S2}}{K_{S2}}\right)N_{S2} - \left(q_S\beta_R\frac{I_{R2}}{N_{R2}} + (1-q_S)\beta_{WS}\frac{I_{W2}}{N_{W2}}\right)S_{S2} - (\mu_S+\bar{a}_S)S_{S2} + \hat{a}_S S_{S3}\\
I'_{S2} &= \left(q_S\beta_R\frac{I_{R2}}{N_{R2}} + (1-q_S)\beta_{WS}\frac{I_{W2}}{N_{W2}}\right)S_{S2} - (\mu_S+\bar{b}_S)I_{S2} + \hat{b}_S I_{S3}\\
S'_{R2} &= r_R\left(S_{R2}+(1-p_R)I_{R2}\right)\left(1-\frac{N_{R2}}{K_{R2}}\right) - \beta_{S2}\frac{I_{S2}}{N_{R2}}S_{R2} - (\mu_R+\bar{a}_R)S_{R2} + \hat{a}_R S_{R3}\\
I'_{R2} &= p_R r_R I_{R2}\left(1-\frac{N_{R2}}{K_{R2}}\right) + \beta_{S2}\frac{I_{S2}}{N_{R2}}S_{R2} - (\mu_R+\bar{b}_R)I_{R2} + \hat{b}_R I_{R3}\\
S'_{G2} &= r_G\left(1-\frac{N_{G2}}{K_{G2}}\right)N_{G2} - \beta_{W2}\frac{I_{W2}}{N_{W2}}S_{G2} - (\mu_G+a_G)S_{G2} + \tilde{a}_G S_{G1}\\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N_{W2}}S_{G2} - (\mu_G+b_G)I_{G2} + \tilde{b}_G I_{G1}\\
S'_{W2} &= r_W\left(S_{W2}+(1-p_W)I_{W2}\right)\left(1-\frac{N_{W2}}{K_{W2}}\right) - \left((1-q_W)\beta_G\frac{I_{G2}}{N_{W2}} + q_W\beta_{SW}\frac{I_{S2}}{N_{W2}}\right)S_{W2}\\
&\quad - (\mu_W+a_W)S_{W2} + \tilde{a}_W S_{W1}\\
I'_{W2} &= p_W r_W I_{W2}\left(1-\frac{N_{W2}}{K_{W2}}\right) + \left((1-q_W)\beta_G\frac{I_{G2}}{N_{W2}} + q_W\beta_{SW}\frac{I_{S2}}{N_{W2}}\right)S_{W2}\\
&\quad - (\mu_W+b_W)I_{W2} + \tilde{b}_W I_{W1}\\
S'_{G1} &= r_G\left(1-\frac{N_{G1}}{K_{G1}}\right)N_{G1} - \beta_W\frac{I_{W1}}{N_{W1}}S_{G1} - (\mu_G+\bar{a}_G)S_{G1} + \hat{a}_G S_{G2}\\
I'_{G1} &= \beta_W\frac{I_{W1}}{N_{W1}}S_{G1} - (\mu_G+\bar{b}_G)I_{G1} + \hat{b}_G I_{G2}\\
S'_{W1} &= r_W\left(S_{W1}+(1-p_W)I_{W1}\right)\left(1-\frac{N_{W1}}{K_{W1}}\right) - \beta_G\frac{I_{G1}}{N_{W1}}S_{W1} - (\mu_W+\bar{a}_W)S_{W1} + \hat{a}_W S_{W2}\\
I'_{W1} &= p_W r_W I_{W1}\left(1-\frac{N_{W1}}{K_{W1}}\right) + \beta_G\frac{I_{G1}}{N_{W1}}S_{W1}\\
&\quad - (\mu_W+\bar{b}_W)I_{W1} + \hat{b}_W I_{W2}
\end{aligned}
\tag{3.1}
$$

## 3.3 Analysis

Little information is known about the migration of the vectors. Thus, we will explore several hypotheses. We will first consider that infected vectors are the only species to migrate. In a study done by Añez and East [1] in 1984 on the effect of parasites on the behavior of the vector, *Rhodnius prolixus*, it was shown that that the parasite, *Trypanosoma rangeli*, hindered the vector's ability to draw blood, thus causing the bug to bite 25 more times than an uninfected vector. Thus, differential behavior of vectors infected with *T. cruzi* may affect the transmission of the parasite as well as vector mobility as mentioned in [43]. In this context, we will consider the possibility that infected vectors move in only one direction (towards more preferred climates) and the possibility that infected vectors move between patches at different rates for different directions. Furthermore, we consider that uninfected vectors also migrate, but at a rate proportional to that of infected vectors. We will not consider host migration to play a significant role. The hosts in the model are bound by habitat constraints, and thus by definition of the patches, we assume the hosts are not likely to move.

In order to get a better understanding of the full model, several special cases will be considered. The main identifying characteristics of the model are vertical transmission, migration, and multiple hosts and vectors.

### 3.3.1 One patch, one host, one vector, no vertical transmission

We begin the analysis of (3.1) by studying the simple system with one host and one vector. By observing system (3.1), we see that when the migration terms $a_i = \bar{a}_i = b_j = \bar{b}_j = 0$, for $i = R, W$, $j = S, G$, the three patches decouple. In this scenario, patch 1 and patch 3 are identical in structure. Thus, analyzing patch 1 and

3 with migration terms set to 0 and $p_W = p_R = 0$, we are analyzing the simple one host-one vector system.

$$\begin{aligned} S_h' &= r_h N_h \left(1 - \frac{N_h}{K_h}\right) - \beta_v \frac{I_v}{N_h} S_h - \mu_h S_h \\ I_h' &= \beta_v \frac{I_v}{N_h} S_h - \mu_h I_h \\ S_v' &= r_v N_v \left(1 - \frac{N_v}{K_v}\right) - \beta_h \frac{I_h}{N_h} S_v - \mu_v S_v \\ I_v' &= \beta_h \frac{I_h}{N_h} S_v - \mu_v I_v \end{aligned} \tag{3.2}$$

Because the vector and host populations always approach an equilibrium, we can consider the limiting system in which $N_h$ and $N_v$ have reached their positive equilibria, $N_h^*$ and $N_v^*$, where

$$N_h^* = K_h \left(1 - \frac{\mu_h}{r_h}\right), \quad N_v^* = K_v \left(1 - \frac{\mu_v}{r_v}\right).$$

As mentioned previously, for each species, we assume $r > \mu$, which guarantees that all disease-free extinction equilibria are unstable. This assumption will be carried out in this and all of the models hereafter. In system (3.2), we define $\tilde{\beta}_v = \beta_v \frac{N_v^*}{N_h^*}$. Results by Thieme [81, 82] guarantee that the behavior of the full system is asymptotic to the limiting system which is given by

$$\begin{aligned} I_h' &= \tilde{\beta}_v \frac{I_v}{N_v^*}(N_h^* - I_h) - \mu_h I_h \\ I_v' &= \beta_h \frac{I_h}{N_h^*}(N_v^* - I_v) - \mu_v I_v \end{aligned} \tag{3.3}$$

This model has been well studied [71, 10] and we give results here. The basic reproductive number, calculated using the next-generation matrix [85], is given by

$$R_0 = \sqrt{\frac{\beta_h \tilde{\beta}_v}{\mu_h \mu_v}},$$

which represents the average number of secondary infections caused by an infected individual introduced into a susceptible population. Because of the vector-host dynamics, $R_0$ represents the geometric mean between the average number of secondary host infections caused by one vector, and the average number of vector infections caused by one host.

In the case when $R_0 < 1$, the disease will die out and the population will approach the disease free equilibrium. When $R_0 > 1$, the population will approach a unique endemic state,

$$\frac{I_v^*}{N_v^*} = \frac{\beta_h\tilde{\beta}_v - \mu_h\mu_v}{\tilde{\beta}_v\beta_h + \tilde{\beta}_v\mu_v} = \frac{R_0^2 - 1}{R_0^2 + \frac{\tilde{\beta}_v}{\mu_h}},$$

$$\frac{I_h^*}{N_h^*} = \frac{\beta_h\tilde{\beta}_v - \mu_h\mu_v}{\tilde{\beta}_v\beta_h + \beta_h\mu_h} = \frac{R_0^2 - 1}{R_0^2 + \frac{\beta_h}{\mu_v}}.$$

### 3.3.2 Patch 2, no migration, no vertical transmission

We will further analyze the decoupled system (3.1) by considering the equations representing patch 2 alone with no vertical transmission or migration. Thus, we analyze system (3.1) with $a_i = \bar{a}_i = 0$, for $i = R, W$, $b_j = \bar{b}_j = 0$, for $j = S, G$, and $p_R = p_W = 0$. The quantities $N_{R2}, N_{S2}, N_{W2}$, and $N_{G2}$ are asymptotically constant to $N_{R2}^*, N_{S2}^*, N_{W2}^*$, and $N_{G2}^*$, respectively, where

$$N_{S2}^* = K_{S2}\left(1 - \frac{\mu_S}{r_S}\right), \quad N_{R2}^* = K_{R2}\left(1 - \frac{\mu_R}{r_R}\right),$$

$$N_{G2}^* = K_{G2}\left(1 - \frac{\mu_G}{r_G}\right), \quad N_{W2}^* = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right).$$

Applying the results from Thieme [81, 82], we can study the limiting system

$$
\begin{aligned}
I'_{S2} &= \left(q_S\beta_R\frac{I_{R2}}{N^*_{R2}} + (1-q_S)\beta_{WS}\frac{I_{W2}}{N^*_{W2}}\right)(N^*_{S2} - I_{S2}) - \mu_S I_{S2},\\
I'_{R2} &= \beta_{S2}\frac{I_{S2}}{N^*_{R2}}(N^*_{R2} - I_{R2}) - \mu_R I_{R2},\\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N^*_{W2}}(N^*_{G2} - I_{G2}) - \mu_G I_{G2},\\
I'_{W2} &= \left((1-q_W)\beta_G\frac{I_{G2}}{N^*_{W2}} + q_W\beta_{SW}\frac{I_{S2}}{N^*_{W2}}\right)(N^*_{W2} - I_{W2}) - \mu_W I_{W2}.
\end{aligned}
\tag{3.4}
$$

In analyzing system (3.4) we observe the disease-free equilibrium is $(I^*_{S2}, I^*_{R2}, I^*_{G2}, I^*_{W2}) = (0,0,0,0)$. We find the basic reproductive number $R_0$ for the system using the next generation matrix method [85]. The work can be seen in Appendix B.1. We determine that

$$R_0 = \sqrt{\frac{1}{2}\left(P + \sqrt{P^2 - 4Q}\right)},$$

where

$$
\begin{aligned}
&P = f_1 + f_2 + f_3, \;\; Q = f_1 f_3,\\
&f_1 = \frac{(1-q_W)\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}, \;\; f_2 = \frac{q_W\beta_{SW} + (1-q_S)\beta_{WS}}{\mu_S\mu_W}\frac{N^*_{S2}}{N^*_{W2}}, \;\; f_3 = \frac{q_S\beta_R\beta_{S2}}{\mu_R\mu_S}\frac{N^*_{S2}}{N^*_{R2}}.
\end{aligned}
\tag{3.5}
$$

In the terms for $R_0$, we see that $f_1$ represents the *T. gerstaeckeri*-woodrat transmission cycle, $f_2$ represents the *T. sanguisuga*-woodrat cycle, and $f_3$ represents the *T. sanguisuga*-raccoon cycle. We further observe that

$$\max\{\sqrt{f_1}, \sqrt{f_2}, \sqrt{f_3}\} < R_0 < \sqrt{f_1 + f_2 + f_3}.$$

At this point, we wish to observe the importance of the overlap of the transmission cycles between *T. sanguisuga* and *T. gerstaeckeri*, and how this overlap affects the ability of the infection to be spread. If there is no overlap, i.e. no *T. sanguisuga*-woodrat cycle, then $f_2 = 0$, so then we may define $\bar{R}_0 = \sqrt{\max\{f_1, f_3\}} < R_0$. Since

$\bar{R}_0$ is always less than $R_0$, we observe the effect of $f_2$ is to increase the value of $R_0$. It is possible that $f_1 < 1$ and $f_3 < 1$, yet $R_0 > 1$. Furthermore, it is also possible that $f_1 + f_3 < 1$, yet $R_0 > 1$.

By investigating the equilibrium conditions, it can be shown that either one or three endemic equilibrium values exist when $R_0 > 1$. The computations can be seen in Appendix B.1.

### 3.3.3 Patches 1 and 2, 1 host 1 vector with vertical transmission and unidirectional migration of infected vectors

In dealing with patches 1 and 2, there are several cases to be considered. We will first consider the scenario with one host and one vector, with vertical transmission, and unidirectional migration of infected vectors. In this case, $N_{R2} = N_{S2} = 0$, $q_W = q_S = 0$, and $a_W = \bar{a}_W = b_G = 0$. Because the woodrat population is asymptotically constant with $N^*_{W1} = K_{W1}\left(1 - \frac{\mu_W}{r_W}\right)$, $N^*_{W2} = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right)$, we can apply Thieme's results [81, 82] and rewrite $I'_{W1}$ and $I'_{W2}$, passing $N_{W1}$ and $N_{W2}$ to their limiting values, $N^*_{W1}$ and $N^*_{W2}$. The system therefore simplifies to

$$\begin{aligned}
N'_{G2} &= r_G N_{G2}\left(1 - \frac{N_{G2}}{K_{G2}}\right) - \mu_G N_{G2} + \tilde{b}_G I_{G1} \\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N^*_{W2}}(N_{G2} - I_{G2}) - \mu_G I_{G2} + \tilde{b}_G I_{G1} \\
I'_{W2} &= p_W \mu_W I_{W2} + \beta_G \frac{I_{G2}}{N^*_{W2}}(N^*_{W2} - I_{W2}) - \mu_W I_{W2} \\
N'_{G1} &= r_G N_{G1}\left(1 - \frac{N_{G1}}{K_{G1}}\right) - \mu_G N_{G1} - \bar{b}_G I_{G1} \\
I'_{G1} &= \beta_W \frac{I_{W1}}{N^*_{W1}}(N_{G1} - I_{G1}) - \mu_G I_{G1} - \bar{b}_G I_{G1} \\
I'_{W1} &= p_W \mu_W I_{W1} + \beta_G \frac{I_{G1}}{N^*_{W1}}(N^*_{W1} - I_{W1}) - \mu_W I_{W1}
\end{aligned} \tag{3.6}$$

Cherif et al. [16] studied a similar model for *T. cruzi* vector transmission dynamics involving two strains (one being more virulent). In their model, a proportion of vectors infected with the more virulent strain migrate to a region in which the less virulent strain is native. However, their model did not include vertical transmission as system (3.6) does.

We determine the disease-free equilibrium for this system (3.6) to be $(N^*_{G2}, 0, 0, N^*_{G1}, 0, 0)$, where

$$N^*_{G2} = K_{G2}\left(1 - \frac{\mu_G}{r_G}\right),\ N^*_{G1} = K_{G1}\left(1 - \frac{\mu_G}{r_G}\right).$$

$R_0$ can be found via the next generation matrix. For system (3.6),$R_0$ is

$$\max\left\{\frac{1}{2}\left(p_W + \sqrt{4\frac{\beta_G\beta_W}{(\mu_G+\bar{b}_G)\mu_W}\frac{N^*_{G1}}{N^*_{W1}} + p_W^2}\right),\ \frac{1}{2}\left(p_W + \sqrt{4\frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}} + p_W^2}\right)\right\}.$$

If $\frac{N^*_{G1}}{N^*_{W1}} \leq \frac{N^*_{G2}}{N^*_{W2}}$ then the second term of $R_0$ is larger of the two because $\bar{b}_G > 0$. In this case $\max(p_W, \frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}) < R_2 < p_W + \frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}$.

To study possible endemic equilibria of system (3.6), we determine that

$$\bar{N}^*_{G2} = K_{G2}\left(1 - \frac{\mu_G - \tilde{b}_G x^*_{G1}}{r_G}\right),\ \bar{N}^*_{G1} = K_{G1}\left(1 - \frac{\mu_G + \bar{b}_G x^*_{G1}}{r_G}.\right)$$

After substituting these values into the equilibrium conditions for system (3.6) (seen in Appendix B), we determine existence of an endemic equilibrium when $R_0 > 1$. We further determine that precisely one endemic equilibria exist in patch 2 alone if and only if $R_1 < 1 < R_2$, and in both patches if $R_1 > 1$ (the patch 2 only endemic equilibrium is unstable in this case).

### 3.3.4 Patches 1 and 2, 1 host 1 vector with vertical transmission and unidirectional migration of all vectors

We may also consider the case in which uninfected vectors move at a reduced rate proportional to that of infected vectors. As mentioned previously, if infected

vectors exhibit differential behavior causing them to migrate more than uninfected vectors, we would consider the effects of having uninfected vectors migrate as well, but at a reduced rate. We note that this scenario is as far as vertical transmission can be treated analytically regarding $R_0$. The vertical transmission terms in the model do not affect the complexity in computing endemic equilibria since $\mu$ is simply replaced with $(1-p)\mu$ in the infected host equations. Thus, in computing $R_0$, the remaining cases will be done without vertical transmission. We let $N_{R2} = N_{S2} = 0$, $q_W = q_S = 0$, $a_W = \bar{a}_W = 0$, and $b_G = \gamma \bar{b}_G$, where $0 < \gamma < 1$. Since $N_{W1}$ and $N_{W2}$ are asymptotically constant to $N^*_{W1} = K_{W1}\left(1 - \frac{\mu_W}{r_W}\right)$, $N^*_{W2} = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right)$, we apply the results of Thieme to this system, passing $N_{W1}$ and $N_{W2}$ to their limiting values, $N^*_{W1}$ and $N^*_{W2}$, and obtaining the following system,

$$\begin{aligned}
S'_{G2} &= r_G\left(1 - \frac{N_{G2}}{K_{G2}}\right)N_{G2} - \beta_{W2}\frac{I_{W2}}{N_{W2}}S_{G2} - \mu_G S_{G2} + \gamma\tilde{b}_G S_{G1} \\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N_{W2}}S_{G2} - \mu_G I_{G2} + \tilde{b}_G I_{G1} \\
I'_{W2} &= p_W\mu_W I_{W2} + \beta_G\frac{I_{G2}}{N^*_{W2}}(N^*_{W2} - I_{W2}) - \mu_W I_{W2} \\
S'_{G1} &= r_G\left(1 - \frac{N_{G1}}{K_{G1}}\right)N_{G1} - \beta_W\frac{I_{W1}}{N_{W1}}S_{G1} - \mu_G S_{G1} - \gamma\bar{b}_G S_{G1} \\
I'_{G1} &= \beta_W\frac{I_{W1}}{N_{W1}}S_{G1} - \mu_G I_{G1} - \bar{b}_G I_{G1} \\
I'_{W1} &= p_W\mu_W I_{W1} + \beta_G\frac{I_{G1}}{N^*_{W1}}(N^*_{W1} - I^*_{W1}) - \mu_W I_{W1}.
\end{aligned} \tag{3.7}$$

Here, we can determine the disease free equilibrium to be $(N^*_{G2}, 0, 0, N^*_{G1}, 0, 0)$, where

$$N^*_{G2} = K_{G2}\sqrt{\left(1 - \frac{\mu_G}{r_G}\right)^2 + 4\frac{\gamma\tilde{b}_G}{r_G}\frac{N^*_{G1}}{K_{G2}}\left(1 - \frac{\mu_G + \gamma\bar{b}_G}{r_G}\right)},$$

$$N^*_{G1} = K_{G1}\left(1 - \frac{\mu_G + \gamma\bar{b}_G}{r_G}\right).$$

Observing the terms of $N^*_{G2}$, we can see that the first term, $K_{G2}\left(1 - \frac{\mu_G}{r_G}\right)$, essentially represents the natural demographic renewal for the population of vectors

in patch 2, while the second term represents the population being brought from vectors in patch 1.

$R_0$ for the system is $R_0 = \max\{R_1, R_2\}$ where

$$R_1 = \frac{1}{2}\left(p_W + \sqrt{4\frac{\beta_G\beta_W}{(\mu_G + \bar{b_G})\mu_W}\frac{N^*_{G1}}{N^*_{W1}} + p_W^2}\right), \text{ and}$$

$$R_2 = \frac{1}{2}\left(p_W + \sqrt{4\frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}} + p_W^2}\right).$$

In general the form of $R_0$ for system (3.7) is the same as that of system (3.6), with a different disease free equilibrium for the vector population due to the unidirectional migration of all vectors. We see that the first term of $R_0$ represents the patch 1 dynamics, while the second term represents patch 2. Similar to the system in section 3.3.3, if $\frac{N^*_{G1}}{N^*_{W1}} \leq \frac{N^*_{G2}}{N^*_{W2}}$ then the second term of $R_0$ is larger.

Determining endemic equilibria for this system is intractable analytically. After a numerical investigation using the parameters estimated in section 3.4, we determine precisely one unique endemic equilibrium exists when $R_0 > 1$.

### 3.3.5 Patch 1 and 2, 1 host 1 vector, no vertical transmission, bidirectional migration of infected vectors

Another scenario we treat in patches 1 and 2 is one host, one vector and bidirectional migration of infected vectors. We have previously assumed that vectors may have a preferred direction so that our migration is unidirectional. However, we know that vectors will move in every direction (although one direction may be preferred over another); thus we consider bidirectional migration. In this system, $p_R = p_W = 0$, $q_W = q_S = 0$, $a_i = \bar{a}_i = 0$ for $i = R, W$. In this system, $N_{W1}$ and $N_{W2}$ are asymptotically constant to $N^*_{W1} = K_{W1}\left(1 - \frac{\mu_W}{r_W}\right)$, $N^*_{W2} = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right)$. Applying the results by Thieme, we obtain the following system in which $N_{W1}$ and $N_{W2}$ have reached their limiting values.

The reduced system becomes

$$\begin{aligned}
N'_{G2} &= r_G N_{G2}\left(1-\frac{N_{G2}}{K_{G2}}\right) - \mu_G N_{G2} - b_G I_{G2} + \tilde{b}_G I_{G1} \\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N^*_{W2}}(N_{G2}-I_{G2}) - (\mu_G + b_G) I_{G2} + \tilde{b}_G I_{G1} \\
I'_{W2} &= \beta_G \frac{I_{G2}}{N^*_{W2}}(N^*_{W2} - I_{W2}) - \mu_W I_{W2} \\
N'_{G1} &= r_G N_{G1}\left(1-\frac{N_{G1}}{K_{G1}}\right) - \mu_G N_{G1} - \bar{b}_G I_{G1} + \hat{b}_G I_{G2} \\
I'_{G1} &= \beta_W \frac{I_{W1}}{N^*_{W1}}(N_{G1}-I_{G1}) - (\mu_G + \bar{b}_G) I_{G1} + \hat{b}_G I_{G2} \\
I'_{W1} &= \beta_G \frac{I_{G1}}{N^*_{W1}}(N^*_{W1} - I_{W1}) - \mu_W I_{W1}
\end{aligned} \tag{3.8}$$

We determine the disease-free equilibrium for this system (3.8) to be of similar form to that of (3.6).

$R_0$ for the system is given as follows:

$$R_0 = \sqrt{\frac{1}{2}\left((g_1+g_2) + \sqrt{(g_1+g_2)^2 - 4g_1g_2\epsilon}\right)}$$

where $\epsilon = \frac{\mu_G(\mu_G+b_G+\bar{b}_G)}{(\mu_G+b_G)(\mu_G+\bar{b}_G)} < 1$, $g_1 = \frac{\beta_G\beta_W}{\mu_G\mu_W}\frac{N^*_{G1}}{N^*_{W1}}\left(\frac{\mu_G+b_G}{\mu_G+b_G+\bar{b}_G}\right)$ and
$g_2 = \frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}\left(\frac{\mu_G+\bar{b}_G}{\mu_G+b_G+\bar{b}_G}\right)$.

We observe that $R_0$ for system (3.8) is of similar form as the $R_0$ for system (3.4), with the exception of the migration terms. It is observed that $\max(\sqrt{g_1}, \sqrt{g_2}) < R_0 < \sqrt{g_1+g_2}$. If either of the migration terms $b_G$ or $\bar{b}_G$ is 0, then $R_0$ reduces to $\max(\sqrt{g_1}, \sqrt{g_2})$. We interpret $g_1$ as the basic reproductive number for patch 1 scaled by the proportion of infected vectors that stay in patch 1, and $g_2$ is the basic reproductive number for patch 2 scaled by the proportion of infected vectors staying in patch 2.

Determining existence of endemic equilibria is intractable analytically, but after a numerical investigation, we verify the hypothesis that in the case of bidirectional migration, only one endemic equilibria is possible if $R_0 > 1$.

### 3.3.6 Patches 1 and 2, 2 hosts 2 vectors, no vertical transmission, unidirectional migration of infected vectors

We end our exploration of two patches by considering the case with 2 patches, 2 hosts, 2 vectors, and unidirectional migration of infected vectors. Then $p_R = p_W = 0$, $a_i = \bar{a}_i = 0$ for $i = R, W$ and $b_S = \bar{b}_S = b_G = 0$. In this scenario, the quantities $N_{S2}, N_{R2}, N_{W2}$, and $N_{W1}$ are asymptotically constant to $N^*_{S2}, N^*_{R2}, N^*_{W2}$, and $N^*_{W1}$, respectively, where

$$N^*_{S2} = K_{S2}\left(1 - \frac{\mu_S}{r_S}\right), \quad N^*_{R2} = K_{R2}\left(1 - \frac{\mu_R}{r_R}\right),$$
$$N^*_{W2} = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right), \quad N^*_{W1} = K_{W1}\left(1 - \frac{\mu_W}{r_W}\right).$$

The model is

$$\begin{aligned}
I'_{S2} &= \left(q_S\beta_R\frac{I_{R2}}{N^*_{R2}} + (1-q_S)\beta_{WS}\frac{I_{W2}}{N^*_{W2}}\right)(N^*_{S2} - I_{S2}) - \mu_S I_{S2}\\
I'_{R2} &= \beta_{S2}\frac{I_{S2}}{N^*_{R2}}(N^*_{R2} - I_{R2}) - \mu_R I_{R2}\\
N'_{G2} &= r_G N_{G2}\left(1 - \frac{N_{G2}}{K_{G2}}\right) - \mu_G N_{G2} + \tilde{b}_G I_{G1}\\
I'_{G2} &= \beta_{W2}\frac{I_{W2}}{N^*_{W2}}(N^*_{G2} - I_{G2}) - \mu_G I_{G2} + \tilde{b}_G I_{G1}\\
I'_{W2} &= \left((1-q_W)\beta_G\frac{I_{G2}}{N^*_{W2}} + q_W\beta_{SW}\frac{I_{S2}}{N^*_{W2}}\right)(N^*_{W2} - I_{W2}) - \mu_W I_{W2}\\
N'_{G1} &= r_G N_{G1}\left(1 - \frac{N_{G1}}{K_{G1}}\right) - \mu_G N_{G1} - \bar{b}_G I_{G1}\\
I'_{G1} &= \beta_W\frac{I_{W1}}{N^*_{W1}}(N^*_{G1} - I_{G1}) - (\mu_G + \bar{b}_G)I_{G1}\\
I'_{W1} &= \beta_G\frac{I_{G1}}{N^*_{W1}}(N^*_{W1} - I_{W1}) - \mu_W I_{W1}
\end{aligned} \tag{3.9}$$

For this scenario, we determine the disease-free equilibrium to be $(0, 0, N^*_{G2}, 0, 0, N^*_{G1}, 0, 0)$, where

$$N^*_{G2} = N^*_{G1} = K_G\left(1 - \frac{\mu_G}{r_G}\right).$$

After calculating $R_0$, the structure seen is similar to that of $R_0$ for system (3.4), and is given by

$$R_0 = \max\{R_1, R_2\},$$

where

$$R_1 = \sqrt{\frac{\beta_G}{(\mu_G + \bar{b}_G)} \frac{\beta_W}{\mu_W} \frac{N^*_{G1}}{N^*_{W1}}}, \quad R_2 = \sqrt{\frac{1}{2}\left(P + \sqrt{P^2 - 4Q}\right)}.$$

$P$ and $Q$ are the same expressions as those in (3.5).

We determine existence of endemic equilibria in patch 2 alone if and only if $R_1 < 1$ and $R_2 > 1$, and in both patches if and only if $R_1 > 1$. The computations can be seen in Appendix B.3.

### 3.3.7 Patch 1, 2, and 3, 2 hosts 2 vectors, no vertical transmission, unidirectional migration of infected vectors

We finally extend our discussion to all three patches. We now analyze the system represented by patch 1, 2, and 3, with 2 hosts, 2 vectors and unidirectional migration of infected vectors. Thus, we consider $p_i = 0$, $a_i = \bar{a}_i = 0$ for $i = R, W$ and $b_S = b_G = 0$. In this scenario, the host populations, $N_{R3}, N_{R2}, N_{W2}$, and $N_{W1}$ are asymptotically constant to the values

$$N^*_{R3} = K_{R3}\left(1 - \frac{\mu_R}{r_R}\right), \quad N^*_{R2} = K_{R2}\left(1 - \frac{\mu_R}{r_R}\right)$$

$$N^*_{W2} = K_{W2}\left(1 - \frac{\mu_W}{r_W}\right), \quad N^*_{W1} = K_{W1}\left(1 - \frac{\mu_W}{r_W}\right).$$

The reduced system becomes

$$
\begin{aligned}
N'_{S3} &= r_S N_{S3}\left(1-\frac{N_{S3}}{K_{S3}}\right)-\mu_S N_{S3}+\tilde{b}_S I_{S2}\\
I'_{S3} &= \beta_R \frac{I_{R3}}{N^*_{R3}}(N_{S3}-I_{S3})-\mu_S I_{S3}+\tilde{b}_S I_{S2}\\
I'_{R3} &= \beta_S \frac{I_{S3}}{N^*_{R3}}(N^*_{R3}-I_{R3})-\mu_R I_{R3}\\
N'_{S2} &= r_S N_{S2}\left(1-\frac{N_{S2}}{K_S}\right)-\mu_S N_{S2}-\bar{b}_S I_{S2}\\
I'_{S2} &= \left(q_S\beta_R \frac{I_{R2}}{N^*_{R2}}+(1-q_S)\beta_{WS}\frac{I_{W2}}{N^*_{W2}}\right)(N_{S2}-I_{S2})-\mu_S I_{S2}-\bar{b}_S I_{S2}\\
I'_{R2} &= \beta_{S2} \frac{I_{S2}}{N^*_{R2}}(N^*_{R2}-I_{R2})-\mu_R I_{R2}\\
N'_{G2} &= r_G N_{G2}\left(1-\frac{N_{G2}}{K_{G2}}\right)-\mu_G N_{G2}+\tilde{b}_G I_{G1}\\
I'_{G2} &= \beta_{W2} \frac{I_{W2}}{N^*_{W2}}(N_{G2}-I_{G2})-\mu_G I_{G2}+\tilde{b}_G I_{G1}\\
I'_{W2} &= \left((1-q_W)\beta_G \frac{I_{G2}}{N^*_{W2}}+q_W\beta_{SW}\frac{I_{S2}}{N^*_{W2}}\right)(N^*_{W2}-I_{W2})-\mu_W I_{W2}\\
N'_{G1} &= r_G N_{G1}\left(1-\frac{N_{G1}}{K_{G1}}\right)-\mu_G N_{G1}-\bar{b}_G I_{G1}\\
I'_{G1} &= \beta_W \frac{I_{W1}}{N^*_{W1}}(N_{G1}-I_{G1})-\mu_G I_{G1}-\bar{b}_G I_{G1}\\
I'_{W1} &= \beta_G \frac{I_{G1}}{N^*_{W1}}(N^*_{W1}-I_{W1})-\mu_W I_{W1}
\end{aligned}
\tag{3.10}
$$

In analysis of system (3.10), we determine the disease free equilibrium to be $(N^*_{S3},0,0,N^*_{S2},0,0,N^*_{G2},0,0,N^*_{G1},0,0)$, where

$$
N^*_{S3}=K_{S3}\left(1-\frac{\mu_S}{r_S}\right),\quad N^*_{S2}=K_{S2}\left(1-\frac{\mu_S}{r_S}\right),
$$

$$
N^*_{G2}=K_{G2}\left(1-\frac{\mu_G}{r_G}\right),\quad N^*_{G1}=K_{G1}\left(1-\frac{\mu_G}{r_G}\right).
$$

We determine $R_0$ for this system to be

$$\begin{aligned} R_0 &= \max\{R_1, R_2, R_3\} \\ &= \max\left\{ \sqrt{\frac{\beta_G}{(\mu_G + \bar{b}_G)} \frac{\beta_W}{\mu_W} \frac{N^*_{G1}}{N^*_{W1}}}, \sqrt{\frac{1}{2}\left(P_2 + \sqrt{P_2^2 - 4Q_2}\right)}, \sqrt{\frac{\beta_R}{\mu_R} \frac{\beta_S}{\mu_S} \frac{N^*_{S3}}{N^*_{R3}}} \right\}. \end{aligned} \tag{3.11}$$

We define

$$R_1 = \sqrt{\frac{\beta_G}{(\mu_G + \bar{b}_G)} \frac{\beta_W}{\mu_W} \frac{N^*_{G1}}{N^*_{W1}}}, \quad R_2 = \sqrt{\frac{1}{2}\left(P_2 + \sqrt{P_2^2 - 4Q_2}\right)}, \quad R_3 = \sqrt{\frac{\beta_R}{\mu_R} \frac{\beta_S}{\mu_S} \frac{N^*_{S3}}{N^*_{R3}}}$$

and $P_2 = h_1 + h_2 + h_3$, $Q_2 = h_1\, h_3$, where

$$\begin{aligned} h_1 &= \frac{(1 - q_W)\beta_G}{\mu_G} \frac{\beta_{W2}}{\mu_W} \frac{N^*_{G2}}{N^*_{W2}}, \quad h_2 = \frac{q_W \beta_{SW}}{(\mu_S + \bar{b}_S)} \frac{(1 - q_S)\beta_{WS}}{\mu_W} \frac{N^*_{S2}}{N^*_{W2}}, \\ h_3 &= \frac{q_S \beta_R}{\mu_R} \frac{\beta_{S2}}{(\mu_S + \bar{b}_S)} \frac{N^*_{S2}}{N^*_{R2}}. \end{aligned} \tag{3.12}$$

Based on the form of $R_0$, we would expect three different scenarios for existence of endemic equilibria. We expect existence of endemic equilibria in all three patches if and only if $R_1 > 1$, in patch 2 and 3 only if and only if $R_2 > 1$ and $R_1 < 1$, and in patch 3 only if and only if $R_3 > 1$, $R_1 < 1$, and $R_2 < 1$. Investigation of these scenarios can be seen in Appendix B.4, in which we are able to show existence of at least one endemic equilibrium for each of the scenarios mentioned above.

### 3.3.8 Synthesis

By analyzing many smaller, sub-models of the original system (3.1), we may make some generalizations regarding the behavior of the full model. We expect that the full system will exhibit classical threshold behavior regarding $R_0$, in which we expect a unique endemic equilibrium for $R_0 > 1$. As stated previously, the form of $R_0$ in a vector-borne disease is a geometric mean between infections caused by hosts and infections caused by vectors. If the system considers multiple hosts and vectors,

the form of $R_0$ will include separate terms for each transmission cycle considered in the model. As described mathematically in section 3.3.3 and discussed in [43], we see vertical transmission has an "almost additive" effect on the basic reproductive number. In the case of unidirectional migration of infected vectors, we observe that $R_0$ consists of as many components as there are patches, and each component for $R_0$ contains parameters for only one patch. Also, as seen in section 3.3.7, multiple endemic equilibria are possible depending on the values of the patchwise reproductive numbers, $R_1, R_2, R_3$. We further note that by examination of (3.7), uninfected vectors migrating in one direction does not complicate the form of $R_0$. With bidirectional migration of infected vectors, we determine that the expression for $R_0$ involves contributions from all patches, rather than having a maximum of several components, which is to be expected since infection is moving in between patches. Based on this determination, we expect the full model to have one component for $R_0$, due to bidirectional migration; thus, it will not take on the form of $\max\{R_1, R_2, R_3\}$. In this case only one endemic equilibrium is possible; either there is no infection in any patch, or infection persists in all patches because all patches are connected by migration of infected vectors.

## 3.4 Numerical Results

### 3.4.1 General demographic parameters

We wish to investigate numerically the results of section 3.3, as well as investigate the behavior of the full model given by system (3.1). In order to do this, we will use biological information to estimate the parameters given in our model. Kribs [44] completed a thorough literature study to estimate demographic and *T. cruzi* infection

related parameters regarding hosts and vectors in the United States. We will use the demographic quantities calculated in Kribs [44], given in Table 3.1.

Table 3.1. Demographic parameters

| Species | $\mu$ | $r$ |
|---|---|---|
| Raccoon | 0.4/yr | 0.90/yr |
| Woodrat | 1/yr | 1.8/yr |
| *T. sanguisuga* | 0.271/yr | 33/yr |
| *T. gerstaeckeri* | 0.562/yr | 100/yr |

We mention here that in our model, each host species has a preferred habitat. That is, the preferred habitat for the raccoons is patch 3; thus we would expect a higher population density of raccoons in patch 3 compared to patch 2. Similarly, the woodrat preferred habitat is prickly pear cactus which predominates in patch 1, with a lower density in patch 2 due to the varying landscapes. Here, we will treat the parameters that differ for each patch.

#### 3.4.1.1 Patch 1

Kribs [44] obtains woodrat densities for Texas based on several sources [11, 68] which estimate woodrat population densities in counties in south and west Texas. These regions, especially the counties in west Texas, are similar to south Texas, dominated by shrub desert, including cactus and honey mesquite. He estimates the woodrat (equilibrium) population density in patch 1 to be 2300 woodrats/$km^2$. Because each population is governed by logistic growth with linear per-capita mortality, the populations approach an equilibrium population density, $N^*$, so that the carrying capacity can be back-calculated using $N^* = K\left(1 - \frac{\mu}{r}\right)$. Kribs estimates the total vector population density in patch 1 to be 31600 vectors/km$^2$. Since we are assum-

Table 3.2. Patch 1 density estimates

| Species | Population density | Carrying capacity |
|---|---|---|
| Woodrat | 2300 rats/km$^2$ | 5200 rats/km$^2$ |
| *T. gerstaeckeri* | 31600 vectors/km$^2$ | 31900 vectors/km$^2$ |

ing that the only vector in our model in patch 1 is *T. gerstaeckeri*, we use this as the density estimate. Using these figures, we obtain patch 1 population density and carrying capacity estimates found in Table 3.2.

#### 3.4.1.2 Patch 2

Estimates for southern plains woodrat density in patch 2 vary by geographical location and study. The woodrat density is affected by landscape, climate, and available materials for den construction. Raun [67] determined a positive correlation for population density of woodrats and density of cactus, although he concluded that cactus is not absolutely necessary to support woodrats. Cactus is the preferred material for den construction and food, but woodrats will use other materials to construct dens if cactus is unavailable [83]. In general, density of woodrat dens is closely associated with overhead cover.

In an 18 month study in Jim Wells Co. part of the Southern Texas Plains, Merkelz and Kerr [55] record a maximum density of 1.5 wr/ha (during spring 1998) using a 10 ha subplot of a 220 ha study site. Density was calculated by using the number of woodrats captured in the area during a single trapping season. Since they do not given any other density calculations, we use this data as part of our data collection. We note that they did not limit their density calculation to only areas with cactus growth so as to include open areas as part of the normal daily range of woodrats.

Conditt and Ribble [18] estimate a range of 1.6-5.8 wr/ha (average 3.7 wr/ha) in Bexar Co located in South Texas. The study was done on a 10 ha area of land with 4 ha dominated by honey mesquite-brush and prickly pear cactus, while the remaining 6 ha dominated by riparian lowland forests. The density was calculated on the 4 ha site due to essentially no woodrats being found on the riparian forested area (cactus-free) of the study site. They mention that the low density may be due to lack of cactus in the region of study and lack of appropriate shelter sites. However, Raymond et al. [68] in 2003 calculate a much higher maximum density of 19.4 wr/ha (with an average of 15.1 wr/ha) in the same county (but a different study site). This study site had limited clumps of prickly pear, but was covered with thick brush and downed trees which served as nest sites. Thus, we see that cactus is not absolutely necessary to maintain a high population density, but rather an abundance of shelter sites.

Although we do not include Oklahoma in patch 2, due to the northern range of *T. gerstaeckeri*, we refer to a study in Harmon County, Oklahoma in the Mesquite Grass Plains region, in which the estimated woodrat density was 13 wr/ha [84]. This region is native to the southern plains woodrat, dominated by Mesquite and prickly-pear cactus. The population density estimate was determined by trapping at 104 of the known 1,129 woodrat houses in the 226 ha study plot, and the density of cactus was not taken into account for the woodrat population density estimate.

Raun [67] estimates a range of 14.8-31.4 woodrats/ha (average 23 wr/ha) in San Patricio Co, part of a transitional region between the Southern Texas Plains and Western Gulf Coastal Plain. Thus, the vegetation in this region is diverse, with the major plant communities being Mesquite-Mixedgrass, Chaparral-Mixedgrass, Live Oak-Chaparral, and Prickly Pear-Short grass, with riparian forests along the rivers [31, 9]. The study site, 9 acres, was reduced to 7.3 acres to estimate the density to

Table 3.3. *Neotoma micropus* population density estimates

| Average density | Location | Ref |
|---|---|---|
| 1.5 wr/ha | Jim Wells Co | [55] |
| 3.7 wr/ha | Bexar Co | [18] |
| 13 wr/ha | Southwestern OK | [84] |
| 15.1 wr/ha | Bexar Co | [68] |
| 23 wr/ha | San Patricio Co | [67] |

eliminate areas that did not support cactus growth. A summary of these results can be found in Table 3.3.

Although there is a broad range of density estimates, we recall that patch 2 is a region with a diverse landscape, including mesquite, cactus, and savanna regions with areas of tree and prairie grassland. Since some estimates were computed in cactus-free regions, it is important to include each estimate in our computation of the average woodrat density. Thus, we include all 5 estimates in computing the average woodrat density in patch 2, obtaining an average population density of 11.3 wr/ha (1130 wr/$km^2$) in patch 2.

There are relatively few papers regarding raccoon distribution in patch 2. In a 3-year study by Gehrt and Fritzell [31], they estimate an average density of 7.3 raccoons/$km^2$ in San Patricio Co in southeast Texas. Since this region is not dominated by forest (as mentioned above), we would expect a lower raccoon density than that estimated by Kribs [44] for raccoon density in southeast USA (including patch 3). Using the equilibrium population densities, $N^* = K\left(1 - \frac{\mu}{r}\right)$, we calculate the carrying capacities for each species in patch 2, found in Table 3.4.

In patch 2, the *T. cruzi* transmission cycles overlap by the association of *T. sanguisuga* in association with both woodrats and raccoons, while *T. gerstaeckeri* feeds

Table 3.4. Patch 2 density estimates

| Species | Population density | Carrying capacity |
|---|---|---|
| Raccoon | 7.3 racc/km$^2$ | 13.1 racc/km$^2$ |
| Woodrat | 1130 rats/km$^2$ | 2542.5 rats/km$^2$ |
| All vectors | 31600 vectors/km$^2$ | 31900 vectors/km$^2$ |

only on woodrats. The *T. sanguisuga* move between the woodrat and raccoon populations regularly enough that we will consider the *T. sanguisuga* a single population. We would like to estimate $q_W$, the proportion of vector-woodrat contacts in patch 2 that are with *T. sanguisuga*, while $1-q_W$ is the proportion of vector-woodrat contacts that are *T. gerstaeckeri*. Also, since *T. sanguisuga* is associated with both hosts, we must estimate $q_S$, the proportion of *T. sanguisuga*-host contacts that are raccoons, with $1 - q_S$ the proportion of *T. sanguisuga*-host contacts made with woodrats.

Eads et al. [27] found 390 vectors from a total of 58 woodrat dens. Of the 390 vectors, 226 were *T. sanguisuga*, 133 were *T. gerstaeckeri*, and 31 were *T. neotomae*. We note here that the proportion of *T. neotomae* is negligible; thus we will normalize so that the proportions of *T. sanguisuga* and *T. gerstaeckeri* sum to 1. Thus, 63% of the vectors found in association with woodrats were *T. sanguisuga* and 37% were *T. gerstaeckeri*. Pippin [65] determined that for 85 woodrat dens, of 229 nymph and adult vectors, 58% were *T. sanguisuga*, while 42% were *T. gerstaeckeri*. If we pool the data, we determine 61% of the vector-woodrat contacts are with *T. sanguisuga*, and if we use the weighted average by number of dens excavated, the percentage is 60%. Thus, we estimate $q_W$ to be 0.605.

To estimate $q_S$, we will define $q_S = \frac{N^*_{R2} \cdot V_R}{N^*_{R2} \cdot V_R + N^*_{W2} \cdot V_W}$, where $N^*_{R2}$ is the patch 2 raccoon density and $N^*_{W2}$ is the patch 2 woodrat density. $V_R$ is the number of vectors per raccoon and $V_W$ is the number of vectors per woodrat (scaled by the proportion

that are *T. sanguisuga*). $V_R$ can be estimated from the estimates given by Kribs [44]. We mention here that there are other hosts in patch 3, so not all *T. sanguisuga* can be found with raccoons. Ideally, we would calculate $V_R$ directly to avoid biasing the estimates. Although raccoons are the preferred host of *T. sanguisuga*, the vector will feed on other hosts including opossums in patch 3. We will estimate the raccoon density equivalent of the opossum density. Based on literature reviews, we determine the population density of opossums to be 10.1/km$^2$ [44]. Thus, we determine the density of opossums is 0.505 times the raccoon density. We will divide the patch 3 *T. sanguisuga* population density, $N^*_{S3}$, estimated by Kribs [44] by the raccoon equivalent total host density. Using the estimates for $N^*_{S3}$ and $N^*_{R3}$ from Kribs [44], we determine $V_R = \frac{N^*_{S3}}{1.505 N^*_{R3}} = 1049.83$ *T. sanguisuga*/raccoon. Since both vector species are associated with woodrats, $V_W$ is $q_W \frac{N^*_{G2}+N^*_{S2}}{N^*_{W2}}$. Kribs [44] estimates that the total *Triatoma* vector population in patch 2 is 31600 vectors/km$^2$. As estimated previously, $N^*_{W2} = 1130$ woodrats/km$^2$. These averages result in an estimate of $V_W$ as 16.9 *T. sanguisuga*/woodrat. Based on this calculation, we arrive at an estimate of $q_S = 0.286$.

#### 3.4.1.3 Patch 3

We use Kribs [44] estimates for the raccoon and *T. sanguisuga* density estimates in patch 3, given in Table 3.5. We note here that the *T. sanguisuga* population density is based on research by Kribs [44]. Kribs identifies only 1 study done by Burkholder et al. [13] regarding population density of *Triatoma* vectors, which estimates *Triatoma* density in relation to woodrat nests. Although *T. sanguisuga* in patch 3 are found with raccoons, we use the same estimate as the total vector population density in patches 1 and 2, due to lack of relevant information on vector population density in patch 3.

Table 3.5. Patch 3 density estimates

| Species | Population density | Carrying capacity |
|---|---|---|
| Raccoon | 20. racc/km$^2$ | 35.6 racc/km$^2$ |
| *T. sanguisuga* | 31600 vectors/km$^2$ | 31900 vectors/km$^2$ |

### 3.4.2 Estimation of infection rate parameters

The model here includes 2 modes of host infection: vertical transmission and direct transmission due to biting and to oral transmission via vector consumption. The vertical transmission parameters can be estimated directly via literature. Kribs [44] estimates the vertical transmission proportion to be 0.01. Then $p_W = 0.01$. We estimate $p_R = 0.1$ due to the adaptation of *T. cruzi* strain type IV to vertical transmission in raccoons. To estimate the direct infection rate parameters, we utilize the technique outlined in [44] to back-calculate the infection rate parameters, by solving for $\beta_R, \beta_S, \beta_{SW}, \beta_{WS}, \beta_G$, and $\beta_W$ using the equilibrium conditions for model (3.1) under the assumption that observed prevalence indicates endemic equilibrium.

Since migration is small compared to the demographic processes, we will estimate the infection rate parameters by patch using model (3.1), assuming no migration. In order to estimate the infection rate parameters, we combine the observed prevalence levels and known demographic parameters in the equilibrium conditions to back-calculate the infection rate parameters, $\beta_R, \beta_S, \beta_{SW}, \beta_{WS}, \beta_{S2}, \beta_{W2}, \beta_G$, and $\beta_W$.

After a thorough literature search, Kribs calculates prevalence levels for each species in each patch. Prevalence levels for *T. gerstaeckeri* and the woodrat are given for Texas and levels for *T. sanguisuga* and raccoons are given for Texas and the southeast U.S. To translate these values to this model, we assume that the patch 1 and 2 prevalence levels for *T. gerstaeckeri* and southern plains woodrat are equivalent

to the Texas estimates found by Kribs. The patch 2 and 3 prevalence levels for *T. sanguisuga* and the raccoon are the same as the Texas and southeast estimates, respectively. A summary of these values is given in Table 3.6.

Table 3.6. *T. cruzi* prevalence estimates from [44]

| Species | Patch 1 | Patch 2 | Patch 3 |
|---|---|---|---|
| Raccoon | - | 0.240 | 0.387 |
| *T. sanguisuga* | - | 0.249 | 0.565 |
| Woodrat | 0.332 | 0.332 | - |
| *T. gerstaeckeri* | 0.454 | 0.454 | - |

We note here that $x^*_{G1} = \frac{I^*_{G1}}{N^*_{G1}}$ is the prevalence value for *T. gerstaeckeri* in patch 1. We utilize a similar notation for the other 3 species in the model.

Using the equilibrium conditions and substituting the population density estimates and prevalence values for patch 1,

$$\beta_W x^*_{W1}(1 - x^*_{G1}) - \mu_G x^*_{G1} = 0$$
$$p_W r_W x^*_{W1}\left(1 - \frac{N^*_{W1}}{K_{W1}}\right) + \beta_G x^*_{G1}\frac{N^*_{G1}}{N^*_{W1}}(1 - x^*_{W1}) - \mu_W x^*_{W1} = 0,$$

we obtain estimates for $\beta_W$ and $\beta_G$, given in Table 3.7.

Substituting the patch 3 population density estimates and prevalence values into the patch 3 equilibrium conditions,

$$\beta_R x^*_{R3}(1 - x^*_{S3}) - \mu_S x^*_{S3} = 0$$
$$p_R r_R x^*_{R3}\left(1 - \frac{N^*_{R3}}{K_{R3}}\right) + \beta_S x^*_{S3}\frac{N^*_{S3}}{N^*_{R3}}(1 - x^*_{R3}) - \mu_R x^*_{R3} = 0,$$

we calculate $\beta_R$ and $\beta_S$, given in Table 3.7.

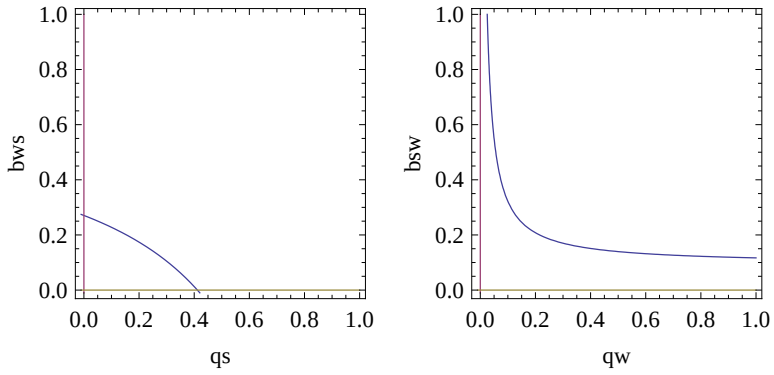


Figure 3.2. Correspondence of $q_S$ and $q_W$ with $\beta_{WS}$, with units 1/yr, and $\beta_{SW}$, with units infected hosts/vector/yr.

Due to the crossover of infection cycles in patch 2, back-calculating the infection rate parameters is complex. After substituting the patch 2 equilibrium population densities and prevalence levels,there are 4 equilibrium conditions remaining

$$
\begin{aligned}
&(q_S\beta_R x^*_{R2} + (1-q_S)\beta_{WS}x^*_{W2}) - \mu_S x^*_{S2} = 0\\
&p_W r_W x^*_{W2}\left(1-\frac{N^*_{W2}}{K_{W2}}\right) + \left((1-q_W)\beta_G x^*_{G2}\frac{N^*_{G2}}{N^*_{W2}} + q_W\beta_{SW}x^*_{S2}\frac{N^*_{S2}}{N^*_{W2}}\right)(1-x^*_{W2})\\
&-\mu_W x^*_{W2}N^*_{W2} = 0\\
&p_R r_R x^*_{R2}\left(1-\frac{N^*_{R2}}{K_{R2}}\right) + \beta_{S2}x^*_{S2}(1-x^*_{R2})\frac{N^*_{S2}}{N^*_{R2}} - \mu_R x^*_{R2} = 0\\
&\beta_{W2}x^*_{W2}(1-x^*_{G2}) - \mu_G x^*_{G2} = 0.
\end{aligned}
\tag{3.13}
$$

We note that there are six $\beta$ values and four equations; thus the system is underdetermined. Consequently, we will assume that $\beta_R$ and $\beta_G$ have the same values in patch 2 as calculated in patches 3 and 1, respectively. After solving system (3.13), we determine values for $\beta_{SW}$, $\beta_{WS}$, $\beta_{S2}$, and $\beta_{W2}$, given in Table 3.7.

After an investigation of the effects of the proportions $q_S$ and $q_W$ on the $\beta_{WS}$ and $\beta_{SW}$ with all of the other parameters held fixed as determined above, we determine

that any value of $q_W$ greater than 0.3 will produce a value of $\beta_{SW}$ in between 0.11/yr and 0.15/yr, which mitigates any inaccuracy in our estimate of $q_W = 0.605$. We also note that, mathematically, $q_S$ needs to be smaller than 0.411 for $\beta_{WS}$ to be positive, as seen in Figure 3.2.

Table 3.7. Stercorarian infection rate parameters

| Parameter | Value | Units |
|---|---|---|
| $\beta_S$ | 0.00025 | $\frac{\text{hosts}}{\text{vector}\cdot\text{yr}}$ |
| $\beta_G$ | 0.079 | |
| $\beta_{SW}$ | 0.132 | |
| $\beta_{S2}$ | 0.00017 | |
| $\beta_R$ | 0.910 | 1/yr |
| $\beta_W$ | 1.408 | |
| $\beta_{WS}$ | 0.116 | |
| $\beta_{W2}$ | 1.408 | |

Of the vector to host transmission terms, $\beta_S$ and $\beta_{S2}$ (vector to raccoon infection rate) are close in value, while $\beta_G$ and $\beta_{SW}$ (vector to woodrat infection rate) are close in value. But, we note here that there is a factor of 1000 by which woodrats are being infected more rapidly than raccoons are infected by *T. sanguisuga*. We note this may be attributed partly to the fact that the population densities for woodrats are approximately 100 times as great as raccoon population densities. All of the host to vector transmission parameters are close in value, with the exception of $\beta_{WS}$ (woodrat infecting *T. sanguisuga*) which is an order of magnitude less than the raccoon to *T. sanguisuga* infection rates, which is to say that woodrats are infecting *T. sanguisuga* at a lower rate than raccoons infecting *T. sanguisuga*.

### 3.4.3 Numerical solutions

#### 3.4.3.1 Trends in migration rate effects

To investigate the effects of migration numerically, we examine prevalence of *T. cruzi* as a function of vector migration. In order to get a clear picture, we first look at unidirectional migration of infected vectors. We then investigate unidirectional migration of all vectors as this is the upper bound of the possible unidirectional migration scenarios for vector migration. We then consider bidirectional migration of infected and all vectors. In each scenario, we consider the effects on *T. cruzi* prevalence as migration increases. To see the effects of the increase, we consider the effects of factor, $k$, multiplied by each migration rate. We calculate the prevalence for each patch as $k$ increases from 0 to 10, where $k = 0$ represents the scenario with no migration and $k$ increasing greater than 1 represents the migration rate increasing past the calculated rate from Chapter 2. For each graph given, the host infected prevalence is represented by the lighter colored graph and the vector infected prevalence is represented by the darker colored graph.

In the case of northward migration of infected vectors (Figure 3.3), we observe a decrease of prevalence in patch 1 since this patch is losing infected vectors and not gaining any, and a slight increase of prevalence in patch 3 as the migration rate increases. The prevalence of *T. sanguisuga* in patch 2 is much lower than that in patch 3, so although patch 3 is gaining infected vectors, the increase in prevalence at equilibrium is minimal.

In this scenario, patch 2 is gaining infected vectors from patch 1(*T. gerstaeckeri*) and losing infected vectors to patch 3(*T. sanguisuga*). As seen in Figure 3.3, the patch 2 prevalence decreases for northward migration. To get a better understanding of why the prevalence decreases in patch 2, we look at the graphs of the prevalence

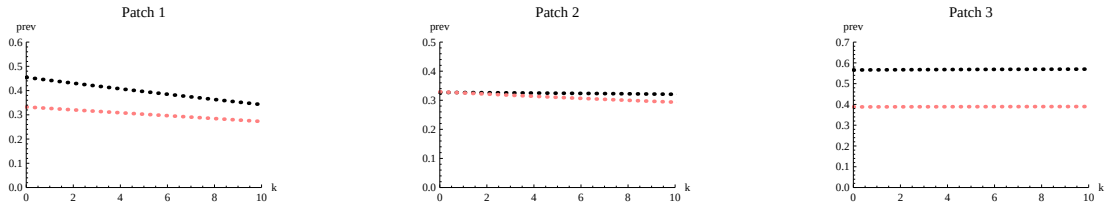


Figure 3.3. Infected vectors migrating north only; dark curve represents vectors, light curve represents hosts.

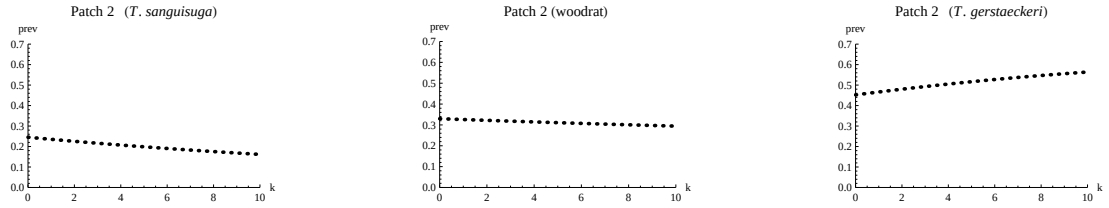


Figure 3.4. Infected vectors migrating north only.

for each vector species in patch 2, seen in Figure 3.4. It can be observed that the *T. sanguisuga* infected prevalence decreases by approximately 35% for high migration rates, while the *T. gerstaeckeri* prevalence increases by approximately 25%. Furthermore, we see a constant decrease in the woodrat prevalence (although *T. gerstackeri* prevalence increases). We note that in patch 2 $N_{S2}$ is approximately 50% greater than $N_{G2}$. Thus, the export of infected *T. sanguisuga* from patch 2 dominates the import of infected *T. gerstaeckeri* from patch 1, causing an overall prevalence decrease in patch 2 in the case of northward migration of infected vectors.

When we consider northward migration of all vectors, the change in prevalence is undetectable by viewing the graph. Patch 1 prevalence is still reduced since it is losing both infected and uninfected vectors, as seen in Figure 3.5. One might expect that the decrease should be less when all vectors are migrating, but after a

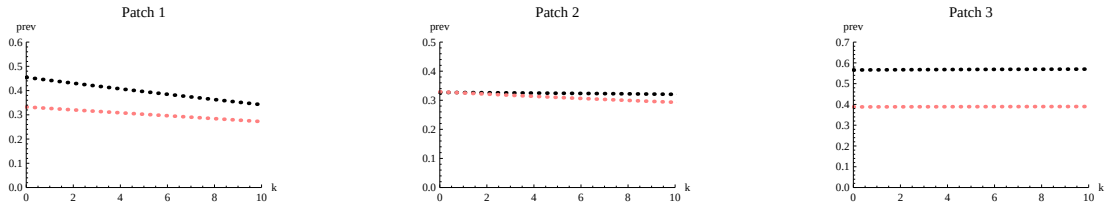


Figure 3.5. All vectors migrating north only; dark curve represents vectors, light curve represents hosts.

numerical investigation of prevalence, we see that the decrease is slightly greater when all vectors are migrating. Also, we would expect the patch 3 prevalence to increase since this patch is gaining both infected and uninfected vectors from patch 2 with a lower prevalence; howeverm the prevalence decreases slightly. If only infected vectors migrate northward, then $R_1$ decreases as $\bar{b}_G$ increases, causing prevalence in patch 1 to decrease. When all vectors are migrating northward, $R_1$ actually decreases more since the vector-host ratio is also decreasing, as the migration rate increases. Thus, the decrease in $R_1$ is amplified when all vectors are migrating northward.

As we observe southward migration of infected vectors, we again view the prevalence levels for each patch, in Figure 3.6. Patch 1 is gaining infected vectors from patch 2; thus we see an increase in prevalence. In patch 3, we see a slight decrease ($< 1\%$) in prevalence due to this patch losing infected vectors. In contrast to northward migration, we see a rise in prevalence for patch 2 in the case of southward migration. To understand why prevalence increases even though this patch is losing infected vectors to patch 1, we graph prevalence levels for each species in patch 2 as a function of migration, seen in Figure 3.7. We observe that prevalence in *T. sanguisuga* increases, as expected since this vector population is gaining infected vectors from patch 3. An interesting observation is that the prevalence in *T. gerstaeckeri* rises for small migration rates ($0 < k < 1$). Then, for higher migration rates, the

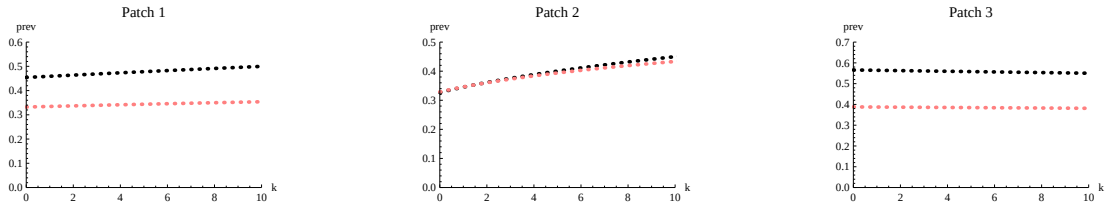


Figure 3.6. Infected vectors moving south only; dark curve represents vectors, light curve represents hosts.


Figure 3.7. Infected vectors migrating south only.

prevalence for *T. gerstaeckeri* in patch 2 decreases. As observed in Figure 3.7, the woodrat *T. cruzi* prevalence increases in patch 2 for southward migration. Thus, for $0 < k < 1$, the sharp rise in prevalence in *T. sanguisuga* (11% increase) in patch 2 for small migration rates may be enough to increase the prevalence in *T. gerstaeckeri* through the woodrats.

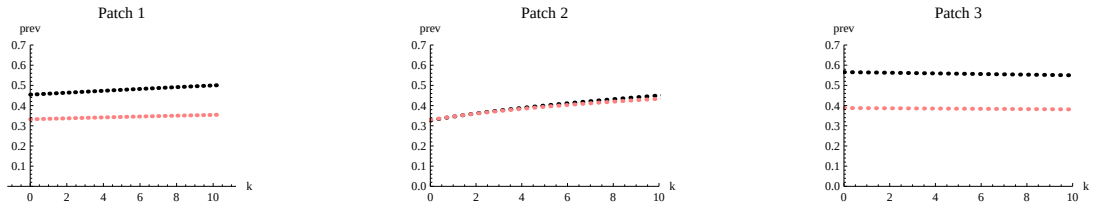


Figure 3.8. All vectors migrating south only; dark curve represents vectors, light curve represents hosts.

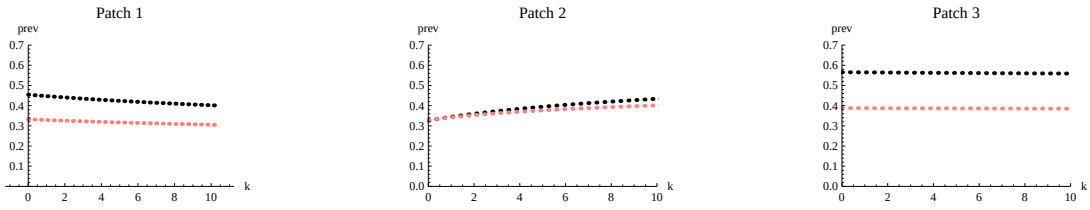


Figure 3.9. Infected vectors bidirectional migration; dark curve represents vectors, light curve represents hosts.

In the case of southward migration of all vectors, the patch 3 prevalence decreases more than when only infected vectors are moving southward. This can be attributed to the decrease in vector-host ratio as the migration rate increases, thereby causing $R_3$ to decrease more than if only infected vectors are migrating. Patch 1 prevalence increases more than if only infected vectors are migrating southward, due to the slight increase in vector-host ratio. The patch 2 prevalence increases more than if only infected vectors are migrating southward, which is primarily due to the increase in vector-host ratio causing $R_2$ to increase more than if only infected vectors are migrating, as seen in Figure 3.8.

We also wish to investigate bidirectional migration of vectors. We first consider bidirectional migration of infected vectors only. As seen in Figure 3.9, prevalence decreases in patches 1 and 3, but increases in patch 2. To better understand why prevalence increases in patch 2, we observe the prevalence graphs for patch 2 only. We observe that both *T. sanguisuga* and *T. gerstaeckeri* prevalence increases in patch 2 (Figure 3.10).

The *T. sanguisuga* population in patch 2 is initially at a lower prevalence than patch 3, and since the *T. sanguisuga* population is gaining vectors from a patch with a higher prevalence, the prevalence increases for this species in patch 2 as seen in


Figure 3.10. Infected vectors bidirectional migration; dark curve represents vectors, light curve represents hosts.

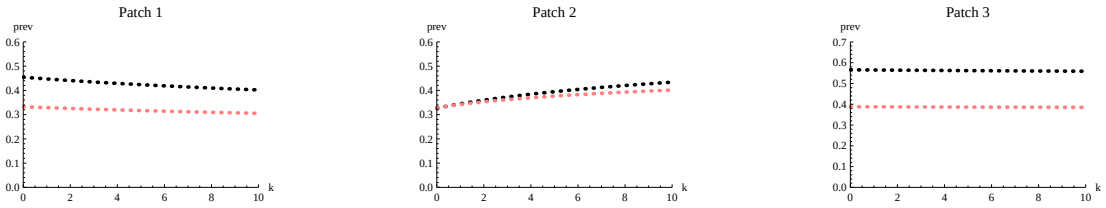


Figure 3.11. All vectors bidirectional migration; dark curve represents hosts, light curve represents vectors.

Figure 3.10. The *T. gerstaeckeri* prevalence also increases, although this population is gaining and losing vectors at the same prevalence. We can most likely attribute the increase in prevalence for *T. gerstaeckeri* in patch 2 due to the increase in woodrat prevalence (seen in Figure 3.10) through the *T. sanguisuga* prevalence increase.

In the case of bidirectional migration of all vectors, the behavior is similar. The patch 1 and 3 prevalence decreases slightly more than in the case of only infected vectors migration, due to the decrease of the vector-host ratio. A similar reason can be given for the patch 2 prevalence increasing more than if infected vectors are moving only.

#### 3.4.3.2 Calculation of the migration rate

We establish in Chapter 2 that vector dispersal can be described by three properties: dispersal distance, preferred direction of dispersal, and degree of preference for a particular direction. Since we do not have clear data on a preference for a direction for *T. sanguisuga* and *T. gerstaeckeri*, we will consider the simplest case in which the vectors have no preference for direction of dispersal. Here we give the vector migration rates calculated in Chapter 2 assuming no preference for direction. We adjust these rates to take into account the area ratios and note that the rates are equal for all vectors (infected and susceptible) in each species.

Table 3.8. Migration rates for no preferred direction (units in 1/year)

| $\bar{m}$ | Species | rate |
|---|---|---|
| $m_{12}$ | *T. gerstaeckeri* | 0.00427 |
| $m_{21}$ | *T. gerstaeckeri* | 0.00385 |
| $m_{23}$ | *T. sanguisuga* | 0.00101 |
| $m_{32}$ | *T. sanguisuga* | 0.000155 |

Table 3.9. Adjusted migration rates for no preferred direction (units in 1/year)

| | Adjusted rate |
|---|---|
| $\tilde{b}_G = \tilde{a}_G$ | 0.00385 |
| $\hat{b}_G = \hat{a}_G$ | 0.00427 |
| $\tilde{b}_S = \tilde{a}_S$ | 0.000155 |
| $\hat{b}_S = \hat{a}_S$ | 0.00101 |

We may also consider that vectors have a preference for direction of migration. Although we do not have clear evidence for the vectors in the model migrating with any clear trend in direction, we consider the possibility of vector migration in a particular direction. In a study on climate change related to Chagas disease distribution, Curto de Casas concludes that higher temperatures may extend the geographical range of the sylvatic vectors of *T. cruzi* [21]. Thus, we may consider vector migration with a northward preference for direction. We note here that when northward is described in section 3.4.3.1, it is referring to one-directional migration from patches

1 to 2 and 2 to 3, not the actual geographical direction north. In this section we are referring to the actual direction north and migration is bidirectional between patches.

Table 3.10. Equilibrium prevalence levels for species based on migration (northward preference for direction)

| Species | Equilibrium prevalence |
|---|---|
| Patch 3 | |
| *T. sanguisuga* | 0.56384 |
| Raccoon | 0.38513 |
| Patch 2 | |
| *T. sanguisuga* | 0.24940 |
| Raccoon | 0.23628 |
| *T. gerstaeckeri* | 0.46649 |
| Woodrat | 0.33459 |
| Patch 1 | |
| *T. gerstaeckeri* | 0.44686 |
| Woodrat | 0.32858 |

The framework for deriving migration rates with a preference for direction is modeled using a sequence of nested ellipses, in which each ring represents a certain range if dispersal distances. We also assume that vectors have a degree of preference for a direction, which represents the eccentricity of each ellipse. Based on results from Chapter 2, we give results for *T. cruzi* prevalence assuming a northward preference of direction (with moderate degree of preference, $e = 0.5$) in Table 3.10. The decrease in prevalence in patch 1 and increase in prevalence for *T. gerstaeckeri* in patch 2 are consistent with what we expect assuming a northward preference for direction. We note that there is approximately a 1% decrease in prevalence for *T. sanguisuga* and raccoons in patch 2 if preferred direction is northward, when compared to migration with no preference for direction. There is a more than 3% increase in *T. gerstaeckeri* prevalence in patch 2 when compared no migration. Although *T. gerstaeckeri* feeds

only on woodrats, the increase in prevalence for *T. gerstaeckeri* is not enough to cause a higher increase in woodrat prevalence; thus we only observe only a slight (<1%) increase in prevalence for woodrats in patch 2 when the preferred direction of migration is northward.

## 3.5 Conclusions

Due to the complexity of system (3.1), several sub-models were analyzed in order to make conclusions regarding the behavior of the full model. The entire system appears to exhibit classical threshold behavior regarding $R_0$, and existence of a unique endemic equilibrium when $R_0 > 1$. In the case of one-directional migration of vectors, $R_0$ consists of as many components as there are patches with $R_0$ being the largest value of the patch-specific $R_0$ values, and multiple endemic equilibria are possible depending on the values of the patch-specific reproductive numbers values. If migration is bidirectional, $R_0$ will consist of only one component for containing parameters from all patches, and only one endemic equilibrium is possible; thus either the whole system reaches a disease-free state or infection persists in all patches. These results are similar to the multi-patch model results in [3], where it is shown that if patch $a$ is at endemic equilibrium, then the disease is at endemic equilibrium in each patch accessible to patch $a$. Similar results are observed in the two-patch model in [2], where it is shown that if $R_0 < 1$, the disease does not persist in either population (patch), and if $R_0 > 1$, the disease persists in both populations.

In section 3.4, we obtained several different sets of results with respect to the effects of vector migration on the prevalence of *T. cruzi* in the geographical region from which we built our model. We examined the effects of one-directional migration, bidirectional migration (at symmetric rates), and bidirectional migration using rates derived from the framework in Chapter 2.

In the case of one-directional migration, the prevalence in the outer patches (patches 1 and 3) varies as expected (e.g. patch 1 prevalence decreases for northward migration only). However, the patch 2 dynamics are different for northward vs. southward migration. In the case of northward migration only, the patch 2 prevalence decreases while for southward migration, the prevalence increases. After closer observation of prevalence for each species in patch 2, it is determined that the *T. sanguisuga* migration has a greater effect on the patch 2 dynamics than the migration of *T. gerstaeckeri* primarily due to the fact that the difference in population density and prevalence in patches 2 and 3 among *T. sanguisuga* is greater than the difference in population density and prevalence in *T. gerstaeckeri* in patches 1 and 2. If vectors are migrating northward only, the export of *T. sanguisuga* dominates the import of *T. gerstaeckeri* thereby causing an overall decrease in patch 2 prevalence; we note that the *T. sanguisuga* population density in patch 2 is approximately 50% greater than the *T. gerstaeckeri* population density. For southward migration rates, the import of *T. sanguisuga* dominates the export of *T. gerstaeckeri* into patch 1. In fact, for lower migration rates, the prevalence for *T. gerstaeckeri* increases slightly due to the sharp increase in *T. sanguisuga* through the connection with the woodrat host. The connection between the vectors is the infection cycle with the woodrat host. Thus, the increase in prevalence for *T. gerstaeckeri* for lower southward migration rates is due to the increase in the prevalence for woodrats through the infection cycle with *T. sanguisuga*.

If migration is bidirectional, the patch 2 prevalence increases. After analysis of one-directional migration, the reason for the patch 2 prevalence increase is more apparent. We note here that the prevalence for *T. sanguisuga* is initially at a lower prevalence in patch 2 than in patch 3, so the *T. sanguisuga* in patch 2 is gaining vectors from a patch with a higher prevalence, thus the increase in *T. sanguisuga*

prevalence. The prevalence for *T. gerstaeckeri* increases in patch 2 in the case of bidirectional migration due to the increase in woodrat prevalence (again through the infection cycle with *T. sanguisuga*). We note here that these trends are observed even if migration rates are considered symmetric (independent of patch size).

Because of the differences in patch sizes (especially the large size of patch 3 compared to patches 1 and 2), the vector density in each patch is affected differently by migration. For example, the change in patch 2 vector density will be greater than the change in patch 3 vector density for bidirectional migration. The patch 3 vector density is minimally affected by migration due to the large patch size.

Based upon these results we can conclude that infection dynamics in patch 2 are sensitive to migration, but primarily driven by the *T. sanguisuga* population. Since the same trends in prevalence change are observed for migration independent of patch size, we should investigate the distinctive transmission characteristics between host and vector. Thus, we consider differences in the *T. sanguisuga* and *T. gerstaeckeri* populations. One major difference in the vector populations is the high prevalence for *T. sanguisuga* in patch 3 compared to patch 2 and the prevalence for *T. gerstaeckeri*. If we assume no difference in prevalence levels for *T. sanguisuga* from patch 3 to 2, we can note that for southward migration, the prevalence in patch 2 still increases for small migration rates, but at a much slower rate, while for northward migration, the prevalence decreases but at a much slower rate than if the prevalence in patch 3 is at its current estimated level. These results give a different picture for the patch 2 dynamics, which allow us to see a dampening effect of the *T. sanguisuga* migration if the prevalence levels for the *T. sanguisuga* populations in patches 2 and 3 are the same. Thus, we can attribute the majority of the patch 2 dynamics when migration is considered to the higher prevalence in *T. sanguisuga* in patch 3.

This aim of this study is to determine the effect of vector migration on *T. cruzi* transmission in the prominent sylvatic cycles ranging from northern Mexico to the southeastern United States. The primary effect of migration is to increase prevalence in the overlap patch where the prevalence is initially at a lower level than the outer (single-cycle) patches. The dominant force is the connection to the large raccoon-*T. sanguisuga* in the southeastern United States, which is affected little on such a large scale, but which can affect dynamics strongly in the overlap region with woodrats and *T. gerstaeckeri*.

As in every study, certain limitations must be noted. Due to the very recent awareness of the need to study *T. cruzi* in the United States, there have been very few studies on the demography of vectors native to the U.S., especially regarding vector population density. To date, there have been virtually no studies on the U.S. *Triatoma* vectors' dispersal or migration capabilities. Thus, with more studies, we may be able to more accurately describe the dispersal capabilities (especially regarding distance and frequency of vector flights). We acknowledge the need for more experimental or field studies with heavy consideration on the *Triatoma* vectors native to the United States, especially *T. gerstaeckeri* and *T. sanguisuga* which we consider to be the primary vectors in the southeast.

Because of the differing patch sizes, a natural question arises to consider effects of migration for higher spatial resolution where all patches are of uniform size. If the geographical region is broken into smaller, same-size patches, we wish to examine to what extent these results would change. Future work in the next chapter of the dissertation is to use cellular automata to see how migration among smaller patches will affect the spread of *T. cruzi* across a global geographic region as well as determine a measure for speed of invasion.

# CHAPTER 4

# INVASION SPEED IN CELLULAR AUTOMATON MODELS FOR *T. CRUZI* VECTOR MIGRATION

## 4.1 Introduction

Chagas' disease, caused by the parasite *Trypanosoma cruzi*, is considered endemic in Central and South America. However, fewer than 10 autochthonous human cases have been diagnosed in the United States, despite the presence of *T. cruzi* in many mammals and triatomine bugs in the U.S. Although there is no clear reason for the paucity of human infections, the lower rates of human *T. cruzi* infection have been attributed to several factors, including fewer domestic vectors, lower North American vector competence (delayed defecation after feeding), and the potential lower virulence of native North American strains of *T. cruzi* [25]. The interaction between the mammals and triatomine species native to the U.S. and triatomine species carrying the more virulent strains in Central and South American is not known. Because the activity level of triatomine bugs is greatly affected by temperature and climate, increasing global temperatures may be likely to extend the vector distribution northward. Lambert et al. [46] studied the potential for the emergence of Chagas' in the U.S., and cite other studies in which increased temperature and climate change have contributed to the spread of certain vector-borne diseases into new locations. Although several factors may play a role in the relatively low number of human cases of Chagas in the U.S., the parasite is highly endemic among sylvatic vector-host populations. Thus, attention should be given to the potential for the spread of the more virulent strains of *T. cruzi* from Central and South America to the sylvatic popula-

tions in the U.S. through the movement of the triatomine vectors. In this study, we wish to investigate the speed of spread of the more virulent strains of *T. cruzi* into the southeastern U.S. through mobility of the vector populations.

Because of the importance of considering the geographic nature of the spread of *T. cruzi*, a spatially explicit model should be considered. Although there are many types of spatially explicit models, we consider a two-dimensional cellular automaton (CA) model. A two-dimensional CA is a regular grid in which each cell is in a finite state at each time step. The state of each cell evolves over time based on a set of rules that take into account the state of each cell and its neighboring cells. Thus, local interactions between neighboring cells yield global results over the entire grid that can be observed over time. For this study, the local CA rules will come from Chapter 3 (the existing metapopulation models for *T. cruzi* transmission). A CA can be either deterministic or stochastic, and in this paper we will consider the deterministic CA.

In biological models, the term *invasion speed* is generally associated with the speed at which a certain population expands over space. The population could be considered an invading population into a region or territory or could refer to an infection invading a spatial region. In this study we consider the effects of an invading strain of *T. cruzi* on the sylvatic vector and host populations in the U.S. In particular, the focus of this study is to consider how the speed and direction of invasion vary with vector mobility. Invasion speed will be examined under different vector migration scenarios. In this study, we will determine a criterion for measuring invasion wavefronts as well as develop a method to calculate invasion speed for varying vector migration characteristics, such as preference for a particular direction. We will compare results across different patches. After reviewing relevant literature on spatially explicit epidemic models and invasion speed, we begin by defining the proposed CA model structures and developing the necessary criteria for invasion. Methods for calculating

invasion speed are described, and invasion speed of the infection are determined for different migration rates.

## 4.2 Literature Review

Several types of spatially explicit models have been used to model epidemiological and ecological problems involving population dispersal. In a spatially explicit model, the spatial habitat can be modeled in a continuous or discrete way. The most common continuous-space models are reaction-diffusion and integro-differential (or difference) equations. Some common models discrete in space are cellular automaton (CA) models and other lattice-based models. Each of these types of models produces different results with characteristics which depend heavily on the scale and structure of the spatial landscape. CA models generally incorporate a significant amount of detail, and thus due to complexity issues are analyzed numerically. Reaction-diffusion and integro-differential equations provide less detail, but are easier to analyze mathematically.

Reaction-diffusion models have been extensively studied in the context of dispersing populations and epidemics [38, 49]. The reaction term is generally a growth (or birth) term, while the diffusion term represents the population dispersal. In reaction-diffusion models, the general assumption is that the dispersal is random and local. Murray et al. [60] investigated a simple reaction-diffusion model for the spread of rabies among foxes in England. Because infected foxes tend to disperse in a more or less random way, a diffusion process is a good choice to model the fox dispersal. Analysis of the model predicts traveling waves of the epidemic. The wave front is defined to be the front in which the largest number of foxes are dying from the disease. The speed of the wave increases in regions with higher fox density and predicts an interesting feature in which there are outbreaks of the epidemic in regions of higher

fox density ahead of the actual wave front. It has been suggested that in certain reaction-diffusion models, the traveling wave speed may actually be underestimated [42]. In this case, the use of integral operators rather than diffusion operators may be used. The result is a system of integro-differential (or integro-difference) equations [58]. In a study by Medlock and Kot [54] and references therein, the spread of disease in space is discussed using two models referred to as the *distributed-contacts* (DC) and *distributed-infectives* (DI) models. In each of these models, either contacts are local with infectives dispersing in space (*distributed-infectives*) or contacts are spatially distributed in space (*distributed-contacts*). Each of these scenarios is modeled by a system of integro-differential equations. Solutions to such models (and reaction-diffusion models) exhibit wavelike behavior (commonly traveling waves). Results show that the epidemic spreads faster in the DC model than in the DI model if the transmission rate is high, while if the rate is low, the DI model causes faster speeds of spread.

Cellular automata and other grid-based models have only recently been used in the study of spatial spread of epidemics. In its most simplistic view, the basic cellular automaton model applied to the spread of a disease consists of a set of grid points (or cells) distributed on a square lattice where each cell has a certain set of neighboring cells called a neighborhood. Each cell typically represents an individual that can be in one of a finite number of states. In the most basic epidemiological models, the states may be susceptible, infected (or infectious), recovered, immune, etc. Individuals move between states at different rates based on the states of the neighboring cells. The transmission of the disease and movement between states may be deterministic or stochastic [77, 80, 91]. The models previously cited obey the more traditional rules for a CA (each cell represents an individual); however, there are other ways of defining a cell. In [30], the cells contain specific populations based on certain

geographical features. The population in each cell has specific demographic features as well as susceptible, infected, and recovered subpopulations for an unspecified SIR infection. The populations and movement between cells are governed by each cell's carrying capacity, whereas the success of the movement of a population between cells depends on whether the destination cell has reached its carrying capacity. Rather than studying the spatial spread of a specific disease across the lattice, [30] gives results for two different experimental disease scenarios. For each scenario, the landscape is focused around imaginary town centers and transport links between towns. The first scenario is one in which all transport links are open and there are no barriers to restrict the spread, while the second scenario considers landscape barriers (implemented as cells with zero carrying capacity). Results show that if the barrier is thin enough and the radius of movement is large enough, the infection can still spread past the barrier into the rest of the susceptible population.

Although there have been many epidemiological CA models, a review of the literature found only one known CA model of Chagas' disease. Slimi et al. [79] implement a CA model for Chagas' disease in order to study how demography and dispersal of Chagas' disease vectors interact to produce variations in bug abundance in a village (covering approximately 1 $km^2$) in the Yucatan peninsula in Mexico. In this model each cell, which represents a fixed area of a village, can either be empty or occupied by adult bugs or larvae. It should be noted that this CA model does not take on the traditional form in which each cell represents an individual, but rather each cell may contain a population of vectors. Adult bugs disperse from the forest surrounding the village during the months of April to June. The model is used to investigate periodic infestations of the vectors into the village resulting in increased Chagas' transmission risk, represented as the number of cells where the bug abundance exceeds more than 1 individual. The transmission risk is between

10% and 40% during the infestation period and remains at that level several months after the immigration period of the bugs has ended. Results of this study may help researchers evaluate current disease control policies (e.g., the spraying of insecticide) around villages. We further mention Devillers et al. [23] who modeled the spread of *T. cruzi* strain types I and IV (formerly known as type IIa) using an agent based model (ABM). ABMs are an extension of CA models, but generally more complex. The agents (or individuals) interact according to a set of rules (similar to a CA), but are not required to update synchronously. Devillers et al. [23] construct a 2-strain model of *T. cruzi* where the agents are two types of hosts: humans and vectors. Infection stages are distinguished in the human hosts, while the vectors are either susceptible or infected. Results of the study show the necessity of a reservoir host species in the transmission cycle to obtain simulation results that agree with observations in nature. Furthermore, the study was designed to show the flexibility of the ABM in modeling situations in which several factors are unknown. In this way, the use of an ABM allows for adding or subtracting compartments in order to hypothesize several situations, such as is the case with this study.

The idea of invasion speed was first studied by Fisher [28] and later by Skellam [78] in the context of ecological reaction-diffusion equations, typically in one spatial dimension. The underlying assumption of dispersal modeled by diffusion is that each member of the dispersing population moves at random. Wave speed is calculated by applying a change of variables to the system of PDEs, resulting in a system of first-order ODEs, whose solutions can be analyzed using a standard approach to stability analysis of equilibrium values. The change of variables method assumes that if a traveling wave must exist, then it can be written as a function of a single argument: for example, in the one-dimensional case $u(x,t) = U(x + ct)$, where $c$ represents the minimum wave speed, and $U$ is positive and increasing. Although

many studies have been done with these models, in some dispersing populations, this idea of random movement may not be realistic. Reaction-diffusion models only produce waves traveling at a constant speed, where the front is either advancing or retreating. Other models used to study invasion speed use either integro-differential or integro-difference equations [42, 54, 58], in which the invasion speed can be calculated by linearizing the system. Systems of this type can produce accelerating waves, which for some populations may be more realistic.

Invasion speed of waves in cellular automata has not been extensively studied. The shape and speed of wave fronts are dependent on the underlying grid. In work by Schönfisch [76], the problem of calculating speed of wave fronts in a simple two-dimensional cellular automaton is addressed. In the epidemic automaton, a simple SI model is considered for two standard neighborhoods, the von Neumann neighborhood and the Moore neighborhood. In this type of classical deterministic cellular automaton, a susceptible individual becomes infected if at least some set number of its neighbors are infected, while in a stochastic cellular automaton the probability for a susceptible individual to become infected increases with the number of infectious neighbors. The shape of a given wave front of infected cells is studied over time. An initial front is determined by a key pattern and the speed of the wave front is determined by the time it takes for the grid points behind the front to reach the infected state. In this problem, distance is measured in units of cells, and time is measured in the time steps of the model. It is concluded that the speed of the front depends on the slope of the front, and multiple fronts can be found for different neighborhoods.

## 4.3 Methods

We will construct a spatially explicit epidemic model based on classical metapopulation models. The model will have a rectangular grid structure based on a two-

dimensional cellular automaton. On a broad scale, the grid is broken up into 3 patches, defined in Chapter 2, each containing specific host and vector populations based on the known habitat of each species in the model. Migration rates between cells will be calculated based on an adaptation of the method in Chapter 2. A deterministic system of ODEs with an SI transmission structure based on the Chapter 3 models will be used to generate numerical solutions for each of the grid points.

The basic model structure is based on the system of ODEs given in Chapter 3 determined by the patches defined in Chapter 2. Each cell in the CA represents a specific area of landscape containing several populations of hosts and vectors. The hosts in the model are the raccoon (*Procyon lotor*) and the southern plains woodrat (*Neotoma micropus*), and the vectors are 2 *Triatoma* species, *Triatoma gerstaeckeri* and *Triatoma sanguisuga*. The hosts and vectors in the model make up 2 distinct sylvatic cycles, *T. gerstaeckeri*-woodrat and *T. sanguisuga*-raccoon. The cycles overlap in patch 2 (cells) where there is also a *T. sanguisuga*-woodrat cycle. Each cell contains up to 2 host species, and up to 2 vector species depending upon the cycles specific to the geographic region containing the particular cell (see Appendix for equations).

The modeling framework is designed by superimposing a grid onto the geographic region identified in Chapter 2. A figure of the grid overlaid on the geographic region can be seen in Figure 4.1. In order to allow for high spatial resolution we wish to make the cells as small as possible while also large enough to treat the populations as continuous via differential equations. Because the lowest population density (raccoons in patch 2) is approximately 7.3 raccoons/km$^2$, we define each cell as a 26.5 km by 26.5 km area of land, with approximately 96 raccoons per cell in patch 2. The rules of the CA are determined by solving a large system of ODEs, with a range of 4 to 8 equations per cell (depending on the number of populations of host and vector in each cell) to be solved, for a total of 9,376 equations.

The populations between cells are connected through migration of vectors. Migration rates will be calculated using the framework described in Chapter 2, but will be adapted to allow for boundary segments to be much shorter than the maximum dispersal distance. The method will also be adapted to calculate migration rates across boundaries defining more than 2 patches (or cells in this case). The vectors migrate within a von Neumann neighborhood, which comprises the four cells adjacent to the central cell. Due to the lack of sufficient data on vector migration, we will consider different migration scenarios pertaining to the direction of migration. In the simplest case, we will consider the possibility that vectors have no preference for direction of migration, in which migration into each of the four orthogonally adjacent cells occurs at the same rate. We will also consider cases in which vectors have a preference for direction of migration, which implies that the rate of migration will be higher into certain cells of the neighborhood than into others.

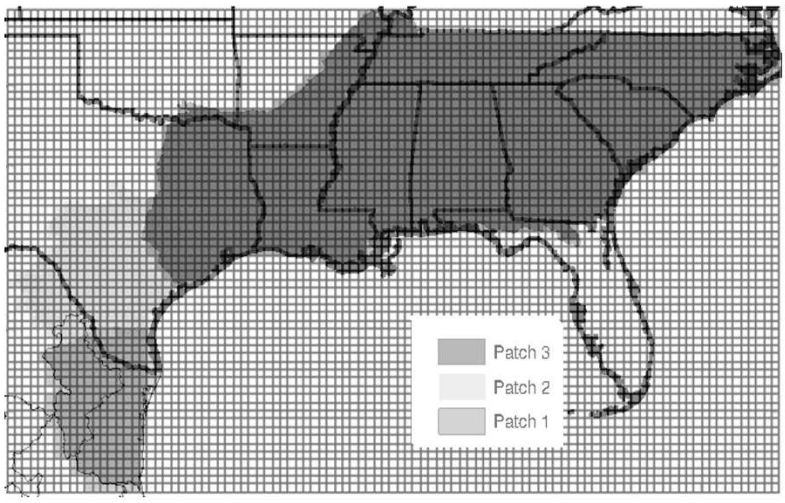


Figure 4.1. Grid framework with cells 26.5 km by 26.5 km.

To determine the effect of vector migration on invasion of *T. cruzi* into sylvatic populations, we will consider a scenario in which a hypothetical strain of *T. cruzi* is introduced into a certain location of the grid. To date, there are 6 strain types of *T. cruzi* identified, but only types I and IV have been identified in the U.S. Some differences have been noted in strains, such as host specificity and rates of vertical transmission in hosts [69]. Furthermore, strain type I is endemic among vector-host cycles in South America and Mexico, and is linked with Chagas' disease. Although both strains are present in the region studied, our focus is not on competition between strains. Climate change leads to vector habitat changes over time which gives cause for concern for invasion of Chagasic strains. In a recent study, Yabsley et al. [94] report higher *T. cruzi* prevalence in South Texas than has previously been reported. The present study uses the CA to measure how vector migration determines invasion speed in such a scenario.

We will consider a one-strain model in which strain I is introduced into patch 1 coming from a region in which it is already endemic. The strain will be introduced into an uninfected (susceptible) vector population. Upon solving the large system of DEs, time-series data will be determined for each grid point. To investigate the invasion speed of the epidemic, we first determine a threshold prevalence level for the presence of the epidemic in a given cell. Invasion speed will then be calculated based on the time it takes for a certain number of cells to reach the threshold prevalence level (units in km/yr). The calculation of the invasion speed will be based on two different methods. We use the speed of the wavefront decomposed into cardinal directions to determine the speed in other directions. We will also consider the method outlined by Schönfisch [76], which utilizes the slope between two grid points on the initial wave front. We will investigate how varying vector migration rates affect the invasion speed of the epidemic as well as determine invasion speed of the infection for each species,

as well as by patch (geographical location), to begin to disentangle the effects of the various factors affecting the speed of the invading strain. In the next section we will develop the method to calculate migration rates for the CA based on the framework given in Chapter 2.

## 4.4 Migration rates for cellular automata

The framework developed in Chapter 2 is designed to translate information about local dispersal into migration rates across a large geographical region. More specifically, the method is used to calculate migration rates for vectors that cross patch boundaries. This method requires the hypothesis that the patch boundary lengths be much longer than the maximum vector dispersal distance. In order to get high spatial resolution, the CA cell lengths will be much shorter than the maximum vector dispersal distance (for the *Triatoma* vectors in this study). Furthermore, the method in Chapter 2 does not take into account the possibility in which 3 patches meet at a corner, as is the case in the grid model presented here.

In general, the method will be adjusted by apportioning all dispersals that exit the cell among the adjacent neighbors by angle. There will be 4 migration rates calculated for a given cell, one for each von Neumann neighbor.

Given an originating dispersal point in a given cell (in polar coordinates) we define the four boundary segments as $f_i$, and the four corner angles as $\eta_i$, for $i = 1, 2, 3, 4$, where each angle is defined counterclockwise from the positive $x$-axis, so that $0 < \eta_i < 2\pi$ (see Figure 4.3). In the global (Cartesian) coordinate system, the cell of origin is in quadrant IV, i.e., the origin is the northwest corner of the cell. As in Chapter 2, in order to model a preferred direction of migration, the framework here uses a sequence of nested ellipses with common focus at the pole. The sequence

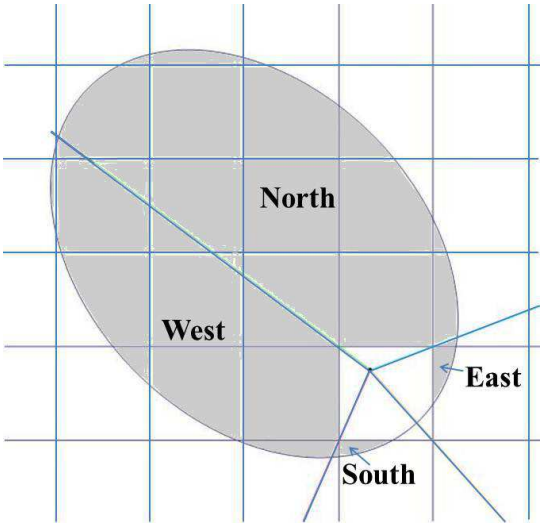


Figure 4.2. Sample regions of integration for a generic vector dispersal ellipse with preferred direction northwest. Note that the angle bounds for northward and westward migration are the corner angles while bounds for southward and eastward migration are defined by the corner angle and the angle at which the ellipse intersects the cell boundary.

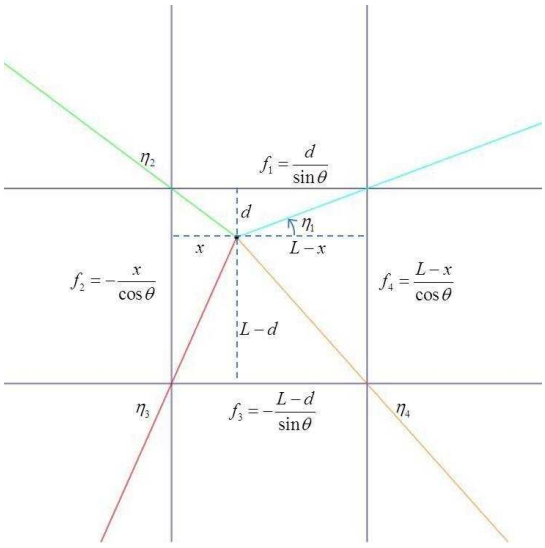


Figure 4.3. Sample cell showing corner angles and equations of boundary segments for a given point of dispersal, $(x, y)$..

of ellipses can be denoted $\{r_j\}_{j=0}^{n}$, where $r_n$ is the outermost ellipse. The formula for a given ellipse, in polar coordinates, is

$$r_j = \frac{b_j\sqrt{1-e^2}}{1 - e\sin(\theta - \theta_0)},$$

where $\theta_0$ is the indicator for preference of direction and is measured $\pi/2$ radians clockwise from the outward normal of the patch boundary, and each $b_j$ is the lateral dispersal distance perpendicular to the preference of direction. We then define a dispersal distribution function, $\Psi(r, \theta)$ (in proportion of vectors per square kilometer), to be piecewise constant on the elliptical rings formed by the nested ellipses. The strength of preference for a given direction is modeled using the eccentricity, $e$ of the ellipse.

For a given point $(x, y)$, we define the proportion of dispersals that cross patch boundary $f_i$ as

$$\tilde{M}_i(x, y) = \int_{\theta_{\text{low}}}^{\theta_{\text{high}}} \int_{f_i}^{r_n} \Psi(r, \theta) r dr d\theta.$$

The angle bounds $\theta_{\text{low}}$ and $\theta_{\text{high}}$ for each direction $i$ depend on whether or not the ellipse intersects the boundary between the corner angles, $\eta_i$ and $\eta_{i+1}$. If the dispersal ellipse remains beyond the cell boundary, the angle bounds cover the entire boundary segment. Specifically, if $r_n(\eta_i) > f_i(\eta_i)$, $\theta_{\text{low}} = \eta_i$. Also, if $r_n(\eta_{i+1}) > f_i(\eta_{i+1})$, $\theta_{\text{high}} = \eta_{i+1}$. However, if the ellipse intersects the cell boundary, then the bounds will be determined by the angle(s) at which the ellipse intersects the boundary. For a given boundary segment $i$, the bounds $\theta_1$ and $\theta_2$ (where $\theta_1 < \theta_2$) define the 2 intersection points of the ellipse with boundary segment $i$. Then, if $\eta_i < \theta_1 < \eta_{i+1}$, then $\theta_{\text{low}} = \theta_1$, and if $\eta_i < \theta_2 < \eta_{i+1}$, then $\theta_{\text{high}} = \theta_2$. Figure 4.2 shows the regions of integration for a generic vector dispersal ellipse with preferred direction northwest.

The function $\tilde{M}_i(x, y)$ is the proportion of patch-crossing dispersals for a given $(x, y)$. Then to sum the dispersals crossing out of the cell in a given direction, we determine the migration rate

$$\bar{m_i} = \frac{m}{A} \int_{\max(-L, -r_{\max_i})}^{0} \int_{0}^{L} \tilde{M}_i dx dy,$$

where $L$ and $A$ are the cell length and area, respectively, and $r_{\max_i}$ is the maximum (perpendicular) distance a vector can disperse and still reach the patch boundary. We note here that $\tilde{M}$ has undefined bounds at 2 points for each of the 4 neighboring directions, so then the double Cartesian integral is always improper, nevertheless, we still use the bounds 0 and $L$. The numerical method used to estimate the rates adjusts for these singularities appropriately. Several rates are given in Table 4.1, and a graph of the rates for varying preferred directions (for northward vector migration) is given in Figure 4.4.

Table 4.1. Specific rates for varying preferred directions, $e = 0.5$, rates in units 1/yr

| Northward vector migration | | |
|---|---|---|
| Preferred direction | *T. gerstaeckeri* | *T. sanguisuga* |
| No preference ($e = 0.5$) | 0.069 | 0.033 |
| North | 0.1310 | 0.0632 |
| East | 0.0138 | 0.0066 |
| South | 0.0097 | 0.0046 |
| West | 0.138 | 0.0066 |

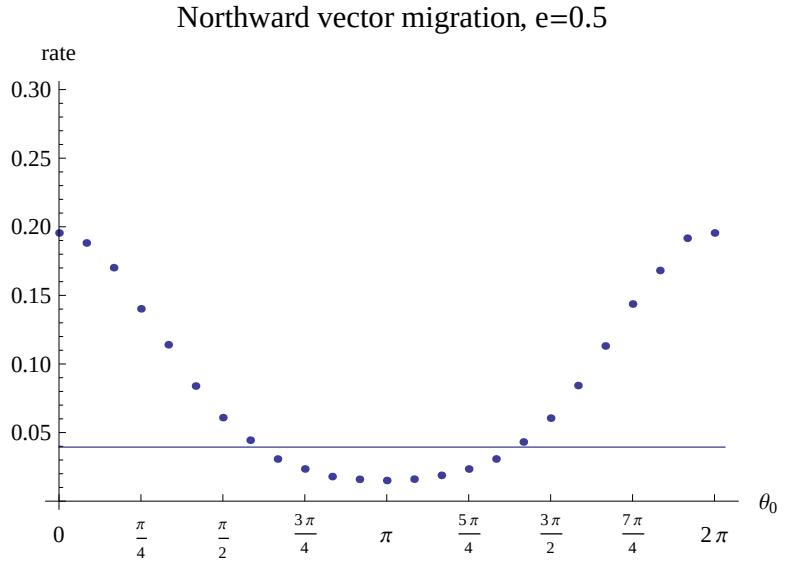


Figure 4.4. Northward migration rate as $\theta_0$ ranges from 0 to $2\pi$ compared with migration rate for no preference of direction (solid line) ($e = 0.5$).

## 4.5 Analysis

### 4.5.1 Determining threshold presence of epidemic

Once the infection enters the susceptible population, the proportion of infected individuals exhibits slower growth at the beginning of the outbreak, then accelerating, and finally approaching equilibrium, with the graph being sigmoid shaped. In order to determine the threshold presence of the epidemic, we consider two specific events during the invasion of the infection. We will determine when the epidemic has officially

arrived into the population, or when the proportion of infected individuals starts the period of accelerated growth, and the time at which the rate of growth levels off. The time it takes for these events to occur in each cell will be calculated as a step to ultimately derive a measure of invasion speed of the epidemic throughout the region.

Each cell in the grid contains both vector and host populations. We define the two epidemic threshold prevalence levels as a proportion of final epidemic prevalence, and base each threshold on the prevalence of infection in the host populations in each cell. Because of the transient nature of the vectors, it is not practical to use the vector population infection prevalence to determine thresholds. In addition, the hosts are not assumed to migrate, and are biologically the more significant population to use as the indicator for arrival of the epidemic. Each threshold level will be determined by approximating two points on the solution curve for the proportion of infected host populations in each cell. We recognize that, although each curve is sigmoid shaped, each population will have slightly different threshold levels, but the level does not vary much between cells.

For each cell, the normalized time series data is split into 3 segments representing the initial growing phase, the accelerated growing phase, and the final flattening phase approaching equilibrium. The data is split by doing a 3-piecewise-linear regression in which each segment of the piecewise function represents one of the 3 phases. The function is constrained to be non-decreasing and the first and third segments must have zero slope. We also note that as the epidemic progresses throughout the region, some of the time series graphs (corresponding to cells very near to or very far from the geographic origin of the invasion) will have more data points at one end of the curve or the other. In order to get a better sampling of points in the curve, we truncate the series so that the linear regression will be performed on a more uniform set of curves. In order to automate this process, we use a fixed number (100) of data

points to the left and right of the time step when the epidemic reaches 50% prevalence level. In the cases in which there are not 100 data points to the left (for example), the left piecewise segment will begin at $t = 0$. This allows for the linear regression to be done on curves with the same shape. Once the function is determined, we use the $t$ coordinates from the breakpoints as indices to find the corresponding $I(t)$ values in the time series data. Those values, normalized by the final $I(t)$ values, are the threshold values. For cells in which there are 2 host species, two threshold values will be calculated, one for each species. For each cell, the threshold value for the host species reaching the epidemic threshold in the earlier time will be used as the epidemic threshold value.

An illustration of this process can be seen in Figure 4.5 with coordinates representing each estimated threshold level based on the intersection points of the 3 linear segments.

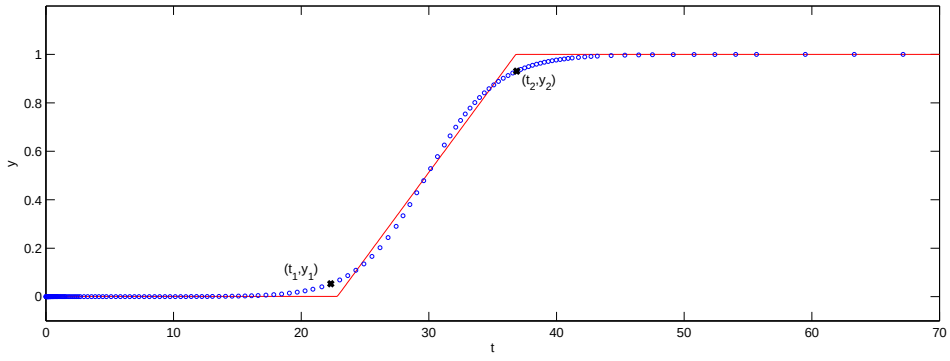


Figure 4.5. Host infection prevalence in a single (sample) cell, expressed as a proportion of final endemic prevalence, with threshold coordinates based on superimposed 3-piecewise linear regression.

### 4.5.2 Invasion speed

To determine the invasion speed of the epidemic, we consider methods which quantify a 2-D velocity vector. To define a velocity vector, one must specify the magnitude (or speed) and direction of travel. There are essentially two ways to quantify a 2-D velocity vector: the use of orthogonal velocity components, or via direct computation of magnitude and direction. In the first method, the overall speed and direction of the velocity can be determined by the resultant vector of the two cardinal components. Here we define the direction of the components to be northward and eastward, so then velocity in the southward and westward directions will be negative. In this section, we will derive a measure of invasion speed and direction using the methods mentioned above.

In theory, speed can be calculated by fixing distance and measuring the time to reach the distance, or by fixing time and calculating distance traveled over the fixed time period. Fixing distances is natural with the component method; thus we will use this approach in section 4.5.2.1 and 4.5.2.2.

#### 4.5.2.1 Decomposition into cardinal directions

One way to determine 2-D velocity is to resolve the vector into orthogonal components determined by using a fixed distance (the length of one cell), and calculating the time it takes for the epidemic to reach (using the threshold epidemic prevalence described in section 4.5.1) neighboring cells. We define the orthogonal components as northward and eastward velocities. The northward velocity of invasion at a point is calculated by dividing the distance between adjacent cell centers by the difference in first threshold times between the given cell and whichever of its north or south neighbors (if either) has the most recent earlier first threshold time than the given cell. If neither neighbor has an earlier threshold time, then the northward velocity

at that point will be 0. If both neighboring cells have the same threshold time (but less than the threshold time of the given cell), then the velocity will be calculated as northward. One must also consider boundary cases. If the given cell does not have a north or south neighbor in the geographical study area, then only one neighboring cell is to be examined, in which the above criterion will determine the appropriate calculation. This process will be repeated for every cell, and the eastward velocity will be calculated in the same manner using the east and west neighboring cells. In order to get a final measure of overall invasion speed (by direction), we will take an average over all cells. To get a patch-specific speed, we will take an average over all cells within a given patch.

Table 4.2. Threshold times

| | | | | |
|---|---|---|---|---|
| 5 | 2 | 6 | 3 | 7 |
| 3 | 6 | 4 | 7 | 9 |
| 4 | 8 | 9 | 12 | 5 |
| 0 | 3 | 5 | 7 | 9 |

Table 4.3. Eastward velocities (cell diameters/$yr$

| | | | | |
|---|---|---|---|---|
| $-\frac{1}{3}$ | 0 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| 0 | $\frac{1}{3}$ | 0 | $\frac{1}{4}$ | $\frac{1}{2}$ |
| 0 | $\frac{1}{4}$ | 1 | $\frac{1}{3}$ | 0 |
| 0 | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

Table 4.4. Northward velocities (cell diameters/$yr$)

| | | | | |
|---|---|---|---|---|
| $\frac{1}{2}$ | 0 | $\frac{1}{2}$ | 0 | 0 |
| 0 | $-\frac{1}{4}$ | 0 | $-\frac{1}{4}$ | $\frac{1}{4}$ |
| $\frac{1}{4}$ | $\frac{1}{5}$ | $-\frac{1}{5}$ | $\frac{1}{5}$ | 0 |
| 0 | 0 | 0 | 0 | 0 |

A visual illustration of this process (using a sample threshold time matrix) can be seen in Tables 4.2-4.4. Regarding the special cases mentioned previously, we note the zero entries in the first column of the eastward velocity matrix (with the exception of the first entry) and the last row of the northward velocity matrix. These entries represent cells on boundaries, in which only one neighboring cell was examined. In the case when the threshold times for neighboring cells are equal (but less than the given cell), we use the positive velocity. The units of each entry in Tables 4.3 and 4.4 are in cell diameters/$yr$. The values for speed can be determined by multiplying each entry by the length of each cell (in km), so the speed will be in units of km/yr.

#### 4.5.2.2 Magnitude and direction

The work by Schönfisch [76] provides a useful framework for decomposing speed of a wavefront in a cellular automaton by directly calculating magnitude and direction. Schönfisch first describes a *path* of cells to be an ordered set of cells such that each cell in the path is in the predefined neighborhood (e.g., von Neumann or Moore) of the previous cell in the path. Two cells are said to be *connected* if there exists a path between them. A *front* is described as the boundary of the set of connected cells. Once a front is established, the *key pattern* of the front is observed. The key pattern is defined based on observing the relative positions of the cells in the front. The relative positions of the cells are defined by a sequence of ordered pairs $(a_i, b_i)$, where $a$ and $b$ represent the horizontal and vertical shifts from one cell to another. As an example, the author describes the sequence $\{(1,0),(1,0),(1,0),...\}$ as a horizontal front. If the sequence is periodic, the key pattern is determined by the period of the sequence of relative positions. The speed of the front is determined by constructing a line parallel to the front at a fixed distance away from the front. Cells on the new line are marked, and the time is measured until the front reaches those cells. Because the front will most likely be irregular, not all of the marked cells will be reached at the same time step. In Schönfisch's method, a sample of 10 cells on the new line are chosen, and speed is calculated using the average time until each cell is reached.

The method developed in this section draws on Schönfisch's framework while establishing definitions that differ in some instances from Schöfisch's, especially in the identification of the pattern of the front as well as application of the method for an irregular shape on the grid. As the epidemic spreads, the shape and direction of the front changes due to the geometry of the grid. Thus, to calculate the invasion speed of the epidemic, we will define the front as a function of time. It is expected that as

the epidemic spreads through the region, the fronts will, 1) be approximately linear, except at patch boundaries, and 2) have similar shapes and slopes within a single patch. We note here the patch boundaries may change the shape of a front in a similar way that light is refracted as it passes from one medium to another. Because each patch has different populations, including vector populations migrating at different rates, the epidemic thresholds differ slightly as the epidemic passes through patch boundaries, which may alter the shape of the front. Because of this, there may at times be two disconnected linear fronts, in particular (given the geography of the region under study here), one front in patch 2 and one in patch 3. In this situation, we will only use the front cells (path) in patch 3, because this is the more relevant direction of the epidemic. In order to identify the front at a particular time, we need to determine the key pattern of the front. We define a front to be the set of cells such that each cell has at least one neighbor that has reached the epidemic threshold and at least one neighbor that has not yet reached the threshold, using the von Neumann neighborhood. This method of determining front cells will ensure that isolated cells that have reached epidemic threshold "ahead" of their neighboring cells not be included in the front. After the front is determined for each time, linear regression will be used on the front cells to approximate a slope for the front. The direction of travel for the front of invasion will then be perpendicular to the linear front.

To maintain consistency with the method of using orthogonal velocity components described in section 4.5.2.1, in this section we will calculate invasion speed by fixing a distance and calculating the time it takes for the epidemic to travel the fixed distance. Once the linear front is identified, a fixed distance will be chosen. Because the method of orthogonal components uses a distance of one cell length, the fixed distance in this method should be as close to this length as possible to be consistent.

However, because the front may not be exactly linear, the fixed distance should be at least the maximum distance between any of the front cells and the linear front line so that nonlinearities in the front do not dominate (or potentially even pre-empt) the calculation of the time taken to reach the advance front. We denote the maximum distance between any of the front cells and the original front line as $D$. In order to satisfy both conditions, we choose the fixed distance to be $max\{2D, (2 * \text{cell diameter})\}$ (in km), A new line will be drawn at that fixed distance away from the identified front line (and thus parallel to it).

When speed is calculated by fixing a distance, the time it takes to reach the distance can be easily calculated for one quantity or particle. However, the time it takes for an epidemic on a landscape to reach a specified distance is not straightforward. Here, we will measure the time it takes until the front reaches each cell through which the new line passes. In general, it will take a different amount of time for the front to reach each individual cell on the new line. Thus, the average time taken for the front to reach all of the cells on the new line will be used. This process can be used to obtain a velocity vector global in space, as a function of time. To obtain an average over the entire time domain, we calculate the average value of the velocity function. We note here that the magnitude and direction method of calculating invasion speed begins with average velocities over one spatial dimension, while the method in section 4.5.2.1 derives velocities locally by cell, and can be averaged in various ways.

## 4.6 Results

In this section, we present results for varying migration rates (based on preference of direction) and provide comparisons across patches and species, as well as present overall statistical measures, such as the mean, median, and standard deviation. We have presented two methods for calculating invasion speed: 1) the use

of cardinal direction (i.e., vector components) and 2) the front (Schönfisch) method. We also compare results for the different methods and report any differences (and similarities) of the numerical results for each method.

### 4.6.1 Basic trends in speed and direction

We assume that the infection is introduced into a cell in patch 1 (in the southwest portion of the patch), while all other cells have only susceptible populations. The infection is spread from cell to cell via migration of vectors. We first present results for the case where vectors have no preference for direction of migration. The results will be given in increasing levels of specificity. We start by reporting values for average overall speed and average speed by patch. In order to avoid oversampling of patch 2 data (because there are two host species present), we calculate patch 2 average speed using one host only. We use the same criteria mentioned in section 4.5.1, in which the host species that reaches the threshold level at the earlier time step will be used. After reporting basic statistics, we look at speed vs. time graphs in order to capture acceleration of the epidemic. This graph, along with the direction field plot, will allow us to understand not only when the speed changes, but where the speed changes.

#### 4.6.1.1 Component method

Each of the methods for measuring invasion speeds relies on determining threshold times representing the beginning and end of the period of fast epidemic growth. Because these threshold values are approximated in each cell, they will vary slightly from cell to cell. We will give results for using the cell specific thresholds as well as using a single (average) threshold over all cells. Basic statistical measures for using cellwise thresholds and average thresholds are given in Tables 4.5 and 4.6. Even

Table 4.5. Statistical measures for speed (no preferred direction of migration) using cellwise threshold prevalence, units km/yr

| Statistical measures | | | |
|---|---|---|---|
| Patch | Mean | Med. | SD |
| All | 7.33 | 6.15 | 40.49 |
| 1 | 8.28 | 8.47 | 1.53 |
| 2 | 4.34 | 4.06 | 1.82 |
| 3 | 7.32 | 6.13 | 46.03 |

Table 4.6. Statistical measures for speed (no preferred direction of migration) using average threshold prevalence of 0.07, units km/yr

| Statistical measures | | | |
|---|---|---|---|
| Patch | Mean | Med. | SD |
| All | 6.27 | 6.13 | 1.62 |
| 1 | 8.45 | 8.41 | 0.77 |
| 2 | 4.05 | 3.85 | 0.79 |
| 3 | 6.28 | 6.13 | 0.93 |

Table 4.7. Statistical measures for speed (no preferred direction of migration) using front method, units km/yr

| Statistical measures | | | |
|---|---|---|---|
| Patch | Mean | Med. | SD |
| All | 6.64 | 6.50 | 1.41 |

though the cellwise threshold prevalence values do not vary by much, even a small change in threshold times may produce outliers in speed (occurring in patch 3), as reflected in the higher means and standard deviations in Table 4.5. Therefore henceforth all graphs and other results will be given using times for average threshold prevalence, unless noted otherwise.

In Figure 4.6, we observe the change in speed from 2 different viewpoints (speed vs. time and by location). At the start of the epidemic, the speed increases through patch 1 and drops as the epidemic crosses into patch 2. This drop in speed is due to several factors. The estimated migration rate for *T. sanguisuga* (only in patches 2 and 3) is roughly half that of *T. gerstaeckeri*, and in patch 2 the population density of *T. sanguisuga* is roughly twice that of *T. gerstaeckeri*. Once the epidemic reaches patch 2, the speed is roughly constant (and lower than in patch 1). As the epidemic moves into patch 3, the speed increases slightly, remaining fairly constant throughout the

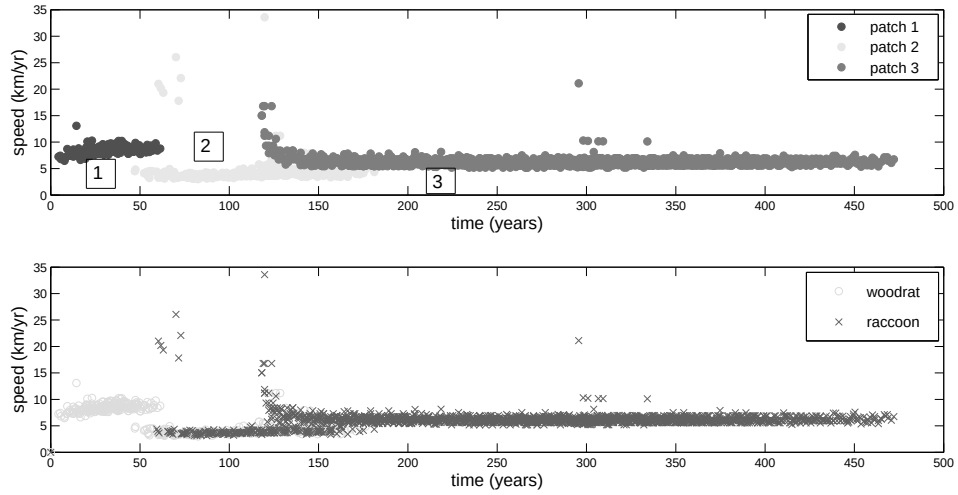


Figure 4.6. Speed vs. time with results by patch (top) and species (bottom)).

remainder of patch 3. Although *T. sanguisuga* migrate at a lower rate, we still observe higher speeds in patch 3 when compared with patch 2 even though the only vector species in patch 3 is the slower migrating *T. sanguisuga*. The increase in invasion speed in patch 3 is due to the higher contact rates between *T. sanguisuga* and raccoon in patch 3 versus patch 2. We also note that the higher speeds in patch 2 correlate to the initial spread of the infection through the raccoon population (Figure 4.6). In fact, if the model is adjusted to have only one host and one vector, the speed remains approximately constant throughout the entire region.

In addition to speed, we also study the direction of invasion. We can observe the change in direction over time, as well as with respect to space (physical location). As observed in Figure 4.7, the direction trends northward, as the epidemic moves from patch 1 to 2, then settles to a northeastward direction (as the epidemic moves through patch 3). We note that there are some directions that do not follow this trend. These are cases in which the front may be moving in two directions, such as at the beginning of the epidemic when the infection travels both northeast and

southeast (to spread to the very southern tip of patch 1) and in the transition from patch 2 to 3.

Results for direction (and speed) with respect to space can by found by using a direction field plot. The direction field plot (see Figure 4.8), shaded to indicated magnitude, confirms the trends seen in Figures 4.6 and 4.7. The speed decreases in patch 2 (the lighter shaded vectors), with a decrease in speed at the patch 1/2 boundary, and an increase in speed at the patch 2/3 boundary. The direction of the epidemic in Figure 4.8 reflects those observed in Figure 4.7, showing a northern trending direction into patch 2 then turning northeast into patch 3. The overall average direction is 33.13° north of due east.

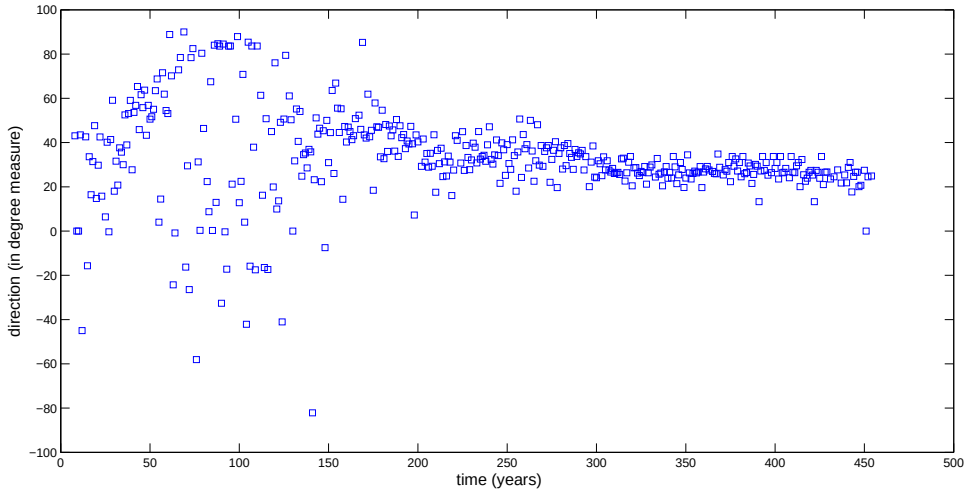


Figure 4.7. Direction vs. time for component method. Direction is measured in degree measures counterclockwise from due east with 90° representing north.

Because the estimate for migration rate depends solely on the vector dispersal rate due to maturation, we also consider the effects of the dispersal rate on the invasion speed. We may consider scaling the dispersal rate $m$ by a factor $k$, where if $k = 1$, the

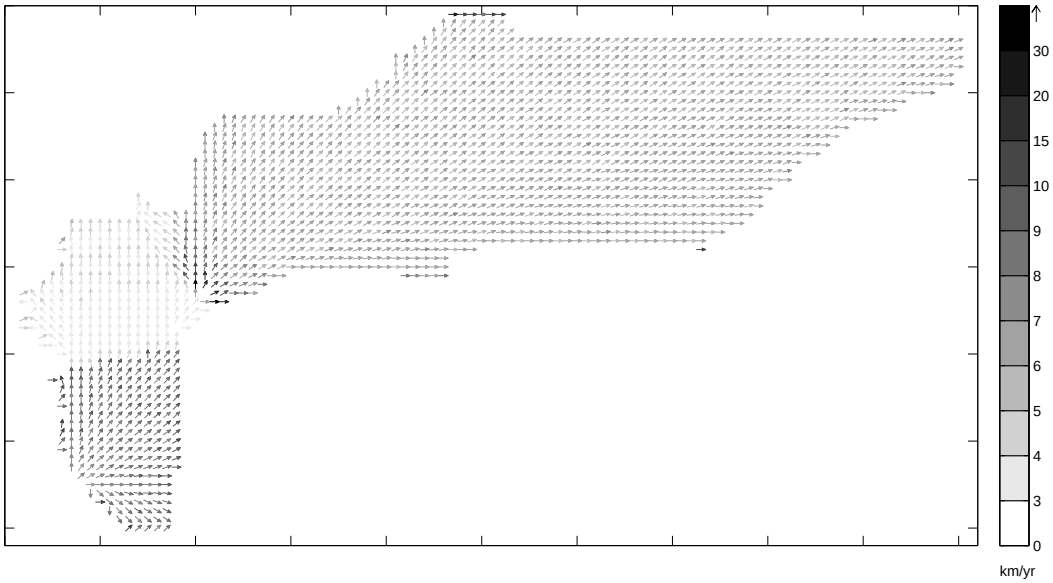


Figure 4.8. Direction field for no preference of direction ($e = 0$) (shading indicates magnitude).

dispersal rate is the current estimated rate used in this study. Based on the dispersal rates estimated in Chapter 2 (and used here in this study), we can conclude that if vectors disperse every 1.8 years (for *T. gerstackeri*) and 3.7 years (for *T. sanguisuga* with no preference for direction, the epidemic advances at approximately 6.3 km/yr. However, if vectors disperse more often than estimated here, we wish to see to what extent invasion speed increases. In Figure 4.9, we report estimated invasion speeds as $k$ ranges between 0 and 15. After observing the graph, using power law regression, we are able to determine that invasion speed increases at $k$ to the 0.44 power.

#### 4.6.1.2 Front method

The front method used to calculate invasion speed is based on the concept that at any given time, the epidemic can be identified by a front (a set of cells with

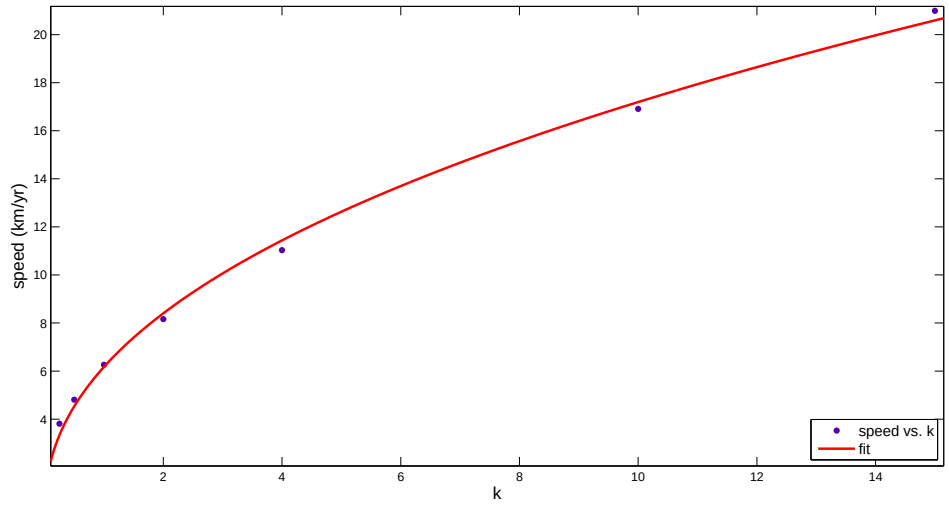


Figure 4.9. Power law regression fit for speed vs. migration rate scaling factor $k$ with $0 < k \leq 15$. The equation for the fit is $6.17k^{0.44}$..

characteristics given in section 4.5.2.2). As a reminder, in this method, speed is calculated by fixing a distance (perpendicular to the linear front) and calculating the time it takes for the front to "move" the given distance. This can be done by calculating the line of best fit through the front and a new line parallel to the front line at the fixed distance away. Cells touching this new line represent the future front. We determine the speed of the invasion at any given time by dividing the fixed distance by the average time it takes for the front to reach the new front line.

Results from using the front method are consistent with results from the component method (differing by no more than 5%) , but we note that the speed changes in the front method in a more continuous way, rather than abrupt drops or increases in speed, as observed in the component method. Figure 4.10 shows the differences in the speed changes from patch to patch.

We can also observe a graph of the direction of the epidemic (in angle measures) as a function of time (see Figure 4.11). As expected (and consistent with the compo-

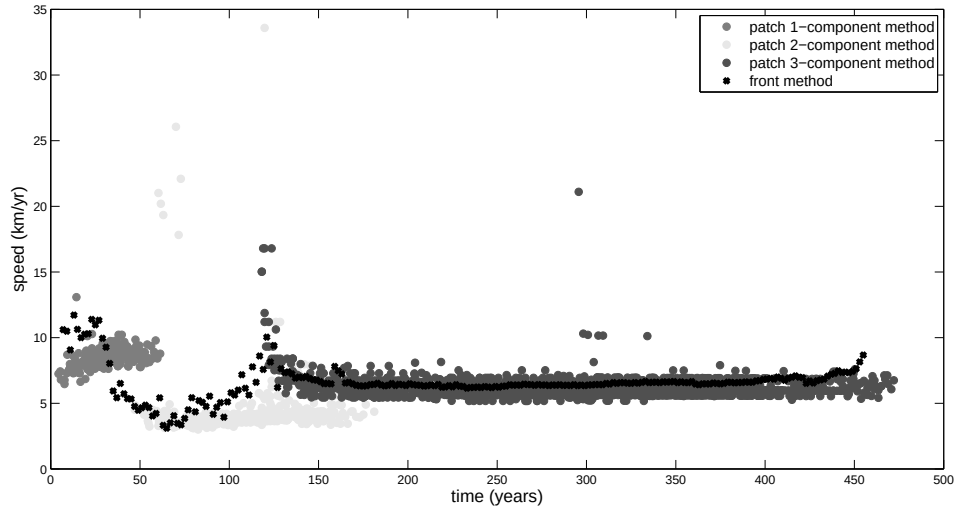


Figure 4.10. Speed vs. time (comparison of methods).

nent method), the direction of travel begins roughly northeast and as the epidemic moves toward patch 2, the direction becomes increasingly northward, moving back toward the northeastern direction. The overall average direction for the front method is 43.65° north of due east.

### 4.6.2 Role of preferred direction

When vectors have a preferred direction of dispersal, as expected, the dynamics of the invasion will change depending on the preference for a particular direction (and strength of preference). In the absence of a preferred direction of vector migration, the epidemic naturally spreads in the direction of the geographical region, and is fastest in patch 1 and slowest in patch 2. When vectors migrate with a preferred direction, these trends may not be the case. In Table 4.8, we report the average speeds and directions overall and by patch with 8 Moore-neighborhood directions (with $e = 0.5$) and no preference for direction in the center. For example, if the preferred direction is

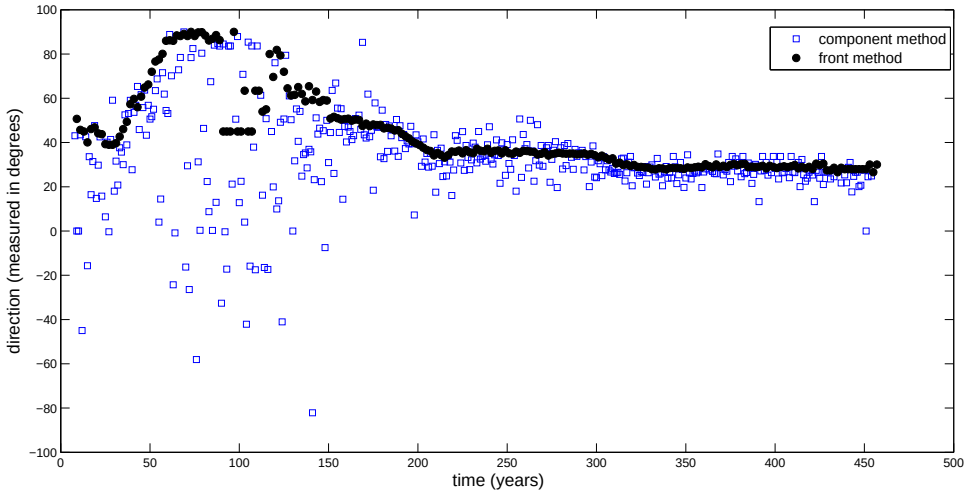


Figure 4.11. Direction vs. time (both methods represented) with direction represented as degree measured counterclockwise from due east, with 90° representing north.

east, the highest speed occurs in patch 3 (see Table 4.8) primarily due to the preferred direction of migration alignment with the geography of the region.

When comparing the effects between the different preferences, we observe the speed is greatest overall if the preferred direction is north to east, and lowest if the preferred direction is south to west. When compared with no preference for direction, the overall average speed is increased by over 50% when the preferred direction is east, while it is decreased by 14% if the preferred direction is west. We note, however, that even though overall average speed is greatest when the preferred direction is east, it is important to look at speed changes on a patch by patch basis. As an illustration, we compare results for changes in speed for preferred directions north and east. If the preferred direction is north, the patch 2 average speed increases, and decreases if the preferred direction is east. The patch 3 average speed increases (from the no preference for direction) in both cases, but with a slightly greater increase if the preference is east (71% vs. 46% increase).

Table 4.8. Average speeds (in units km/yr) and directions (in degree measures north of due east) by patch and overall. When a preferred direction is given, $e = 0.5$.

| **Northwest** | | | **North** | | | **Northeast** | | |
|---|---|---|---|---|---|---|---|---|
| | speed | direc. | | speed | direc. | | speed | direc. |
| overall | 5.32 | 17.06° | overall | 8.91 | 25.13° | overall | 9.03 | 37.41° |
| patch 1 | 6.56 | 15.46° | patch 1 | 9.18 | 17.16° | patch 1 | 11.41 | 32.79° |
| patch 2 | 5.44 | 66.33° | patch 2 | 7.18 | 78.06° | patch 2 | 5.34 | 96.47° |
| patch 3 | 5.11 | 9.59° | patch 3 | 9.15 | 17.97° | patch 3 | 8.89 | 28.82° |
| **West** | | | **No Preference** | | | **East** | | |
| | speed | direc. | | speed | direc. | | speed | direc. |
| overall | 5.41 | 19.23° | overall | 6.27 | 40.83° | overall | 9.77 | 49.13° |
| patch 1 | 5.44 | 24.84° | patch 1 | 8.29 | 37.25° | patch 1 | 9.68 | 51.07° |
| patch 2 | 4.12 | 68.31° | patch 2 | 4.03 | 91.62° | patch 2 | 3.53 | 108.19° |
| patch 3 | 5.32 | 10.78° | patch 3 | 6.28 | 33.39° | patch 3 | 10.74 | 39.63° |
| **Southwest** | | | **South** | | | **Southeast** | | |
| | speed | direc. | | speed | direc. | | speed | direc. |
| overall | 3.96 | 45.23° | overall | 4.51 | 69.87° | overall | 5.35 | 67.01° |
| patch 1 | 5.74 | 39.12° | patch 1 | 5.75 | 53.51° | patch 1 | 7.38 | 58.36° |
| patch 2 | 3.31 | 93.55° | patch 2 | 3.21 | 108.56° | patch 2 | 2.56 | 111.04° |
| patch 3 | 3.75 | 38.53° | patch 3 | 4.57 | 66.11° | patch 3 | 5.53 | 61.34° |

The direction of invasion is also affected by vectors migrating with a preferred direction. Without vector preferred direction, the epidemic naturally spreads northward into patch 2 and in the northeast direction through patch 3. Because of the natural shape of the regions, the direction of invasion cannot be greatly affected, but there are specific effects on the direction changes that should be noted. We illustrate these effects using the northeast preferred direction. If the preferred direction is northeast, we would expect the epidemic to essentially be pulled into the northeastern direction via the higher northern and eastward vector migration rates. This can clearly be observed in patch 1 (see Figure 4.12). In the absence of a preferred direction, the epidemic naturally will move northward (Figure 4.7 in section 4.6.1), but if the preferred direction is northeast, the epidemic moves through patch 1 primarily in

the northeastern direction. However, this effect is not as clear in locations in which the patch shape and boundaries alter the direction of the invasion. For example, when the epidemic reaches patch 2, the vectors migrating will move the epidemic northeast as quickly as the patch geography will allow (along the patch borders). However, this means that at some point the front will necessarily spread into the remainder of patch 2, and will do so in the northwestern direction, albeit at a much slower speed. This effect can also be seen in Figure 4.14 by observing the location of the front cells at a particular moment in time in which the epidemic is being pulled along the patch 2/3 border.

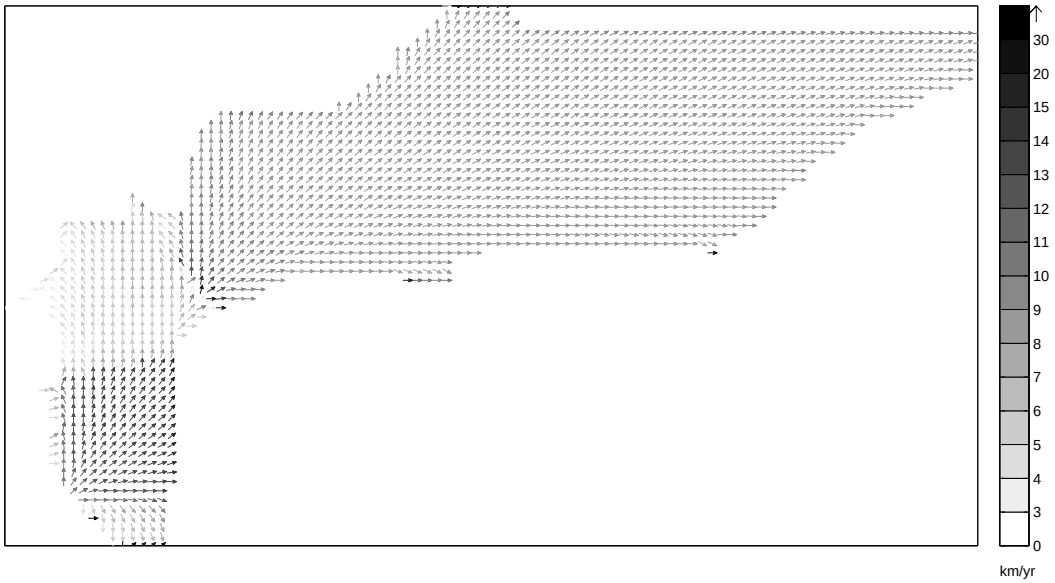


Figure 4.12. Direction field plot for northeast preferred direction of migration ($e = 0.5$). Shading indicates overall higher speeds (when compared with Figure 4.8), and the epidemic is being pulled into the direction of preference (this effect is most noticeable in patch 1).

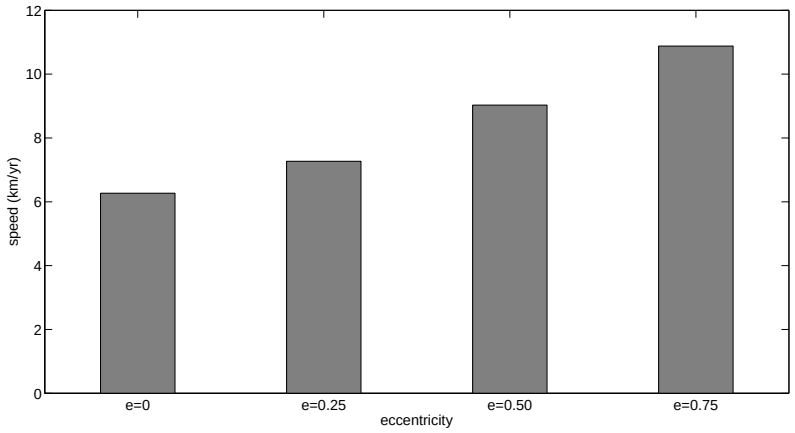


Figure 4.13. Bar graph of average invasion speed vs. degree of preference (eccentricity) with preferred direction northeast.

To capture how the strength of preference increases (or decreases) speed, we consider degrees of strength of preference for a particular direction. To illustrate the effects, we consider a northeast preferred direction with 3 increasing strengths of preference indicated by eccentricity of the dispersal ellipse. As expected, the stronger the preference, the greater the deviation from the results for no preference for direction. If the strength of preference is low ($e = 0.25$), the speed increases by approximately 16%, when compared with no preference for direction, while if the strength of preference is high ($e = 0.75$), the speed increases by over 70%. Table 4.9 summarizes the results. We also observe the increase in overall speed using a bar graph of speed vs. strength of preference (Figure 4.13).

## 4.7 Conclusions

In this study, we develop a spatially explicit model to study the spread of *T. cruzi* in sylvatic host and vector populations across a geographical region separated into patches. We consider the scenario in which a hypothetical strain of *T. cruzi*

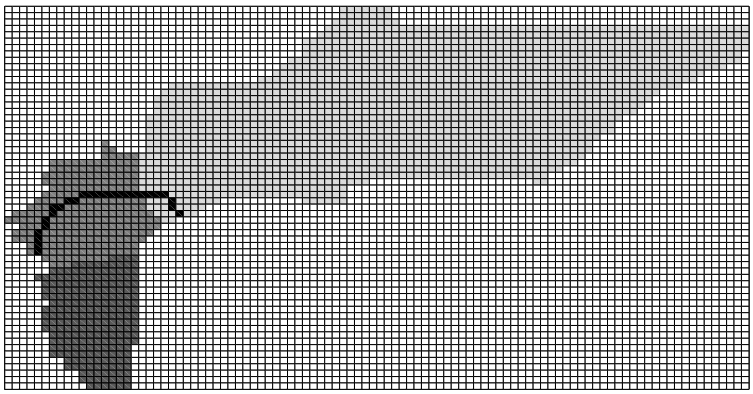

Figure 4.14. A given front for northeastern preferred direction of migration. Colors indicate patches. Here the front signals a rapid expansion into patch 3 and a slower growth into the northwest portion of patch 2..

Table 4.9. Average speed (units km/yr) and direction (degrees north of due east) overall and by patch for varying levels of eccentricity.

| $e = 0$ | | | $e = 0.25$ | | | $e = 0.50$ | | | $e = 0.75$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | speed | direc. | | speed | direc. | | speed | direc. | | speed | direc. |
| all | 6.27 | 40.83° | all | 7.27 | 38.52° | all | 9.03 | 37.41° | all | 10.88 | 34.64° |
| 1 | 8.29 | 37.25° | 1 | 9.67 | 35.26° | 1 | 11.41 | 32.79° | 1 | 17.07 | 28.51° |
| 2 | 4.03 | 91.62° | 2 | 4.33 | 93.76° | 2 | 5.34 | 96.47° | 2 | 6.05 | 94.28° |
| 3 | 6.28 | 33.39° | 3 | 7.44 | 31.63° | 3 | 8.89 | 28.82° | 3 | 10.64 | 26.18° |

is introduced into the southwest part of the region in order to capture speed and direction of invasion across the region. The infection spreads among populations by way of migrating *Triatoma* vectors. We first estimate the vector migration rate by adapting the framework in Chapter 2 to develop a high resolution grid based model. Invasion speed in this model is defined using two distinct methods. The component method determines the orthogonal velocity components cell by cell, from which magnitude and direction can be calculated. The front method constructs a vector by first

determining magnitude and direction using a linear regression to approximate the front.

Results from both methods are consistent, yet each method has its advantages and disadvantages. The component method calculates invasion speed and direction for each cell in the grid, therefore is able to capture sudden changes in direction (and speed). However, because the method uses individual cell information, there can be significant variability in the results (outliers in speed and direction), in part due to numerical artifacts of the threshold time estimates. The front method gives a more holistic perspective by using a group of cells (front cells), as an estimate for a linear front at each moment in time. Because this method uses groups of cells rather than individual cells, it gives a more complete description of the direction of the epidemic. However, by the same token some information, such as when there is a split front (in two directions), may be missed. To use the front method, it is assumed that the front is approximately linear, when in reality there may be some nonlinear fronts. Nonlinear, irregularly shaped fronts occur when a front stretches across a patch boundary, or when the geometry of the grid changes abruptly. Although there should be few of these irregular fronts, the method is limited in its ability to capture only the approximately linear front.

In general, there are specific locations in which the epidemic changes speed (and direction). The speed in patch 1 is greatest, while the speed in patch 2 is lowest due to the greater density of slower moving vectors (*T. sanguisuga*) and lower contact rates in patch 2. The epidemic naturally spreads northward into (and throughout) patch 2, and in the northeastern direction in patch 3. We also observe that if the preferred direction is more in alignment with the natural geography of the regions (as well as the natural spread of the epidemic given the invasion scenario), the increase in speed when compared with no preference of direction is greater than the decrease

in speed if the preferred direction is not aligned with the geography of the region. These results are consistent with the results in Chapter 2, when there is a greater change in migration rate if the preferred direction is out of the native patch.

The results in this chapter use vector dispersal and migration rates given in Chapter 2. These rates are based on estimates for what causes vector dispersal, which may be higher if vectors disperse more often than assumed in Chapter 2. The estimate for the dispersal rate is based on the rate at which vectors mature. We also assume that vectors may disperse for a maximum of 5 weeks. If vectors fly more often than this (i.e., are more impatient to find food), our estimates for invasion speed are a lower bound, especially if we also model host dispersal. In this model, we do not consider hosts to disperse due to the habitat constraints defined by the patches. In our model, the patch transitions are abrupt, reflected by the sharp changes in results in speed and direction. To mitigate these sharp changes in results, it may be possible to blend the patches in a smoother way. On a map, the ecoregion changes are abrupt, but the real geographic changes may actually be more gradual in some places. To incorporate gradual changes, the cell parameters defining the patch location would be blended, rather than change abruptly to be one ecoregion or another.

# CHAPTER 5

# CONCLUSION

The parasite *Trypanosoma cruzi* infects over 100 mammalian species in the United States, is transmitted via insect vectors from the *Triatoma* species, and is responsible for causing Chagas' disease. In the United States, Chagas' disease is named a Neglected Parasitic Infection. Although the disease is well-known in South America and Mexico, relatively little attention has been given to Chagas' disease in humans in the United States, much less in sylvatic settings. Although human infection with the disease is of major concern, there have been few human cases of Chagas' in the U.S. However, the disease remains endemic in sylvatic settings. There are several distinct transmission cycles, and in this study we consider the major cycles in the southeastern portion of the U.S. ranging to the very northern part of Mexico. Communication between the cycles occurs via the movement or migration of the insect vectors. There have been only a handful of studies regarding migration of *T. cruzi* vectors in South America, and even fewer on North American vectors. The majority of studies have been experimental field studies designed to obtain vector flight distance capabilities. Because of the limited amount of biological data, mathematical models provide an invaluable tool to the study of the effects of migration on prevalence of *T. cruzi*.

In order to examine the effects of migration of *Triatoma* vectors on *T. cruzi* transmission in sylvatic settings, we first develop a framework designed to calculate vector migration rates if certain local dispersal parameters are known. The framework developed in Chapter 1 is unique in the sense that it uses local small-scale

dispersal information to derive migration rates over a large scale geographic region. This framework also converts spatially continuous data to data that can be used in spatially discrete models. To study the effects of migration on *T. cruzi* transmission, we create a deterministic system of ODEs that describes *T. cruzi* transmission in specific sylvatic settings. This model is an SI compartmental model with 3 patches represented by distinct transmission cycles (overlapping in the second patch). The patches are connected via migration of infected and uninfected *Triatoma* vectors. The model is analyzed using standard threshold analysis techniques for epidemiological compartmental models. Numerical results (using the migration rate estimates derived in Chapter 1 and demographics and infection parameter estimates) are obtained to see effects of migration on prevalence of *T. cruzi* over time in each of the patches. Much insight can be gained from this model, but because of the importance of the spatial aspect of *T. cruzi* spread, we also develop a cellular automoton model to study the spatial (and temporal) dynamics of sylvatic *T. cruzi* transmission, from which results such as invasion speed of the epidemic can be investigated.

Each of the chapters contained in this study is a work toward the goal of understanding the dynamics of the spread of *T. cruzi* via the migration of *Triatoma* insect vectors. One of the first goals is to appropriately define the regions and develop a framework that can be used to estimate migration rates. Vectors may migrate for various reasons and have certain flying distance capabilities. The framework in Chapter 1 takes into account reasons for vector dispersal, flying distance capabilities, preferred direction of dispersal (if any), and strength of preference. This framework is unique in its ability to take small-scale dispersal information that can be used in models with a more global perspective in population epidemiology and ecology.

The compartmental model in Chapter 2 models the dynamics of *T. cruzi* infection in 3 patches represented by 3 host-vector transmission cycles connected via

migration of vectors. In 2 of the patches, there is one host and one vector species, with an overlap patch in between them containing both hosts and both vector species. The goal of this model is to determine to what extent migration affects the prevalence of *T. cruzi*. In general, the model compares patch-wise prevalence levels in the absence of migration vs. vector migration connecting the cycles. Overall results indicate that the patch 2 dynamics are sensitive to increases in migration rate, and the dynamics in the patch are driven primarily by the *T. sanguisuga* vector population. Because the model is not spatially explicit, it is unable to capture the spread of sylvatic *T. cruzi* transmission across the actual geographic regions defined in Chapter 1.

The model in Chapter 3 is a spatially explicit, high resolution cellular automaton. Each cell in the automaton represents a specific host-vector transmission cycle according to its geographic location (patch) defined in Chapter 1 and contains a system of ODEs (specified by patch) based on the Chapter 2 model with migration serving as a link between the cells. Due to the large size of the dynamical system, this model is only analyzed numerically. One of the primary goals of this model is to develop a method to measure invasion speed. Two unique methods for calculating invasion speed are presented, each of which is able to capture speed and direction of the epidemic over time.

Results indicate that migration lowers $R_0$ overall (with a decrease in prevalence in patches 1 and 3), yet prevalence increases in the overlap patch. We recall that patch 2 has 2 hosts, and is gaining vectors from a region (patch 3) with a higher *T. cruzi* prevalence. Although patch 2 prevalence increases, results from the spatially explicit model indicate that the epidemic wave front travels slowest in this patch. Future work should focus on complex regions like patch 2, as well as regions in which the *T. cruzi* prevalence is lower, yet connected to regions with higher prevalence, as migration is likely to increase prevalence in regions of lower prevalence.

Although the cellular automaton model is spatially explicit, it is also continuous in time, and is based on the model presented in Chapter 2. The model is able to capture prevalence (by patch) over time. In order to compare the results from Chapter 2 and 3, we consider the invasion scenario hypothesized in Chapter 3 in which the infection is introduced into patch 1. If we consider this same scenario in the Chapter 2 model, we are able to compare to what extent the higher resolution spatial component of Chapter 3 enhances or changes the results. Some of the obvious differences in the models are the times it takes for the epidemic phase to begin in each of the models. Because the cellular automaton is spatially explicit, the time it takes for the epidemic to spread from one cell to another can be seen, whereas in the 3-cell automaton (Chapter 2 model), time progression of the epidemic within a given patch is not depicted. Thus, in general, the times in the cellular automaton model will be longer. As seen in Figures 5.1 and Figures 5.2, the epidemic takes longer to arrive in patch 3 as expected, but the Chapter 2 model predicts the growth phase in patch 3 to be much shorter than in the high resolution model. Using the 3-piece linear regression method described in Chapter 3, we can compare the durations of the growth phases for each model (see Table 5.1). One major difference in results from the two models is that the growth phase for patch 3 in the Chapter 2 model is shorter (by a factor of .28) than the patch 2 growth phase, while in the Chapter 3 model, the patch 3 growth phase is longer (by a factor of 2.8) than the patch 2 growth phase. This difference in result is due to the increased spatial resolution in patch 3 depicting the large size of the patch. When comparing the prevalence graphs over time for each patch, we can view in Figure 5.3 the progressions of each prevalence in each patch over time. We see that in the Chapter 2 model graphs (dashed lines), the epidemic arrives in each patch very close to the same time, whereas the Chapter 3 model depicts the natural (more realistic) spatial spread through the region. When viewed on the same time

Table 5.1. Epidemic growth phases for Chapter 2 and Chapter 3 models. Units are in years.

| **Epidemic Growth Phase** | | | |
|---|---|---|---|
| | | start time | duration |
| patch 1 | Chapter 2 | 6 | 11 |
| | Chapter 3 | 15 | 42 |
| patch 2 | Chapter 2 | 24 | 42 |
| | Chapter 3 | 74 | 101 |
| patch 3 | Chapter 2 | 38 | 12 |
| | Chapter 3 | 143 | 287 |

scale, the different shapes of the graphs can also be observed. It can be seen that the prevalence vs. time graph in patch 3 from the Chapter 3 model is not sigmoid shaped like the others. This suggests a slower initial rate of increase in prevalence in patch 3 when compared with the other model, which predicts a faster rate of increasing prevalence in patch 3 (compared with the other 2 patches).

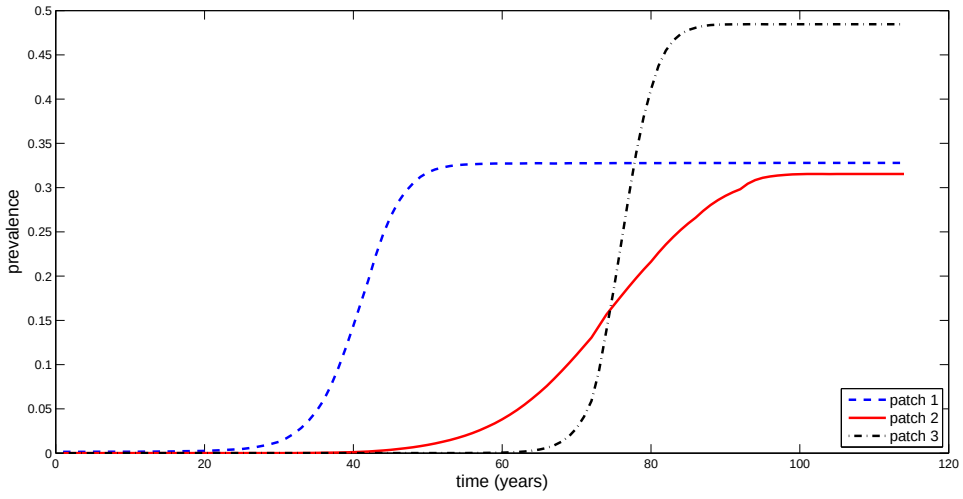


Figure 5.1. Comparison of prevalence vs. time by patch for 3-cell model in Chapter 2.

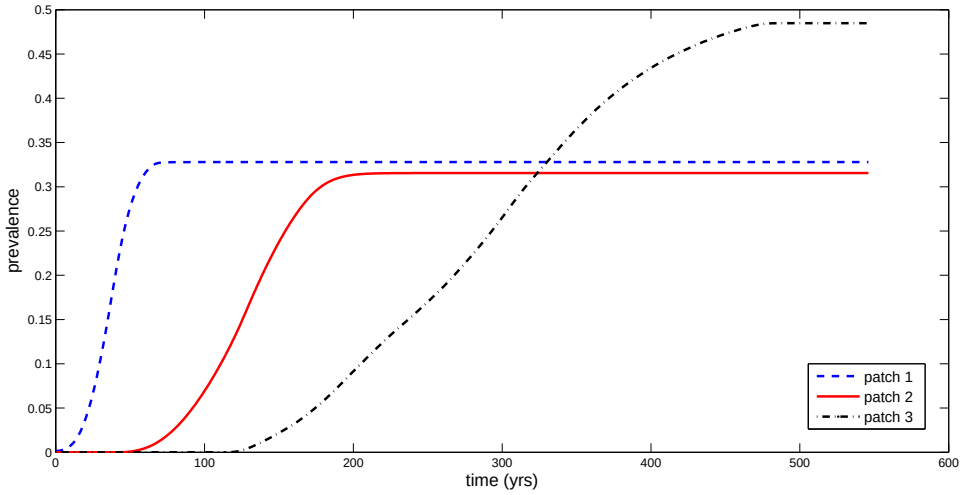


Figure 5.2. Comparison of prevalence vs. time by patch for cellular automaton model in Chapter 3.

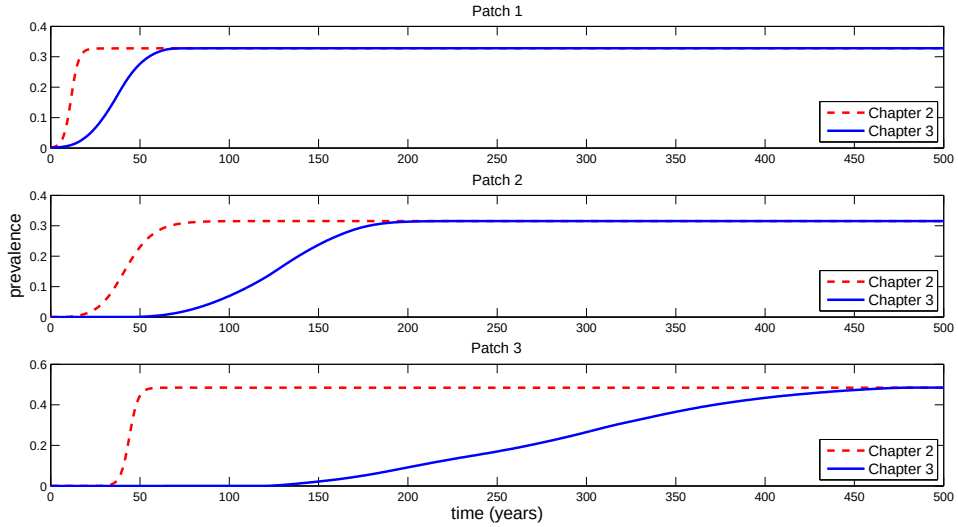


Figure 5.3. Comparison of prevalence vs. time results in Chapter 2 and 3. Results are presented by patch..

Each model (and method) presented here plays an important role in the goal of identifying and understanding effects of vector migration on *T. cruzi* sylvatic transmission. Sylvatic *T. cruzi* transmission is complex due to the different host-vector

cycles and the modes by which the infection is spread and possible infection with other strain types (in the same population). This model focuses solely on infection with one strain, yet both strain types are known to circulate in the United States (in other hosts not studied here). Further work may consider multiple strain types as well as other host-vector transmission cycles carrying other strain types of *T. cruzi*. We also remark that the cellular automaton model presented here has only been analyzed numerically. Further work may be to develop a system of integro-difference equations to model the dispersal, which can be analyzed analytically to determine an expression for invasion speed (to be compared with our results). We also note that there is a limited amount of biological (experimental) data regarding vector dispersal rates. Due to this limited information, in this study, the vector dispersal rate is estimated solely based upon vector life cycle information (maturation rate). Vectors may disperse for reasons other than maturation, such as when the current host dies or fails to return to the nest, yet we do not have any data to estimate this dispersal rate. There is a need for further research that develops a more accurate measure of the rate at which vectors disperse to find a new host.

# APPENDIX A

# ERROR CORRECTIONS FOR MIGRATION RATE CALCLATION

The method developed in this chapter generates some error near the regions where the patch boundary segments intersect. In this appendix, we will define the types of errors generated and calculate the adjusted migration rates.

## A.1 Sources of errors

Each boundary segment intersection point can be described in terms of concave and convex corners. The errors generated at each corner can be described in terms of dispersals that are either an "overcount" (including dispersals that should not be counted by the approximation) or an "undercount"(failing to include dispersals that should be counted by the approximation). Recall that each boundary segment, $j$ (numbered clockwise), has its own coordinate system with dispersals originating from each point $(x_j, y_j)$. The method described in this chapter integrates dispersals over a certain proportion of an ellipse at each point in the $j^{th}$ coordinate system. Error is generated at corners where boundary segments intersect in which there is a changeover in coordinate systems between the $j^{th}$ and $(j+1)^{st}$ boundary segments as seen in Figure A.1. If the corner is convex, there will be a region with points in both coordinate systems, while for a concave corner, there will be a region in which points are not in either system. For each corner, there will be different regions in which the approximation overcounts or undercounts the dispersals made from each point in the region.

For each convex corner, there are 3 sources of error: 2 regions generating an undercount with a region generating a double count between them as seen in Figure A.1. If we consider the $j^{th}$ and $(j+1)^{st}$ boundary segments, there is an undercount generated on the right by points in the $j^{th}$ coordinate system (see Figure A.1, error type 1). For this type of error, dispersals are only being counted as crossing the $j^{th}$ boundary segment, but in fact some dispersals originating in the $j^{th}$ coordinate system

cross the $(j+1)^{st}$ boundary segment, but are not being accounted for. A similar error undercount is generated on the left at points only in the $(j+1)^{st}$ coordinate system (Figure A.1, error type 3), in which some dispersals originating in the $(j+1)^{st}$ coordinate system cross the $j^{th}$ boundary segment, but are not counted. Because the coordinate systems intersect at a convex corner, there is also an overcount at points which are on both coordinate systems (the region between the intersection of the inward normals of the boundary segments) (Figure A.1, error type 2). Some dispersals originating in the region of intersection of the $j^{th}$ and $(j+1)^{st}$ coordinate systems cross both the $j^{th}$ and $(j+1)^{st}$ patch boundaries and thus are counted twice, once for each coordinate system.

At a concave corner, there are also 3 sources of error: 2 regions generating an overcount with a region that remains uncounted between them. If we again consider the corner to be the intersection of segments $j$ and $j+1$, there is an overcount on the right at points on the $j^{th}$ coordinate system (Figure A.1, error type $\bar{1}$). For this type of error, some dispersals originating in the $j^{th}$ coordinate system are being counted as crossing the $j^{th}$ boundary segment, but in fact actually cross the extension of segment $j$ to the right of the intersection point, but do not cross the patch boundary. This generates an overcount of dispersals that end up above the $j^{th}$ boundary segment, but below the $(j+1)^{st}$ segment. Similarly, there is an overcount on the left at points on the $(j+1)^{st}$ coordinate system (Figure A.1, error type $\bar{3}$), in which dispersals are counted as crossing the $(j+1)^{st}$ segment, but actually cross an extension of the $(j+1)^{st}$ segment. Because of the nature of the concave corner, there is a region not a part of either coordinate system (the region between the intersection of the inward normals of the boundary segments) (Figure A.1, error type $\bar{2}$), so no dispersals originating from within this region are counted as crossing either patch boundary segment. A visual depiction of these types of errors is given in Figure A.1.

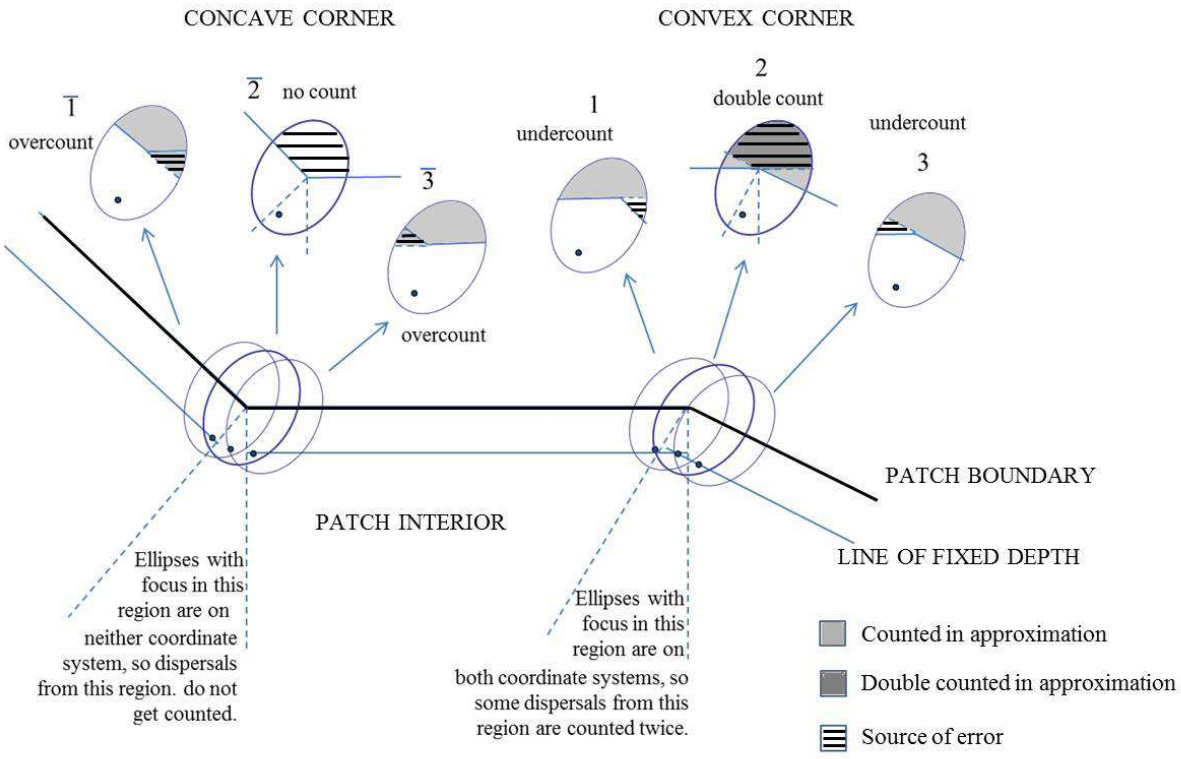


Figure A.1. Errors generated at a concave corner (left) and a convex corner (right).

In either case, we will define the total error of dispersals at the corner of the $j^{th}$ and $(j+1)^{st}$ boundary segments to be the sum of the 3 types of errors using the notation $\epsilon_j = \epsilon_{j1} + \epsilon_{j2} + \epsilon_{j3}$ if the corner is convex, and $\bar{\epsilon}_j = \bar{\epsilon}_{j1} + \bar{\epsilon}_{j2} + \bar{\epsilon}_{j3}$ if the corner is concave. In the remaining sections we will develop expressions for of error at each type of corner.

## A.2 Error analysis at convex corner

Consider a convex corner created by the $j^{th}$ and $(j+1)^{st}$ linear boundary segments. In this section, we will focus our attention only on error types 2 and 3 (since type 1 is symmetric to type 3).

### A.2.1 Global error boundaries for a convex corner

At each corner, there are certain regions in which the error of a given type is generated. As stated previously, for a convex corner, there are two regions of undercount with a double count region between them. For each convex corner, there are 3 regions in which originating dispersals may either not be counted when they should be or may be counted twice (when they should only be counted once). The regions (seen in Figure A.2) are comprised of an elliptical curve (an ellipse with one focus at the corner and orientation directly opposite of $\theta_0$) and two linear segments tangent to the elliptical curve defined as the lines $y_{(j+1)} = -r_{max_{j+1}}$ and $y_j = -r_{max_j}$. Ellipses with focus on either of the two linear segments will be tangent to the respective patch boundary segment. An ellipse with focus on the elliptical curve between the two linear segments will pass through the corner with a tangent to an alternate line whose slope is between those of the two patch boundaries. We define the region bounded by the patch boundaries and the elliptical curve as $U_0$, the region bounded by the left boundary segment, the elliptical curve, and the line segment parallel to the right boundary segment as $U_1$ and the region bounded by the right segment, the elliptical curve, and the line segment parallel to the left boundary segment as $U_2$. We also define 2 relevant points, $P_1 = (x_{P1}, y_{P1})$ and $P_2 = (x_{P2}, y_{P2})$, as the points of tangency of the elliptical curve and each of the two linear segments defined previously. The ellipse with focus at either of these coordinates represents an ellipse that is tangent to the corresponding patch boundary at the corner(i.e., an ellipse that has focus at $P_1$ will be $-r_{max_j}$ distance from patch boundary $j$). An ellipse with focus in $U_0$ includes the corner in its interior and crosses both boundary segments. An ellipse with focus in $U_1$ will cross the $j^{th}$ boundary segment and also the extension of the $(j+1)^{st}$ boundary segment outside of the patch. An ellipse with focus in $U_2$ will cross the $(j+1)^{st}$ boundary segment as well as the extension of the $j^{th}$ boundary segment

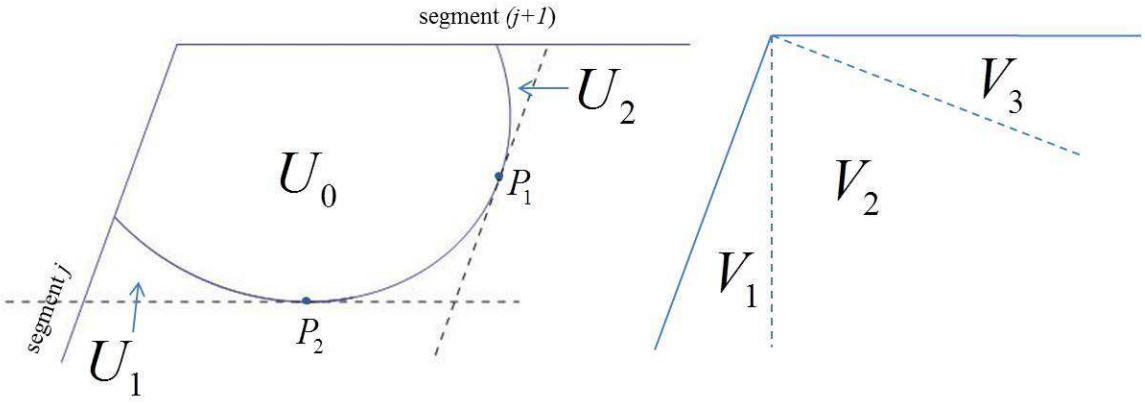


Figure A.2. Convex error regions.

outside of the patch. The type of error varies on where the point falls with respect to the 2 segments coordinate systems.

We also name 3 other regions defined by the inward normals of the boundary segments, $V_1$, $V_2$, and $V_3$ (seen in Figure A.2). $V_1$ is the region between the left boundary segment and the inward normal of the right segment, $V_3$ is the region between the right boundary segment and the inward normal of the left segment, and $V_2$ is the region between the inward normals of the 2 boundary segments. Undercounts are generated from some dispersals originating in region $V_1$ and $V_3$, and double counts are generated from some dispersals originating in region $V_2$. Dispersals from a given point are counted incorrectly by the method described in section 2.4.1, when the corresponding ellipse crosses the opposite segment on the opposite side of the corner. Error is also generated when the point of origin falls into either two consecutive segments' regions of integration (for a convex corner) or no regions of integration (for a concave corner). The global error regions are summarized in Table A.1.

Table A.1. Global error regions for convex corner

| Type of error | Error regions |
|---|---|
| 1 (Left undercount) | $W_1 = V_1 \bigcap [U_0 \bigcup U_2]$ |
| 2 (Double count) | $W_2 = V_2 \bigcap [U_0 \bigcup U_1 \bigcup U_2]$ |
| 3 (Right undercount) | $W_3 = V_3 \bigcap [U_0 \bigcup U_1]$ |

### A.2.2 Error definition for error type 3

The regions given in Table A.1 can be used to determine the global $x, y$ integral bounds for each type of error generated. An ellipse with focus in $U_0$ crosses both boundary segments, but are only counted as crossing the $(j+1)^{st}$ segment, failing to count the dispersals that cross the $j^{th}$ segment. Dispersals originating in $U_1$ pass to the left of the corner and are not counted as crossing the $j^{th}$ boundary segment, when in fact they should be. To define error type 3, we use the $(j+1)^{st}$ coordinate system. For a given $y_{(j+1)}$, we define the $x_{(j+1)}$ which generates minimal undercount error on the right to be $x_{(j+1),c}$, the focus of the ellipse passing through the point where the boundary segments intersect, and thus the right boundary of $U_0$, while $x_{(j+1),b}$ lies on the inward normal of the $j^{th}$ boundary segment (i.e., the boundary between $V_2$ and $V_3$), and is the focus of the ellipse generating maximum undercount error (see Figure A.3). Then,

$$x_{(j+1),b} = y_{(j+1)} \tan \zeta_j, \tag{A.1}$$

for $\zeta_j$ defined to be the angle turn from the $(j+1)^{st}$ to the $j^{th}$ boundary segments in the clockwise direction, i.e., the angle between the inward normals of the boundary segments.

We also determine

$$x_{(j+1),c} = y_{(j+1)} \tan(\theta_{2,(j+1)} - \pi/2) = -y_{(j+1)} \cot(\theta_{2,(j+1)}), \tag{A.2}$$

where $\theta_{2,(j+1)}$ is the maximum angle at which the ellipse with focus at $(x_{(j+1)}, y_{(j+1)})$ intersects boundary segment $(j+1)$ for any $x_{(j+1)}(\geq x_{(j+1),b})$, as shown in Figure A.3. We note here that $\theta_{2,(j+1)}$ is a function of $y_{(j+1)}$, but from this point forward will not explicitly express this dependence to avoid confusing notation. The functions $x_{(j+1),b}$ and $x_{(j+1),c}$ are the left and right boundaries of $W_3$.

Based on the region $W_3$, we define $y_{low,3}$ as $-r_{max_{j+1}}$ if $P_2 \in V_3$ or $y_{(j+1)}$ : $x_{(j+1),b} = x_{(j+1),c}$ if $P_2 \in V_1 \bigcup V_2$.

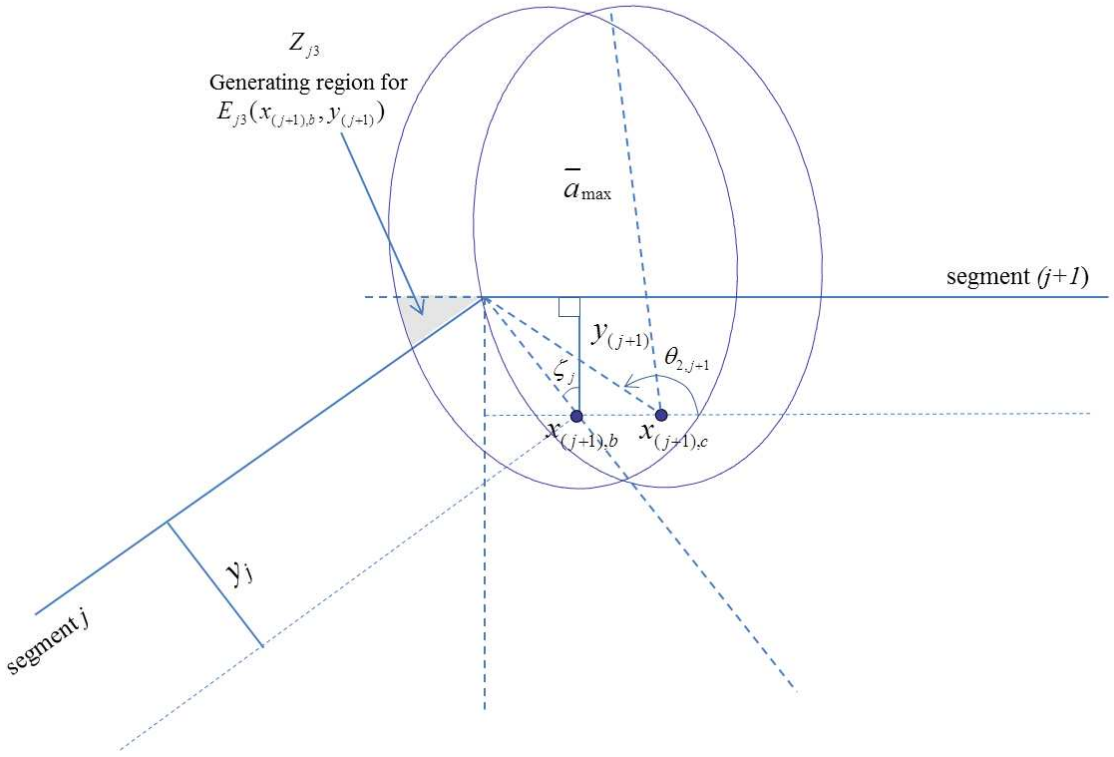


Figure A.3. Error of type 3 (undercount) generated at a convex corner.

For $x_{(j+1)}, y_{(j+1)} \in W_3$, we define $E_{j3}(x_{(j+1)}, y_{(j+1)})$ as the error generated by undercount of vectors originating at $(x_{(j+1)}, y_{(j+1)})$ in the $(j+1)^{st}$ coordinate system and crossing the $j^{th}$ boundary segment. The relevant region, $Z_{j3}(x_{(j+1)}, y_{(j+1)})$ (see Figure A.3), is the region into which vectors dispersing from $(x_{(j+1)}, y_{(j+1)})$ are not counted as crossing the $j^{th}$ patch boundary, when in fact they should be.

To describe $Z_{j3}$ in terms of the local (polar) coordinate system, we introduce several new quantities and equations. We define $\gamma_1$ and $\gamma_2$ as the angles at which the ellipse intersects opposite boundary segment, while $\theta_1$ and $\theta_2$ are previously defined as the angles at which the ellipse intersects boundary segment for the native coordinate system. Derivations for the expressions for $\theta_1$ and $\theta_2$ are given in section 2.4.1. The expressions for $\gamma_1$ and $\gamma_2$ can be derived similarly. In this calculation, these angles are functions of $y_{j+1}$, but we will not explicitly write this dependence. For a given $(x_{j+1}, y_{j+1})$, we also consider the angles defined by the intersection of the boundary segments using right triangle trigonometry, $\beta_{j+1} = \pi - \arctan\left(-y_{(j+1)}/x_{(j+1)}\right)$. Finally, we define the two relevant boundary segments, $r = F_j(\theta)$ and $r = F_{j+1}(\theta)$ in polar coordinates as lines passing through the pole. Depending on the eccentricity and turn angle between segments, the ellipse may generate error without actually containing the corner (cf., Figures A.4 and A.5). Thus, we define $Z_{j3}$ as $Z_{j3} = \{(r,\theta) | \{F_j < r < \min\{F_{j+1}, r_n\}\}, \{\max\{\beta_{(j+1)}, \gamma_1\} < \theta < \gamma_2\}\}$.

Then

$$E_{j3}(x_{(j+1)}, y_{(j+1)}) = \iint_{Z_{j3}} \Psi(r,\theta)\, r\, dr\, d\theta,$$

and the total undercount of this sort at this corner is

$$\epsilon_{j3} = \iint_{W_3} E_{j3}(x_{(j+1)}, y_{(j+1)}) dx_{(j+1)} dy_{(j+1)}.$$

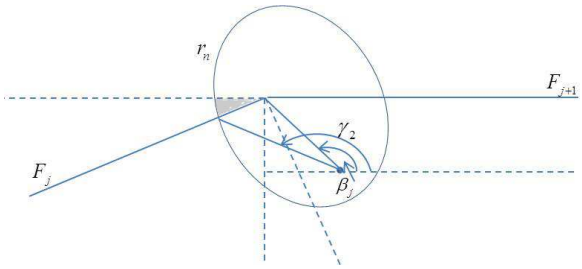


Figure A.4. Error type 3 (ellipse contains corner), point of origin in $U_0$.

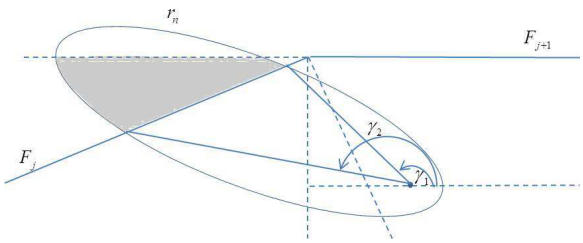


Figure A.5. Error type 3 (ellipse does not contain corner),point of origin in $U_2$.

The error definition for type 1 is similar to type 3, but uses the $j^{th}$ coordinate system.

### A.2.3 Error definition for type 2

At this point we also consider the region in which some dispersals originating in both coordinate systems are counted twice, once for each coordinate system. Because the origin of these dispersals is in both coordinate systems, we arbitrarily choose to use the $(j+1)^{st}$ coordinate system to define the error. An ellipse with focus in $W_2$ lies in the intersection region of the coordinate systems, $V_2$, and crosses both boundary segments, thus will lie in either $U_0, U_1$, or $U_2$. Dispersals in $U_0$ contain the corner, and cross both segments (as well as the extensions of each segment), thus dispersals are counted as crossing the extensions of each segment, when they should not be counted. An ellipse in $U_1$ will cross the $(j+1)^{st}$ segment and the extension of the $j^{th}$ segment to the right of the corner. The dispersals that cross the extension of the $j^{th}$ boundary are counted twice. In a similar manner, dispersals originating in $U_2$ that cross the extension of the $(j+1)^{st}$ segment are counted twice.

To determine the global bounds, we first define the line $y_j = -r_{max_j}$ in the $(j+1)^{st}$ coordinate system as $x_{(j+1),d} = x_{P1} + (y_{(j+1)} - y_{P1}) \cot \zeta_1$. Then $(x_{(j+1)}, y_{(j+1)}) \in W_2$ corresponds to the global integral bounds $x_{(j+1)} \in [0, x_{right,2}]$, where

$$x_{right,2} = \begin{cases} \min x_{(j+1),b}, x_{(j+1),d}, & y_{P1} < y_{(j+1)} < 0 \\ x_{(j+1),c}, & -r_{max_{j+1}} \leq y_{(j+1)} < y_{P1} \end{cases}, \tag{A.3}$$

and $y_{(j+1)} \in [y_{low,2}, 0]$, where

$$y_{low,2} = \begin{cases} -r_{max_{j+1}}, & x_{P2} > 0 \\ y_{(j+1)} : x_{(j+1),c} = 0, & x_{P2} < 0, x_{P1} > 0 \\ y_{(j+1)} : x_{(j+1),d} = 0, & x_{P1} < 0 \end{cases} \tag{A.4}$$

For $(x_{(j+1)}, y_{(j+1)}) \in W_2$, we define $E_{j2}(x_{(j+1)}, y_{(j+1)})$ as the error generated by overcount (double count) of vectors originating at $x_{(j+1)}, y_{(j+1)}$ in the $(j+1)^{st}$ coordinate system and crossing both boundary segments. The relevant region, $Z_{j2} = \{(r, \theta) | \{\max\{F_j, F_j + 1\} < r < r_n\}, \{\max\{\gamma_1, \theta_1\} < \theta < \min\{\gamma_2, \theta_2\}\}$ (seen in Figures A.7, A.8, and A.9) is the region into which vectors dispersing from $(x_{(j+1)}, y_{(j+1)})$ are counted twice.

Then,

$$E_{j2}(x_{(j+1)}, y_{(j+1)}) = \iint_{Z_{j2}} \Psi(r, \theta)\, r\, dr\, d\theta.$$

We define the double count error to be

$$\epsilon_{j2} = \iint_{W_2} E_{j2}(x_{(j+1)}, y_{(j+1)}) dx_{(j+1)} dy_{(j+1)}.$$

## A.3 Error generated at concave corner

### A.3.1 Global bounds for concave corner

For a concave corner, there are 6 regions in which originating dispersals may generate error. The regions are determined by the same elliptical curve and 2 linear

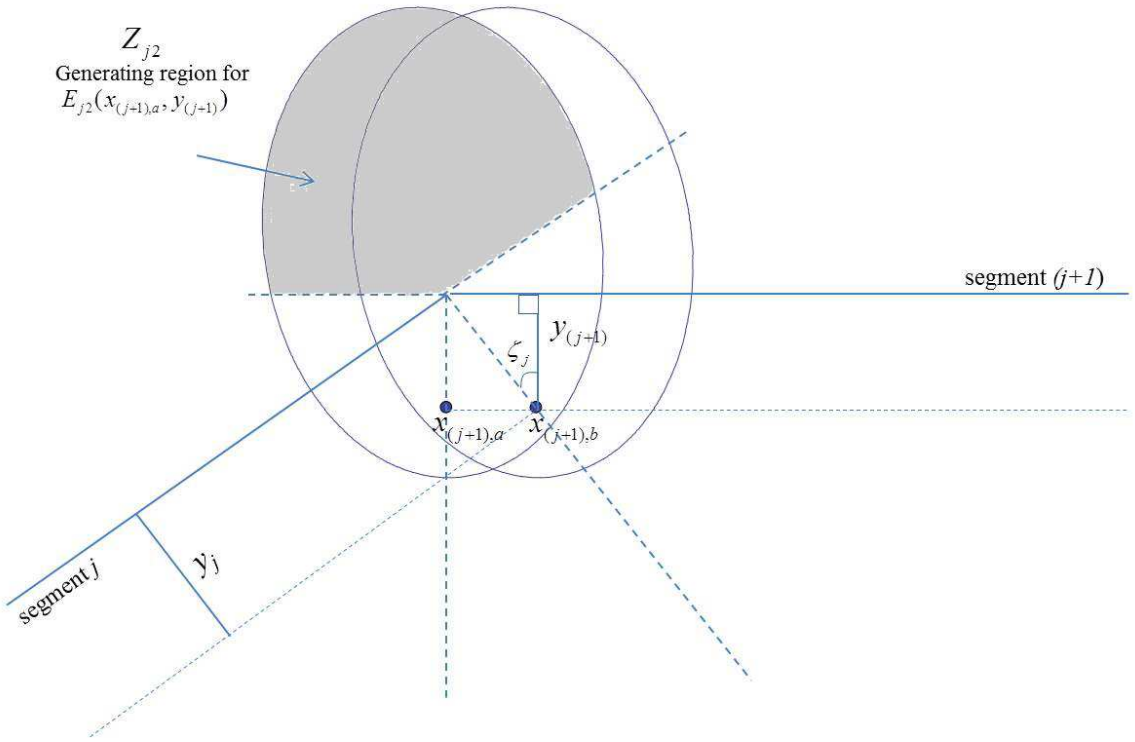


Figure A.6. Error of type 2 (double count) generated at a convex corner.

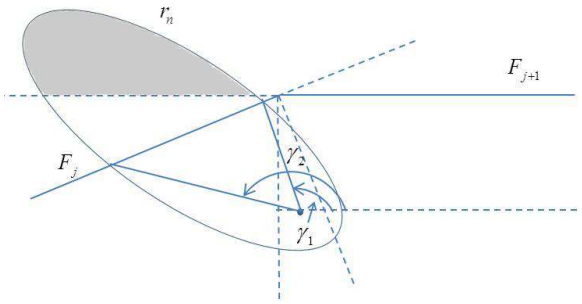


Figure A.7. Error type 2 (ellipse does contain corner), point of origin in $U_1$ .

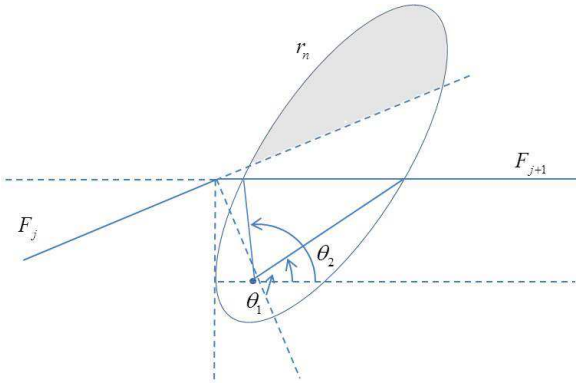


Figure A.8. Error type 2 (ellipse does not contain corner), point of origin in $U_2$.

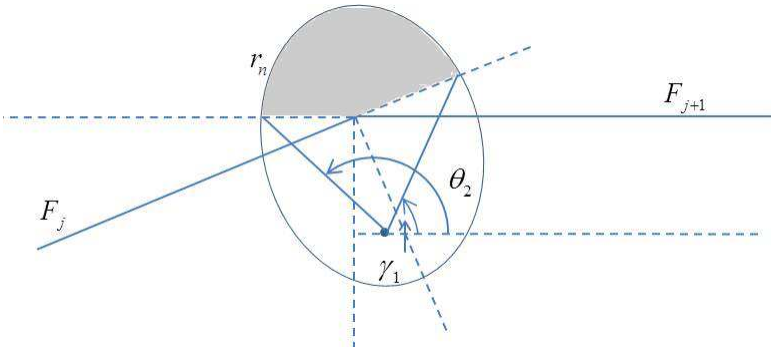


Figure A.9. Error type 2 (ellipse contains corner), point of origin in $U_0$.

segments as described in section A.2.1, but we note the presence of 3 relevant regions lying below the regions described in section A.2.1, in which originating dispersals may still generate error. We also name 3 additional regions defined by the inward normals of the boundary segments. Because of the nature of a concave corner, $\bar{V}_1 = \{(x,y)|\{0 < x < l_j\}, \{y_j < 0\}\}$, $\bar{V}_3 = \{(x,y)|\{0 < x_{(j+1)} < l_{(j+1)}\}, \{y_{(j+1)} < 0\}\}$, and $\bar{V}_2$ is the region lying between the inward normals of the 2 coordinate systems (i.e. the region not in either coordinate system). A visual depiction of these named regions is given in Figure A.10 and a summary is given in Table A.2.

An ellipse with focus in $U_0$ will contain the corner, and thus crosses both boundary segments. An ellipse with focus in $U_1$ will cross the $j^{th}$ boundary segment an the extension of the $(j+1)^{st}$ segment inside of the patch. Ellipses generated in $U_3$ do not contain the corner, and cross extensions of the boundary segments inside the the patch, but do not cross patch boundaries. An ellipse with focus in $U_4$ will cross the extension of boundary segment $(j+1)^{st}$ to the left of the corner, and will not cross segment $j$. Then, an ellipse with focus in $U_5$ will cross the extension of boundary segment $j$ to the right of the corner, and will not cross segment $j+1$. Some dispersals in $\bar{V}_1$ will generate overcounts on the left, and some dispersals in $\bar{V}_3$ will generate

Table A.2. Global error regions for concave corner

| Type of error | Error regions |
|---|---|
| $\bar{1}$ (Left overcount) | $\bar{W}_1 = \bar{V}_1 \bigcap [U_0 \bigcup U_2 \bigcup U_3 \bigcup U_5]$ |
| $\bar{2}$ (No count) | $\bar{W}_2 = \bar{V}_2 \bigcap [U_0 \bigcup U_1 \bigcup U_2]$ |
| $\bar{3}$ (Right overcount) | $\bar{W}_3 = \bar{V}_3 \bigcap [U_0 \bigcup U_1 \bigcup U_3 \bigcup U_4]$ |

overcounts on the right. Finally, no dispersals in $\bar{V}_2$ are counted when some of them (coinciding with the appropriate $U_i$ regions) should be.

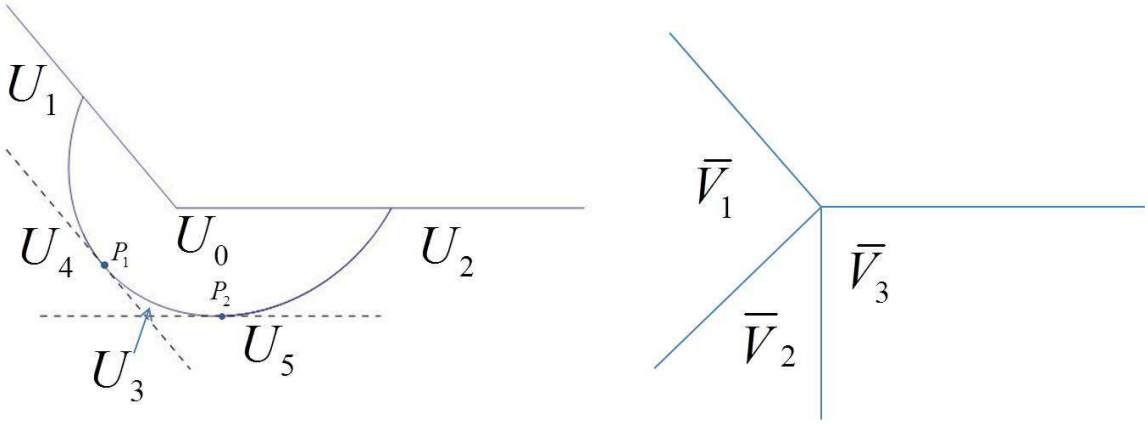


Figure A.10. Concave error regions.

### A.3.2 Error definition for type $\bar{3}$

To define error type $\bar{3}$, the overcount of dispersals on the right, we use coordinate system $(j+1)$. An overcount is generated on the right by an ellipse in $W_1$. Some dispersals in $\bar{V}_3$ will generate an overcount of vectors on the right. Because dispersals originating in $U_0$ contain the corner, crossing both boundary segments, dispersals are counted as crossing the extension of the $j+1^{st}$ segment, when they should not be counted. The method integrates over the part of the ellipse above the $j+1^{st}$

segment, but should actually be counted as crossing the $j^{th}$ segment. Ellipses with focus in $U_1$ (generating error) will pass to the left of the corner and cross the $j^{th}$ boundary segment. The method integrates over the part of the ellipse above the extension of segment $(j+1)$, rather than the part of the ellipse over segment $j$. Dispersals originating in $U_3$ that generate error are counted as crossing the extension of the $j+1^{st}$ segment, when they should not be. Ellipses with focus in $U_4$ cross the extension of boundary segment $j+1$ to the left of the corner, and with regard to type $\bar{3}$, these dispersals are counted when in fact they should not be. For a given $y_{(j+1)}$, we define the $x_{(j+1)}$ which generates minimal overcount on the right as $\bar{x}_{(j+1),b}$ given by the right half of the ellipse separating $U_0$ from $U_2$, which can be written using right triangle trigonometry (c.f. Figure A.11) as

$$\bar{x}_{(j+1),b} = y_{(j+1)} \tan(\theta_{2,(j+1)} - \pi/2) = -y_{(j+1)} \cot(\theta_{2,(j+1)}), \tag{A.5}$$

where $\theta_{2,(j+1)}$ is the maximum angle at which the ellipse with focus at $(x_{(j+1)}, y_{(j+1)})$ intersects boundary segment $(j+1)$ for any $x_{(j+1)} (\geq x_{(j+1),b})$, as shown in Figure A.11. We also observe the focus of the ellipse generating maximum overcount (for fixed $y_{(j+1)}$) error to be on the inward normal of the $(j+1)^{st}$ coordinate system, at $\bar{x}_{(j+1),a} = 0$, separating $\bar{V}_2$ from $\bar{V}_3$. We note $\theta_{2,(j+1)}$ and $\bar{x}_{(j+1),b}$ dependance on $y_{(j+1)}$ will be left implicit from this point forward.

Given $(x_{(j+1)}, y_{(j+1)}) \in \bar{W}_3$, error is generated for $0 \leq x_{(j+1)} \leq \bar{x}_{(j+1),b}$ and $y_{(j+1)} \in [\bar{y}_{low,3}, 0]$, where $\bar{y}_{low,3}$ is $-r_{max_{j+1}}$ if $P_2 \in \bar{V}_3$ or $y_{(j+1)} : \bar{x}_{(j+1),b} = 0$ if $P_2 \in \bar{V}_1$ or $\bar{V}_2$.

$$\bar{y}_{low,3} = \begin{cases} -r_{max}, & P2 \in \bar{V}_2 \\ x_{(j+1),c} = 0, & P2 \in \bar{V}_3 \end{cases} \tag{A.6}$$

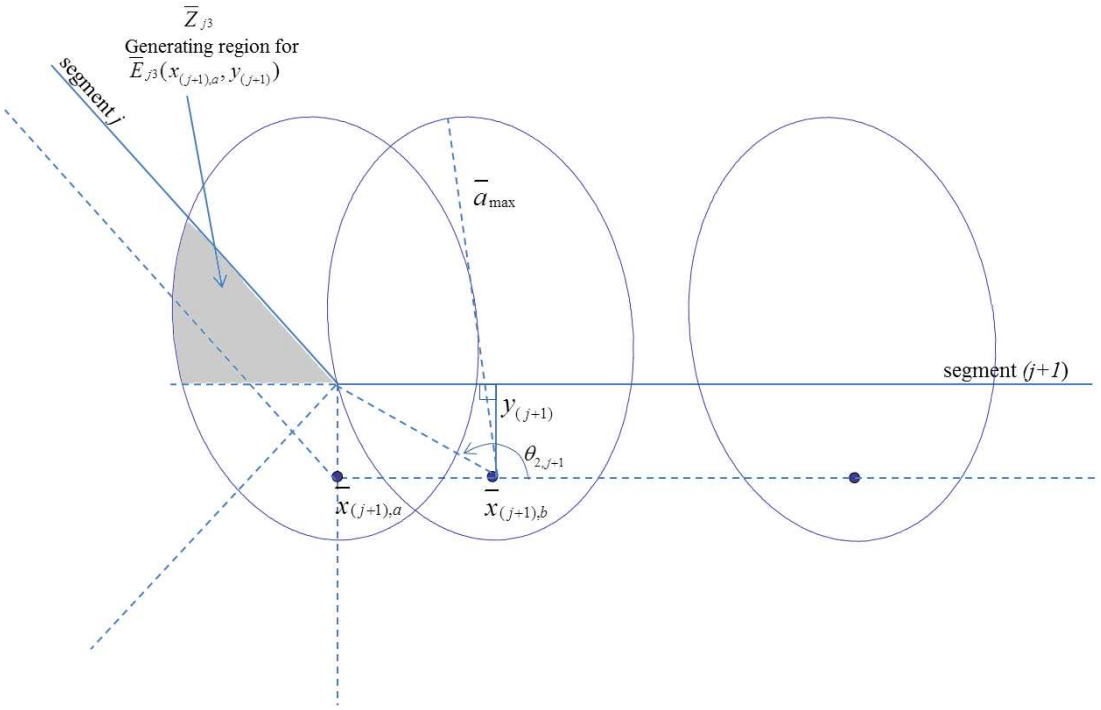


Figure A.11. Error of type 3 (overcount) generated at a concave corner.

For, $(x_{(j+1)}, y_{(j+1)}) \in \bar{W}_3$, we define $\bar{E}_{j3}(x_{(j+1)}, y_{(j+1)})$ as the overcount error of vectors dispersing from $(x_{(j+1)}, y_{(j+1)})$ into a corresponding region $\bar{Z}_{j3}$ (see Figure A.11), where

$\bar{Z}_{j3} = \{(r,\theta) | F_{j+1} < r < \min\{F_j, r_n\}, \max\{\beta_{j+1}, \theta_1\} < \theta < \theta_2\}$, so

$$\bar{E}_{j3}(x_{(j+1)}, y_{(j+1)}) = \iint_{\bar{Z}_{j3}} \Psi(r,\theta)\, r\, dr\, d\theta.$$

Then the total error generated by overcount of vectors originating in the $(j+1)^{st}$ coordinate system and failing to cross the $j^{th}$ boundary segment is

$$\bar{\epsilon}_{j3} = \iint_{\bar{W}_3} \bar{E}_{j3}(x_{(j+1)}, y_{(j+1)}) dx_{(j+1)} dy_{(j+1)}.$$

The error definition for type $\bar{1}$ is similar to type $\bar{3}$, but uses the $j^{th}$ coordinate system.

### A.3.3 Error definition for type $\bar{2}$

We now focus on the error generated by vectors dispersing from within the region not contained in either coordinate system. In order to describe (and calculate) this error, we may appropriately extend either coordinate system into the designated region between coordinate systems. We will use the $(j+1)^{st}$ system. Error type $\bar{2}$ occurs because some dispersals in region $\bar{V}_2$ (not part of either coordinate system). Dispersals in $U_0$ cross both patch boundaries, and are not counted by the method when they should be. Dispersals originating in $U_1$ cross patch boundary $j$, and are not counted by the method, when they should be. In a similar manner, dispersals in $U_2$ are not counted, but cross segment $j+1$, and should be counted.

Given $(x_{(j+1)}, y_{(j+1)}) \in \bar{W}_2$, then $y_{(j+1)} \in \left[\bar{y}_{low,2}, y_{(j+1),b}\right]$, where

$$\bar{y}_{left,2} = \begin{cases} y_{(j+1),a}, & x_{(j+1)} < x_{P1} \\ g(x_{(j+1)}), & x_{P1} < x_{(j+1)} < x_{P2} \\ -r_{max_{j+1}}, & x_{(j+1)} > x_{P2} \end{cases}, \tag{A.7}$$

where $g(x)$ represents the Cartesian equation for the lower branch of the elliptical curve with focus at corner (i.e., the origin) with opposite orientation as $\theta_0$ and $x_{(j+1)} \in [\bar{x}_{left,2}, 0]$, where $\bar{x}_{left,2}$ is the $x$-value for which $\bar{y}_{low,2} = \bar{y}_{(j+1),a}$.

So, for a given $(x_{(j+1)}, y_{(j+1)}) \in \bar{W}_2$, we define $\bar{E}_{j2}(x_{(j+1)}, y_{(j+1)})$ as the error generated by failing to count dispersals originating in $\bar{W}_2$ dispersing into $\bar{Z}_{j2} = \{(r,\theta) | \max\{F_j, F_{j+1}\} < r < r_n, \max\{\theta_1, \gamma_1\} < \theta < \min\{\theta_2, \gamma_2\}\}$.

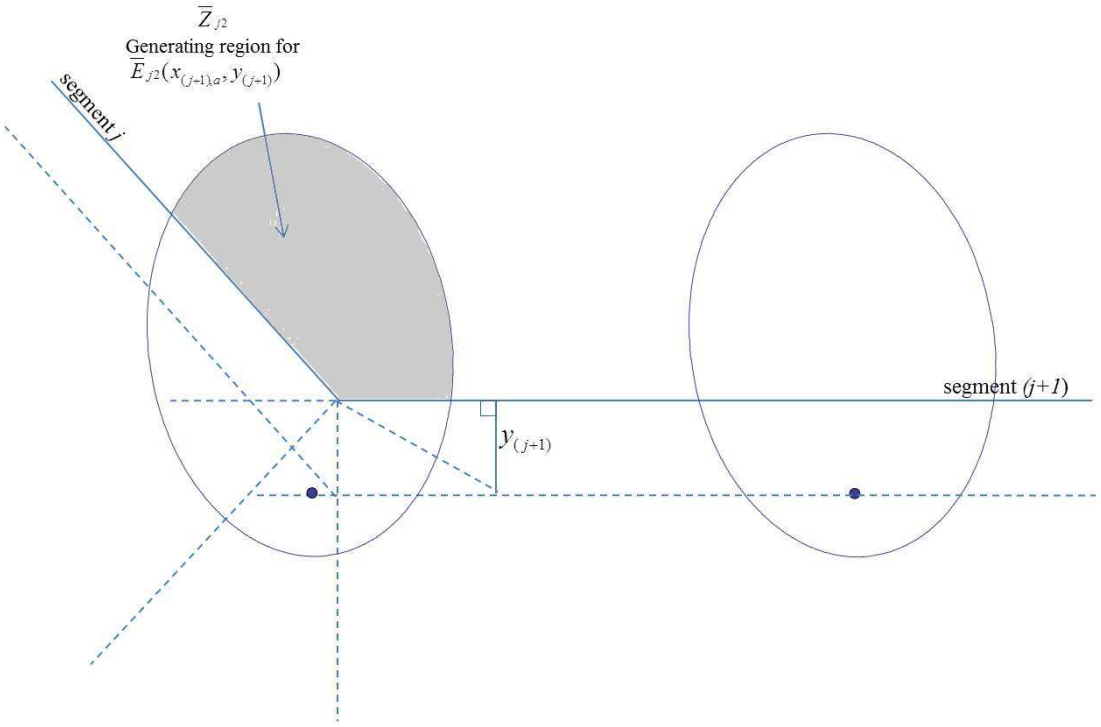


Figure A.12. Error of type $\bar{2}$ (no count) generated at a concave corner.

Then,

$$\bar{E}_{j2}(x_{(j+1)}, y_{(j+1)}) = \iint_{\bar{Z}_{j2}} \Psi(r, \theta)\, r\, dr\, d\theta.$$

# APPENDIX B

# R0 AND ENDEMIC PREVALENCE CALCULATIONS

B.1 Patch 2 with $p = 0$ and no migration

We determine the reproductive number, $R_0$, for section 3.3.2, patch 2 with no vertical transmission or migration, using the next generation matrix method [85].

Based on the next generation matrix method, we rewrite system (3.4), $\frac{dX}{dt} = f(X)$ in terms of two vectors: $\frac{dX}{dt} = \mathcal{F}_0 - \mathcal{V}_0$. $\mathcal{F}_0$ represents the terms generating new infections, while $\mathcal{V}_0$ consists of the remaining terms. After computing the derivatives of $\mathcal{F}_0$ and $\mathcal{V}_0$, we obtain

$$F_0 = \begin{pmatrix} 0 & q_S\beta_R\frac{N^*_{S2}}{N^*_{R2}} & 0 & (1-q_S)\beta_{WS}\frac{N^*_{S2}}{N^*_{W2}} \\ \beta_{S2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_{W2}\frac{N^*_{G2}}{N^*_{W2}} \\ q_W\beta_{SW} & 0 & (1-q_W)\beta_G & 0 \end{pmatrix}$$

and

$$V_0 = \begin{pmatrix} \mu_S & 0 & 0 & 0 \\ 0 & \mu_R & 0 & 0 \\ 0 & 0 & \mu_G & 0 \\ 0 & 0 & 0 & \mu_W \end{pmatrix}.$$

The dominant eigenvalue of $F_0V_0^{-1}$ is

$$R_0 = \sqrt{\frac{1}{2}\left(P + \sqrt{P^2 - 4Q}\right)},$$

where

$$P = f_1 + f_2 + f_3, \;\; Q = f_1f_3,$$
$$f_1 = \frac{(1-q_W)\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}, \;\; f_2 = \frac{q_W\beta_{SW}(1-q_S)\beta_{WS}}{\mu_S\mu_W}\frac{N^*_{S2}}{N^*_{W2}}, \;\; f_3 = \frac{q_S\beta_R\beta_{S2}}{\mu_R\mu_S}\frac{N^*_{S2}}{N^*_{R2}}. \tag{B.1}$$

It remains to be shown that $\sqrt{P^2-4Q}$ is real. Thus, we must show that $P^2-4Q>0$, as follows:

$$\begin{aligned}
P^2-4Q &= (f_1+f_2+f_3)^2-4f_1f_3 \\
&= f_1^2+f_2^2+f_3^2+2f_1f_2+2f_2f_3+2f_1f_3-4f_1f_3 \\
&= f_1^2-2f_1f_3+f_3^2+2(f_1f_2+f_2f_3)+f_2^2 \\
&= (f_1-f_3)^2+2(f_1f_2+f_2f_3)+f_2^2>0.
\end{aligned}$$

To determine the existence of endemic equilibria, we determine the equilibrium conditions to be

$$\begin{aligned}
(q_S\beta_R x^*_{R2}+(1-q_S)\beta_{WS}x^*_{W2})(1-x^*_{S2})N^*_{S2}-\mu_S x^*_{S2}N^*_{S2}=0,\\
\beta_{S2}x^*_{S2}N^*_{S2}(1-x^*_{R2})-\mu_R x^*_{R2}N^*_{R2}=0,\\
\beta_{W2}x^*_{W2}(1-x^*_{G2})N^*_{G2}-\mu_G x^*_{G2}N^*_{G2}=0,\\
((1-q_W)\beta_G x^*_{G2}N^*_{G2}+q_W\beta_{SW}x^*_{S2}N^*_{S2})(1-x^*_{W2})-\mu_W x^*_{W2}N^*_{W2}=0,
\end{aligned} \tag{B.2}$$

where $x^*_{S2}=\frac{I^*_{S2}}{N^*_{S2}}$, $x^*_{R2}=\frac{I^*_{R2}}{N^*_{R2}}$, $x^*_{G2}=\frac{I^*_{G2}}{N^*_{G2}}$, and $x^*_{W2}=\frac{I^*_{W2}}{N^*_{W2}}$.

To simplify, we divide the 1st and 3rd equations by the nonzero values $N^*_{S2}$ and $N^*_{G2}$ respectively and make the substitutions, $\tilde{\beta}_G=(1-q_W)\beta_G\frac{N^*_{G2}}{N^*_{W2}}$, $\tilde{\beta}_{SW}=q_W\beta_{SW}\frac{N^*_{S2}}{N^*_{W2}}$, and $\tilde{\beta}_{S2}=\beta_{S2}\frac{N^*_{S2}}{N^*_{R2}}$.

We solve the 2nd and 3rd equations for $x^*_{R2}$ and $x^*_{G2}$:

$$x^*_{R2}=\frac{\tilde{\beta}_{S2}x^*_{S2}}{\tilde{\beta}_{S2}x^*_{S2}+\mu_R} \quad \text{and} \quad x^*_{G2}=\frac{\beta_{W2}x^*_{W2}}{\beta_{W2}x^*_{W2}+\mu_G}$$

We then solve the 1st equation in order to isolate $x^*_{W2}$. We obtain

$$(\beta_R x^*_{R2}+(1-q_S)\beta_{WS}x^*_{W2})(1-x^*_{S2})-\mu_S x^*_{S2}=0.$$

Substituting the result for $x^*_{R2}$ from above, we obtain

$$\left[q_S\beta_R\left(\frac{\tilde{\beta}_{S2}x^*_{S2}}{\tilde{\beta}_{S2}x^*_{S2}+\mu_R}\right)+(1-q_S)\beta_{WS}x^*_{W2}\right](1-x^*_{S2})-\mu_S x^*_{S2}=0.$$

Multiplying through on both sides by the denominator and combining like terms results in

$$\begin{aligned}-(q_S\beta_R\tilde{\beta}_{S2}+\tilde{\beta}_{S2}\mu_S)(x^*_{S2})^2-(1-q_S)\beta_{WS}\tilde{\beta}_{S2}x^*_{W2}(x^*_{S2})^2+((1-q_S)\beta_{WS}\tilde{\beta}_{S2}\\ -(1-q_S)\beta_{WS}\mu_R)x^*_{W2}x^*_{S2}\\ +((1-q_S)\beta_R\tilde{\beta}_{S2}-\mu_S\mu_R)x^*_{S2}+(1-q_S)\beta_{WS}\mu_Rx^*_{W2}=0\end{aligned} \tag{B.3}$$

We then divide everything by $\mu_R$ and $\mu_S$ to obtain,

$$\begin{aligned}-\left(\frac{q_S\beta_R\tilde{\beta}_{S2}}{\mu_R\mu_S}+\frac{\tilde{\beta}_{S2}}{\mu_R}\right)(x^*_{S2})^2+\left(\frac{q_S\beta_R\tilde{\beta}_{S2}}{\mu_R\mu_S}-1\right)x^*_{S2}-\left(\frac{(1-q_S)\beta_{WS}\tilde{\beta}_{S2}}{\mu_R\mu_S}(x^2_{S2})\right)x^*_{W2}\\ -x^*_{W2}\left(\frac{(1-q_S)\beta_{WS}\tilde{\beta}_{S2}}{\mu_R\mu_S}-\frac{(1-q_S)\beta_{WS}}{\mu_S}\right)x^*_{S2}-\left(\frac{(1-q_S)\beta_{WS}}{\mu_S}\right)x^*_{W2}=0.\end{aligned} \tag{B.4}$$

To simplify, we make the substitutions $A=\frac{q_S\beta_R}{\mu_S}$, $B=\frac{\tilde{\beta}_{S2}}{\mu_R}$, and $C=\frac{(1-q_S)\beta_{WS}}{\mu_S}$ and solve for $x^*_{W2}$ to obtain

$$x^*_{W2}=\frac{x^*_{S2}\left[(A+1)x^*_{S2}-A+\frac{1}{B}\right]}{C(1-x^*_{S2})\left(x^*_{S2}+\frac{1}{B}\right)}. \tag{B.5}$$

We finally solve the 4th equation; substituting $x^*_{G2}$, we obtain

$$\left[\tilde{\beta}_G\left(\frac{\beta_{W2}x^*_{W2}}{\beta_{W2}x^*_{W2}+\mu_G}\right)+\tilde{\beta}_{SW}x^*_{S2}\right](1-x^*_{W2})-\mu_Wx^*_{W2}=0.$$

Multiplying through by the denominator on both sides, we obtain

$$\begin{aligned}-(\tilde{\beta}_G\beta_{W2}+\beta_W\mu_W)(x^*_{W2})^2-\tilde{\beta}_{SW}\beta_{W2}x^*_{S2}(x^*_{W2})^2+(\tilde{\beta}_{SW}\beta_{W2}-\tilde{\beta}_{SW}\mu_G)x^*_{S2}x^*_{W2}\\ +(\tilde{\beta}_G\beta_{W2}-\mu_G\mu_W)x^*_{W2}-\tilde{\beta}_{SW}\mu_Gx^*_{S2}=0,\end{aligned}$$

and dividing all terms by $\mu_W$ and $\mu_G$ to get

$$
\begin{aligned}
&-\left(\frac{\tilde{\beta}_G\beta_{W2}}{\mu_G\mu_W}+\frac{\beta_{W2}}{\mu_G}\right)(x^*_{W2})^2+\left(\frac{\tilde{\beta}_G\beta_{W2}}{\mu_G\mu_W}-1\right)x^*_{W2}\\
&-\left(\frac{\tilde{\beta}_{SW}\beta_{W2}}{\mu_G\mu_W}(x^*_{W2})^2-\left(\frac{\tilde{\beta}_{SW}\beta_{W2}}{\mu_G\mu_W}-\frac{\tilde{\beta}_{SW}}{\mu_W}\right)x^*_{W2}-\frac{\tilde{\beta}_{SW}}{\mu_W}\right)x^*_{S2}=0
\end{aligned}
\tag{B.6}
$$

We make the substitutions $D=\frac{\tilde{\beta}_{SW}}{\mu_W}$, $E=\frac{\beta_W}{\mu_G}$, and $F=\frac{\tilde{\beta}_G}{\mu_W}$ and rewrite the previous equation as

$$
(1-x^*_{W2})\left(x^*_{W2}+\frac{1}{E}\right)x^*_{S2}=\frac{1}{D}x^*_{W2}\left[(F+1)x^*_{W2}-F+\frac{1}{E}\right].
$$

We can now determine a polynomial in $x^*_{S2}$, say $\phi(x^*_{S2})$. The resulting polynomial is of 4th degree, once the disease free equilibrium has been divided out, with constant term $\phi(0)=C(P-Q-1)$, with $P,Q$ as defined in (3.5), which is positive for $R_0>1$ by the following result

$$
\begin{aligned}
R_0^2&>1\\
\frac{1}{2}\left[P+\sqrt{P^2-4Q}\right]&>1\\
\sqrt{P^2-4Q}&>2-P\\
P^2-4Q&>4-4P+P^2\\
P-Q&>1
\end{aligned}
$$

Therefore, since $\sqrt{P^2-4Q}$ is real, we have $R_0>1\Leftrightarrow P-Q>1$. Furthermore, it is observed that $\phi(1)<0$ which implies that $\phi(x^*_{S2})$ crosses the x-axis 1 or 3 times between 0 and 1. By inspection of the form of $x^*_{R2}$ and $x^*_{G2}$, we see that $x^*_{R2}$ is in (0,1) if $x^*_{S2}$ is in (0,1) and $x^*_{G2}$ is in (0,1) if $x^*_{W2}$ is. Alternatively, we can use the equations (B.4) and (B.6) to obtain a polynomial in $x^*_{W2}$ and apply the same technique. Thus, there are either 1 or 3 endemic equilibria for this system when $R_0>1$. Thus, if $(x^*_{S2},x^*_{W2})\in[0,1]\times[0,1]$, then the remaining values, $x^*_{R2}$ and $x^*_{G2}$ are also in $[0,1]$.

## B.2 Patches 1 and 2, 1 host 1 vector, with VT, unidirectional migration of infected vectors

To determine the basic reproductive number for section 3.3.3, we apply the next generation matrix method [85] to (3.6). We determine the relevant matrices to be

$$F_1 = \begin{pmatrix} 0 & 0 & \beta_{W2}\frac{N^*_{G2}}{N^*_{W2}} & 0 \\ 0 & 0 & 0 & \beta_W\frac{N^*_{G1}}{N^*_{W1}} \\ \beta_G & 0 & p_W r_W\left(1-\frac{N^*_{W2}}{K_W}\right) & 0 \\ 0 & \beta_G & 0 & p_W r_W\left(1-\frac{N^*_{W1}}{K_W}\right) \end{pmatrix}$$

and

$$V_1 = \begin{pmatrix} \mu_G & -\tilde{b}_G & 0 & 0 \\ 0 & \mu_G+\bar{b}_G & 0 & 0 \\ 0 & 0 & \mu_W & 0 \\ 0 & 0 & 0 & \mu_W \end{pmatrix}.$$

After computing $F_1V_1^{-1}$, we obtain the dominant eigenvalue as

$$\max\left\{\frac{1}{2}\left(p_W+\sqrt{4\frac{\beta_G\beta_W}{(\mu_G+\bar{b}_G)\mu_W}\frac{N^*_{G1}}{N^*_{W1}}+p_W^2}\right),\ \frac{1}{2}\left(p_W+\sqrt{4\frac{\beta_G\beta_{W2}}{\mu_G\mu_W}\frac{N^*_{G2}}{N^*_{W2}}+p_W^2}\right)\right\}. \tag{B.7}$$

To study the existence of endemic equilibria, we first determine

$$\bar{N}^*_{G2} = K_G\left(1-\frac{\mu_G-\tilde{b}_G x^*_{G1}}{r_G}\right),\ \ \bar{N}^*_{G1} = K_G\left(1-\frac{\mu_G+\bar{b}_G x^*_{G1}}{r_G}.\right)$$

The remaining equilibrium conditions are given by

$$\begin{aligned} \beta_{W2}x^*_{W2}(1-x^*_{G2})\bar{N}^*_{G2} - \mu_G x^*_{G2}\bar{N}^*_{G2} + \tilde{b}_G x^*_{G1}\bar{N}^*_{G1} &= 0 \\ \beta_G x^*_{G2}\bar{N}^*_{G2}(1-x^*_{W2}) - \mu_W x^*_{W2}N^*_{W2}(1-p_W) &= 0 \\ \beta_W x^*_{W1}(1-x^*_{G1})\bar{N}^*_{G1} - \mu_G x^*_{G1}\bar{N}^*_{G1} - \bar{b}_G x^*_{G1}\bar{N}^*_{G1} &= 0 \\ \beta_G x^*_{G1}\bar{N}^*_{G1}(1-x^*_{W1}) - \mu_W x^*_{W1}N^*_{W1}(1-p_W) &= 0 \end{aligned} \tag{B.8}$$

As $\bar{N}^*_{G1} = K_G\left(1 - \frac{\mu_G + \bar{b}_G x^*_{G1}}{r_G}\right) \neq 0$ since $r_G > \mu_G + \bar{b}_G$, we can divide both sides of the 3rd equation by $\bar{N}^*_{G1}$ to obtain

$$x^*_{G1} = \frac{\beta_W x^*_{W1}}{\beta_W x^*_{W1} + \mu_G + \bar{b}_G}.$$

It can be seen that if $0 \leq x^*_{W1} < 1$, then $0 \leq x^*_{G1} < 1$. To show $0 < x^*_{W1} < 1$, we substitute this expression for $x^*_{G1}$ into the 4th equation in the equilibrium conditions and expand to obtain a cubic polynomial in $x^*_{W1}$. We can factor out the disease free equilibrium (in patch 1), in which $x^*_{W1} = 0$ leads us to the simple one host one vector patch 2 only endemic equilibrium. Thus, if $R_1 < 1 < R_2$, we have existence of precisely one endemic equilibrium. Otherwise, for $x^*_{W1} > 0$, our resulting quadratic polynomial is $\zeta(x^*_{W1}) = a_2 {x^*_{W1}}^2 + a_1 x^*_{W1} + a_0 = 0$ with

$$a_0 = r_G \mu_W N^*_{W1} (\bar{b}_G + \mu_G)^2 \left(\frac{\beta_G \beta_W N^*_{G1}}{(\bar{b}_G + \mu_G)\mu_W N^*_{W1}} + p_W - 1\right),$$

$$a_1 = -\beta_W[\beta_G K_G((\mu_G + \bar{b}_G)(r_G - \mu_G) - \beta_W(r_G - (\mu_G + \bar{b}_G))) + 2(\mu_G + \bar{b}_G) r_G \mu_W N^*_{W1}(1 - p_W)],$$

$$a_2 = -\beta^2_W[\beta_G K_G(r_G - (\mu_G + \bar{b}_G)) + r_G \mu_W N^*_{W1}(p_W - 1)]$$

We first observe that the constant term, $\zeta(0) = a_0$ is positive if and only if $R_1 > 1$ by the result below:

$$\begin{aligned}
R_1 &> 1 \\
\frac{1}{2}\left(p_W + \sqrt{4\frac{\beta_G \beta_W N^*_{G1}}{(\mu_G + \bar{b}_G)\mu_W N^*_{W1}} + p^2_W}\right) &> 1 \\
\left(p_W + \sqrt{4\frac{\beta_G \beta_W N^*_{G1}}{(\mu_G + \bar{b}_G)\mu_W N^*_{W1}} + p^2_W}\right) &> 2 \\
\sqrt{4\frac{\beta_G \beta_W N^*_{G1}}{(\mu_G + \bar{b}_G)\mu_W N^*_{W1}} + p^2_W} &> 2 - p_W \\
4p_W + 4\frac{\beta_G \beta_W N^*_{G1}}{(\mu_G + \bar{b}_G)\mu_W N^*_{W1}} - 4 &> 0 \\
p_W + \frac{\beta_G \beta_W N^*_{G1}}{(\mu_G + \bar{b}_G)\mu_W N^*_{W1}} - 1 &> 0
\end{aligned}$$

Furthermore $\zeta(1) < 0$. Thus, we can conclude that $\zeta(x^*_{W1})$ crosses the x-axis precisely once between 0 and 1.

In order to show that $0 \leq x^*_{G2} < 1$ and $0 \leq x^*_{W2} < 1$, we solve the first equilibrium condition for $x^*_{G2}$ in terms of $x^*_{W2}$ and $x^*_{G1}$,

$$\begin{aligned} x^*_{G2} &= \frac{\beta_{W2} x^*_{W2} \bar{N}^*_{G2} + \tilde{b}_G x^*_{G1} N^*_{G1}}{\beta_{W2} x^*_{W2} \bar{N}^*_{G2} + \mu_G \bar{N}^*_{G2}} \\ &= \frac{\beta_{W2} x^*_{W2} + \tilde{b}_G x^*_{G1} \frac{\bar{N}^*_{G1}}{\bar{N}^*_{G2}}}{\beta_{W2} x^*_{W2} + \mu_G} \end{aligned}$$

We note that if $x^*_{W2}$ is in $(0,1)$, $\frac{\bar{N}^*_{G1}}{\bar{N}^*_{G2}} < 1$ and $\mu_G > \bar{b}_G$, then $0 \leq x^*_{G2} < 1$. To show $0 \leq x^*_{W2} < 1$, we substitute the expression for $x^*_{G2}$ into the 2nd equilibrium condition from (B.8):

$$\begin{aligned} &\eta(x^*_{W2}) = -(\beta_G \beta_{W2} \bar{N}^*_{G2} + \beta_{W2} \mu_W (1 - p_W) N^*_{W2})(x^*_{W2})^2 \\ &+ [\beta_G \beta_{W2} \bar{N}^*_{G2} - \mu_G \mu_W (1 - p_W) N^*_{W2} - \beta_G \tilde{b}_G x^*_{G1} \bar{N}^*_{G1}] x^*_{W2} + \beta_G \tilde{b}_G x^*_{G1} \bar{N}^*_{G1} = 0. \end{aligned}$$

Since we have shown $0 \leq x^*_{G1} < 1$ and $\bar{N}^*_{G2}, \bar{N}^*_{G1}$, and $N^*_{W2}$ are all nonzero, positive constants, it is clear that $\eta(0) = \beta_G \tilde{b}_G x^*_{G1} \bar{N}^*_{G1} > 0$ and $\eta(1) = -(1 - p_W)\mu_W N^*_{W2}(\beta_{W2} + \mu_G) < 0$. Thus, $\eta(x^*_{W2})$ crosses the x-axis precisely once between 0 and 1. Thus, $0 \leq x^*_{W2} < 1$ which implies that $0 \leq x^*_{G2} < 1$, so that there exists precisely one endemic equilibrium value for system (3.6) when $R_1 > 1$.

If $R_1 < 1$, we observe that the constant term, $\zeta(0) = a_0$, is negative. Furthermore, $\zeta(1)$ is also negative. Thus, $\zeta$ has either 0 or 2 roots between 0 and 1. After a numerical investigation using the parameters estimated in section 3.4, we determine that both roots of $\zeta$ are always negative for $R_1 < 1$, thus no roots are in (0,1).

B.3 Patches 1 and 2, 2 hosts 2 vectors, no VT, unidirectional migration of infected vectors

We wish to determine existence of endemic equilibria for section 3.3.6, system (3.9). We first determine that

$$N^*_{G2} = K_{G2}\left(1 - \frac{\mu_G - \tilde{b}_G x^*_{G1}}{r_G}\right), \quad N^*_{G1} = K_{G1}\left(1 - \frac{\mu_G + \bar{b}_G x^*_{G1}}{r_G}\right).$$

We then wish to solve the remaining equilibrium conditions given by

$$\begin{aligned}
(q_S\beta_R x^*_{R2} + (1-q_S)\beta_{WS}x^*_{W2})(1-x^*_{S2})N^*_{S2} - \mu_S x^*_{S2}N^*_{S2} &= 0\\
\beta_S x^*_{S2}N^*_{S2}(1-x^*_{R2}) - \mu_R x^*_{R2}N^*_{R2} &= 0\\
\beta_{W2}x^*_{W2}(1-x^*_{G2})N^*_{G2} - \mu_G x^*_{G2}N^*_{G2} + \tilde{b}_G x^*_{G1}N^*_{G1} &= 0\\
((1-q_W)\beta_G x^*_{G2}N^*_{G2} + \tilde{\beta}_{SW}x^*_{S2}N^*_{S2})(1-x^*_{W2}) - \mu_W x^*_{W2}N^*_{W2} &= 0\\
\beta_W x^*_{W1}(1-x^*_{G1})N^*_{G1} - (\mu_G + \bar{b}_G)x^*_{G1}N^*_{G1} &= 0\\
\beta_G x^*_{G1}N^*_{G1}(1-x^*_{W1}) - \mu_W x^*_{W1}N^*_{W1} &= 0
\end{aligned} \tag{B.9}$$

To simplify, we make the substitution $\tilde{\beta}_{SW} = q_W\beta_{SW}\frac{N^*_{S2}}{N^*_{W2}}$ and $\tilde{\beta}_{S2} = \beta_{S2}\frac{N^*_{S2}}{N^*_{R2}}$.

$$\begin{aligned}
(q_S\beta_R x^*_{R2} + (1-q_S)\beta_{WS}x^*_{W2})(1-x^*_{S2}) - \mu_S x^*_{S2} &= 0\\
\tilde{\beta}_{S2}x^*_{S2}(1-x^*_{R2}) - \mu_R x^*_{R2} &= 0\\
\beta_{W2}x^*_{W2}(1-x^*_{G2}) - \mu_G x^*_{G2} + \tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}} &= 0\\
((1-q_W)\beta_G x^*_{G2}\frac{N^*_{G2}}{N^*_{W2}} + \tilde{\beta}_{SW}x^*_{S2})(1-x^*_{W2}) - \mu_W x^*_{W2} &= 0\\
\beta_W x^*_{W1}(1-x^*_{G1})N^*_{G1} - (\mu_G + \bar{b}_G)x^*_{G1}N^*_{G1} &= 0\\
\beta_G x^*_{G1}N^*_{G1}(1-x^*_{W1}) - \mu_W x^*_{W1}N^*_{W1} &= 0
\end{aligned} \tag{B.10}$$

In system (B.10), the last 2 equations decouple from the system, so we may study those equations separately. By inspection, we see that the last two equations are identical to the last two equations of (B.8) with $p_W = 0$. From analysis of (B.8)

we see that either $x^*_{G1} = x^*_{W1} = 0$ or $R_1 > 1$. In this patch alone, one unique endemic equilibrium exists if and only if $R_1 > 1$ (as shown in Appendix B.2 for system (B.8)).

If $x^*_{G1} = x^*_{W1} = 0$, we determine the remaining equilibrium conditions

$$\begin{aligned}
(q_S\beta_R x^*_{R2} + (1-q_S)\beta_{WS}x^*_{W2})(1-x^*_{S2}) - \mu_S x^*_{S2} &= 0 \\
\tilde{\beta}_{S2}x^*_{S2}(1-x^*_{R2}) - \mu_R x^*_{R2} &= 0 \\
\beta_{W2}x^*_{W2}(1-x^*_{G2}) - \mu_G x^*_{G2} + \tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}} &= 0 \\
((1-q_W)\beta_G x^*_{G2}\frac{N^*_{G2}}{N^*_{W2}} + \tilde{\beta}_{SW}x^*_{S2})(1-x^*_{W2}) - \mu_W x^*_{W2} &= 0
\end{aligned} \tag{B.11}$$

which are identical to the equilibrium conditions (B.2) for patch 2 alone, in which we have determined that if $R_2 > 1$, either 1 or 3 endemic equilibrium exists, in addition to the disease free equilibrium (which always exists).

If instead, $R_1 > 1$, we assume that $x^*_{G1}$ is the positive equilibrium given in Appendix B.2 for system (3.6) with $p_W = 0$. Thus, in determining possible endemic equilibrium values for this system, we may solve the 3rd equation of (B.10) to obtain

$$x^*_{G2} = \frac{\beta_{W2}x^*_{W2} + \tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}}}{\beta_{W2}x^*_{W2} + \mu_G} \tag{B.12}$$

We substitute $x^*_{G2}$ into the 4th equation obtaining

$$\left[\tilde{\beta}_G\left(\frac{\beta_{W2}x^*_{W2} + \tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}}}{\beta_{W2}x^*_{W2} + \mu_G}\right) + \tilde{\beta}_{SW}x^*_{S2}\right](1-x^*_{W2}) - \mu_W x^*_{W2} = 0$$

and multiplying through by the denominator we obtain

$$\begin{aligned}
&-\left(\tilde{\beta}_G\beta_{W2} + \beta_W\mu_W\right)(x^*_{W2})^2 - \tilde{\beta}_{SW}\beta_{W2}x^*_{S2}(x^*_{W2})^2 + \left(\tilde{\beta}_{SW}\beta_G - \tilde{\beta}_{SW}\mu_G\right)x^*_{S2}x^*_{W2} \\
&\quad + \left(\tilde{\beta}_G\beta_{W2} - \tilde{\beta}_G\tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}} - \mu_G\mu_W\right)x^*_{W2} + \tilde{\beta}_{SW}\mu_G x^*_{S2} + \tilde{\beta}_G\tilde{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}} = 0,
\end{aligned}$$

and dividing every term by $\mu_W, \mu_G$, the expression becomes

$$
\begin{aligned}
&-\left(\frac{\tilde{\beta}_G\beta_{W2}}{\mu_G\mu_W}+\frac{\beta_{W2}}{\mu_G}\right)(x^*_{W2})^2+\left(\frac{\tilde{\beta}_G\beta_{W2}}{\mu_G\mu_W}-1-\frac{\tilde{\beta}_G\bar{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}}}{\mu_W\mu_G}\right)x^*_{W2}\\
&-\left(\frac{\tilde{\beta}_{SW}\beta_{W2}}{\mu_G\mu_W}(x^*_{W2})^2-\left(\frac{\tilde{\beta}_{SW}\beta_{W2}}{\mu_G\mu_W}-\frac{\tilde{\beta}_{SW}}{\mu_W}\right)x^*_{W2}-\frac{\tilde{\beta}_{SW}}{\mu_W}\right)x^*_{S2}+\frac{\tilde{\beta}_G\bar{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}}}{\mu_W\mu_G}=0.
\end{aligned}
\tag{B.13}
$$

Applying the same substitutions in Appendix B.2, where the first two equations of (B.10) are identical to the first two equations of (B.2) in Appendix B.1 and letting $X=\frac{\bar{b}_G x^*_{G1}\frac{N^*_{G1}}{N^*_{G2}}}{\mu_G}$, we obtain

$$
D(1-x^*_{W2})(x^*_{W2}-\frac{1}{E})x^*_{S2}=(F+1)(x^*_{W2})^2-\left(F+\frac{X+1}{E}\right)x^*_{W2}-\frac{X}{E}. \tag{B.14}
$$

Finally, we use (B.14) and (B.5) from Appendix B.1 to obtain a polynomial in $x^*_{S2}$,

$$
\begin{aligned}
g(x^*_{S2})=&E(F+1+Dx^*_{S2})[A(B+1)x^*_{S2}+(1-AB)]^2(x^*_{S2})^2\\
&-(FX+Dx^*_{S2})[C(Bx^*_{S2}+1)(1-x^*_{S2})]^2+[1-EF+FX+(1-E)Dx^*_{S2}]\\
&[A(B+1)x^*_{S2}+(1-AB)]x^*_{S2}[C(Bx^*_{S2}+1)(1-x^*_{S2})]=0.
\end{aligned}
\tag{B.15}
$$

We note that

$$
x^*_{W2}=\frac{x^*_{S2}\left[(A+1)x^*_{S2}-A+\frac{1}{B}\right]}{C(1-x^*_{S2})\left(x^*_{S2}+\frac{1}{B}\right)}. \tag{B.16}
$$

is positive if and only if

$$
x^*_{S2}>\frac{AB-1}{AB+A},
$$

in which case $AB>1$ for the expression to be positive (biologically significant).

Thus, we determine that $g(0)<0$, $g(\frac{AB-1}{AB+A})<0$ and $g(1)>0$, so that $g$ has at least one root in $\left(\frac{AB-1}{AB+A},1\right)$.

## B.4 Patches 1, 2, and 3, 2 hosts 2 vectors, no VT, unidirectional migration of infected vectors

To study the existence of endemic equilibria for system (3.10) in section 3.3.7, we first determine that

$$N^*_{S3} = K_{S3}\left(1 - \frac{\mu_S - \tilde{b}_S x^*_{S2}}{r_S}\right), \quad N^*_{S2} = K_{S2}\left(1 - \frac{\mu_S + \bar{b}_S x^*_{S2}}{r_S}\right).$$

$$N^*_{G2} = K_{G2}\left(1 - \frac{\mu_G - \tilde{b}_G x^*_{G1}}{r_G}\right), \quad N^*_{G1} = K_{G1}\left(1 - \frac{\mu_G + \bar{b}_G x^*_{G1}}{r_G}\right).$$

The remaining equilibrium conditions are given by

$$\begin{aligned}
\beta_R x^*_{R3}(1 - x^*_{S3})N^*_{S3} - \mu_S x^*_{S3} + \tilde{b}_S x^*_{S2} N^*_{S2} &= 0 \\
\beta_S x^*_{S3} N^*_{S3}(1 - x^*_{R3}) - \mu_R x^*_{R3} N^*_{R3} &= 0 \\
((1 - q_S)\beta_R x^*_{R2} + q_S \beta_{WS} x^*_{W2})(1 - x^*_{S2})N^*_{S2} - (\mu_S + \bar{b}_S)x^*_{S2} N^*_{S2} &= 0 \\
\beta_{S2} x^*_{S2} N^*_{S2}(1 - x^*_{R2}) - \mu_R x^*_{R2} N^*_{R2} &= 0 \\
\beta_{W2} x^*_{W2}(1 - x^*_{G2})N^*_{G2} - \mu_G x^*_{G2} N^*_{G2} + \tilde{b}_G x^*_{G1} N^*_{G1} &= 0 \\
((1 - q_W)\beta_G x^*_{G2} N^*_{G2} + q_W \beta_{SW} x^*_{S2} N^*_{S2})(1 - x^*_{W2}) - \mu_W x^*_{W2} N^*_{W2} &= 0 \\
\beta_W x^*_{W1}(1 - x^*_{G1})N^*_{G1} - (\mu_G + \bar{b}_G)x^*_{G1} N^*_{G1} &= 0 \\
\beta_G x^*_{G1} N^*_{G1}(1 - x^*_{W1}) - \mu_W x^*_{W1} N^*_{W1} &= 0
\end{aligned} \tag{B.17}$$

If $x^*_{G1} = x^*_{W1} = 0$, the remaining equilibrium conditions are

$$\begin{aligned}
\beta_R x^*_{R3}(1 - x^*_{S3})N^*_{S3} - \mu_S x^*_{S3} + \tilde{b}_S x^*_{S2} N^*_{S2} &= 0 \\
\beta_S x^*_{S3} N^*_{S3}(1 - x^*_{R3}) - \mu_R x^*_{R3} N^*_{R3} &= 0 \\
((1 - q_S)\beta_R x^*_{R2} + q_S \beta_{WS} x^*_{W2})(1 - x^*_{S2})N^*_{S2} - (\mu_S + \bar{b}_S)x^*_{S2} N^*_{S2} &= 0 \\
\beta_{S2} x^*_{S2} N^*_{S2}(1 - x^*_{R2}) - \mu_R x^*_{R2} N^*_{R2} &= 0 \\
\beta_{W2} x^*_{W2}(1 - x^*_{G2})N^*_{G2} - \mu_G x^*_{G2} N^*_{G2} &= 0 \\
((1 - q_W)\beta_G x^*_{G2} N^*_{G2} + q_W \beta_{SW} x^*_{S2} N^*_{S2})(1 - x^*_{W2}) - \mu_W x^*_{W2} N^*_{W2} &= 0
\end{aligned} \tag{B.18}$$

If $R_1 < 1$, $R_2 > 1$, $x^*_{S2}, x^*_{W2}$, and $x^*_{G2}$ are all positive, then it is clear that $x^*_{R3} = x^*_{S3} = 0$ is not a solution to system (B.18). Thus, we have existence of 1 or more endemic equilibrium in patches 2 and 3 alone. On the other hand, if $x^*_{G1}$ and $x^*_{W1}$ are positive, then infection must persist in all patches, i.e. $x^*_{G2} = x^*_{W2} = x^*_{S2} = x^*_{R2} = x^*_{R3} = x^*_{S3} = 0$ is not a solution to system (B.17). Intuition suggests that there exists one endemic equilibria for patch 3 alone if $R_1 < 1$ and $R_2 < 1$. In this case, if we set $I^*_{S2} = I^*_{R2} = I^*_{W2} = I^*_{G2} = 0$, system (3.10) breaks down to the simple one host one vector model in which there exists precisely one endemic equilibrium if and only if $R_3 > 1$.

# APPENDIX C

# MODEL EQUATIONS AND PARAMETERS

## C.1 Equations and assumptions

The system contains 9,376 equations. Each cell contains anywhere from 4 to 8 equations, depending on the specific patch location. The equations in (C.1) are representative of a patch 2 interior cell, which contains both species of vector and both hosts. Each state variable contains two subscripts: one identifying the species (S-*T. sanguisuga*, G-*T. gerstaeckeri*, R-raccoon, W-woodrat) and the other representing the cell location ($i$ represents the current cell, while $n$, $s$, $e$, $w$ represents the 4 possible adjacent cells (north, south, east, and west) to the current cell. Each vector equation contains 8 migration terms, representing the bidirectional movement between cells assuming a von Neumann radius. Not every equation in every cell will contain all migration terms. Specific migration terms will be 0 if there is no corresponding vector species population in the specific adjacent cell or if the current cell is on a grid boundary.

Table C.1. Parameter definitions and values

| Param. | Value | Definition | Source |
|---|---|---|---|
| Demographic parameters | | | |
| $r_S$ | $33/yr$ | species specific growth rates | [44] |
| $r_G$ | $100/yr$ | | |
| $r_R$ | $0.9/yr$ | | |
| $r_W$ | $1.8/yr$ | | |
| $\mu_S$ | $0.271/yr$ | species specific natural mortality rates | [44] |
| $\mu_G$ | $0.562/yr$ | | |
| $\mu_R$ | $0.4/yr$ | | |
| $\mu_W$ | $1/yr$ | | |
| $K_{S3}$ | 31900 vec/km$^2$ | species specific carrying capacities | [44] |
| $K_{R3}$ | 35.6 racc/km$^2$ | | |
| $K_{S2}$ | 19277 vec/km$^2$ | | |
| $K_{R2}$ | 13.1 racc/km$^2$ | | |
| $K_{G2}$ | 12553 vec/km$^2$ | | |
| $K_{W2}$ | 2542.5 wr/km$^2$ | | |
| $K_{G1}$ | 31900 vec/km$^2$ | | |
| $K_{W1}$ | 5200 wr/km$^2$ | | |
| Infection parameters | | | |
| $\beta_S$ | 0.00025 hosts/vec/yr | *T. sanguisuga* to raccoon inf. rate (patch 3) | Ch. 3 |
| $\beta_G$ | 0.079 hosts/vec/yr | *T. gerstaeckeri* to woodrat inf. rate (patch 1) | |
| $\beta_{SW}$ | 0.132 hosts/vec/yr | *T. sanguisuga* to woodrat inf. rate (patch 2) | |
| $\beta_{S2}$ | 0.00017 hosts/vec/yr | *T. sanguisuga* to raccoon inf. rate (patch 2) | |
| $\beta_R$ | $0.910/yr$ | raccoon to *T. sanguisuga* inf. rate (patch 3) | |
| $\beta_W$ | $1.408/yr$ | woodrat to *T. gerstaeckeri* inf. rate (patch 1) | |
| $\beta_{WS}$ | $0.116/yr$ | woodrat to *T. sangiusuga* inf. rate (patch 2) | |
| $\beta_{W2}$ | $1.408/yr$ | woodrat to *T. gerstaeckeri* inf. rate (patch 2) | |
| Migration parameters | | | |
| $M_{sN}$ | | *T. sanguisuga* migration rate-north | Ch. 2 |
| $M_{sE}$ | | *T. sanguisuga* migration rate-east | |
| $M_{sS}$ | | *T. sanguisuga* migration rate-south | |
| $M_{sW}$ | | *T. sanguisuga* migration rate-west | |
| $M_{gN}$ | | *T. gerstaeckeri* migration rate-north | |
| $M_{gE}$ | | *T. gerstaeckeri* migration rate-east | |
| $M_{gS}$ | | *T. gerstaeckeri* migration rate-south | |
| $M_{gW}$ | | *T. gerstaeckeri* migration rate-west | |

$$
\begin{aligned}
S'_{Si} &= r_S \left(1 - \frac{N_{Si}}{K_{Si}}\right) N_{Si} - \left(q_S \beta_R \frac{I_{Ri}}{N_{Ri}} + (1 - q_S)\beta_{WS} \frac{I_{Wi}}{N_{Wi}}\right) S_{Si} - \mu_S S_{Si} \\
&\quad - (M_{sN} + M_{sE} + M_{sS} + M_{sW}) S_{Si} + M_{sN} S_{Ss} + M_{sE} S_{Sw} + M_{sS} S_{Sn} + M_{sW} S_{Se} \\
I'_{Si} &= \left(q_S \beta_R \frac{I_{Ri}}{N_{Ri}} + (1 - q_S)\beta_{WS} \frac{I_{Wi}}{N_{Wi}}\right) S_{Si} - \mu_S I_{Si} - (M_{sN} + M_{sE} + M_{sS} + M_{sW}) I_{Si} \\
&\quad + M_{sN} I_{Ss} + M_{sE} I_{Sw} + M_{sS} I_{Sn} + M_{sW} I_{Se} \\
S'_{Ri} &= r_R \left(S_{Ri} + (1 - p_R) I_{Ri}\right) \left(1 - \frac{N_{Ri}}{K_{Ri}}\right) - \beta_{S2} \frac{I_{Si}}{N_{Ri}} S_{Ri} - \mu_R S_{Ri} \\
I'_{Ri} &= p_R r_R I_{Ri} \left(1 - \frac{N_{Ri}}{K_{Ri}}\right) + \beta_{S2} \frac{I_{Si}}{N_{Ri}} S_{Ri} - \mu_R I_{Ri} \\
S'_{Gi} &= r_G \left(1 - \frac{N_{Gi}}{K_{Gi}}\right) N_{Gi} - \beta_{W2} \frac{I_{Wi}}{N_{Wi}} S_{Gi} - \mu_G S_{Gi} S_{Gi} - (M_{gN} + M_{gE} + M_{gS} + M_{gW}) S_{Gi} \\
&\quad + M_{gN} S_{Gs} + M_{gE} S_{Gw} + M_{gS} S_{Gn} + M_{gW} S_{(Ge)} \\
I'_{Gi} &= \beta_{W2} \frac{I_{Wi}}{N_{Wi}} S_{Gi} - \mu_G I_{Gi} - (M_{gN} + M_{gE} + M_{gS} + M_{gW}) I_{Gi} \\
&\quad + M_{gN} I_{(Gs)} + M_{gE} I_{Gw} + M_{gS} I_{Gn} + M_{gW} I_{Ge} \\
S'_{Wi} &= r_W \left(S_{Wi} + (1 - p_W) I_{Wi}\right) \left(1 - \frac{N_{Wi}}{K_{Wi}}\right) - \left((1 - q_W)\beta_G \frac{I_{Gi}}{N_{Wi}} + q_W \beta_{SW} \frac{I_{Si}}{N_{Wi}}\right) S_{Wi} \\
&\quad - \mu_W S_{Wi} \\
I'_{Wi} &= p_W r_W I_{Wi} \left(1 - \frac{N_{Wi}}{K_{Wi}}\right) + \left((1 - q_W)\beta_G \frac{I_{Gi}}{N_{Wi}} + q_W \beta_{SW} \frac{I_{Si}}{N_{Wi}}\right) S_{Wi} - \mu_W I_{Wi}
\end{aligned}
\tag{C.1}
$$

REFERENCES


[1] Añez, N., East, J.S. Studies on *Trypanosoma rangeli* Tejera 1920 II. Its effect on feeding behaviour of triatomine bugs. *Acta tropica* 41 (1984) 93–95.

[2] Allen, L.J., Wesley, C.L., Owen, R.D., Goodin, D.G., Koch, D., Jonsson, C.B., Chu, Y., Hutchinson, J.M., and Paige, R.L. A habitat-based model for the spread of hantavirus between reservoir and spillover species. *Journal of Theoretical Biology*. 260 (2009) 510–522.

[3] Arino, J., Davis, J.R., Hartley, D., Jordan, R., Miller, J.M., and van den Driessche, P. A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology* 22 (2005) 129–142.

[4] Barbu, C., Dumonteil, E., Gourbière, S. Characterization of the dispersal of non-domiciliated *Triatoma dimidiata* through the selection of spatially explicit models. *PLoS Neglected Tropical Diseases*. 4(8):e777 (2010) 1–11.

[5] Beard, C.B., Pye, G., Steurer, F.J., Rodriguiz, R., Campman, R., Townsend Peterson, A., Ramsey, J., Wirtz, R.A., Robinson, L.E. Chagas disease in a domestic transmission cycle in southern Texas, USA. *Emerging Infectious Diseases*, 9 (2003) 103–105.

[6] Beasley, J.C., DeVault, T.L., Retamosa, M.I., Rhodes, O.E. Jr. A hierarachical analysis of habitat selection by raccoons in Northern Indiana. *Wildlife Damage Management, Internet Center for USDA National Wildlife Research Center-Staff Publications*. (2007)

[7] Bern, C., Montgomery, S.P. An estimate of the burden of Chagas disease in the United States. *Clinical Infectious Diseases*, 49 (2009) 52–54.

[8] Bottazzi, M.E., Dumonteil, E., Valenzuela, J.G., Betancourt-Cravioto, M., Tapia-Conyer, R., Hotez, P.J., Bridging the innovation gap for neglected tropical diseases in Mexico: capacity building for the development of a new generation of antipoverty vaccines. *Boletín médico del Hospital Infantil de México*, 68 (2011) 138–146.

[9] Box, T.W. Denisty of plains wood rat dens on four plant communities in south Texas. *Ecology*, 40 (1959) 715–716.

[10] Brauer, F. Castillo-Chavez, C., and Velasco-Hernández, J.X., Recruitment effects in heterosexually transmitted disease models. *International Journal of Applied Science and Computation*. 3 (1996) 78–90.

[11] Braun, J.K., Mares, M.A., Neotoma micropus. *Mammalian Species*, 330 (1989) 1–9.

[12] Brown, J.H., Lieverman, G.A., Dengler, W.F. Woodrats and cholla: dependence of a small mammal population on the density of cacti. *Ecology*. 53 (1972) 310–313.

[13] Burkholder, J.E. Allison, T.C. Kelly, V.P. *Trypanosoma cruzi* (Chagas) (Protozoa: Kinetoplastida) in invertebrate, reservoir, and human hosts of the Lower Rio Grande Valley of Texas. *Journal of Parasitology*, 66 (1980) 305–311.

[14] Cantrell, R.C., Cosner, C. *Spatial Ecology via Reaction-Diffusion Equations*. Chichester, West Sussex, England: John Wiley & Sons Ltd, 2003.

[15] Centers for Disease Control and Prevention (2012) Chagas Disease, Retrieved from http://www.cdc.gov/parasites/chagas

[16] Cherif, A., Garcia Horton, V., Melendez Rosario, G. Feliciano, W., A tale of two regions: A mathematical model for Chagas' disease. *MTBI Technical Report* MTBI 05-05M. Arizona State University 2008.

[17] Clark, C.G., Pung, O.J. Host Specificity of Ribosomal DNA Variation in Sylvatic *Trypanosoma Cruzi* from North America. *Molecular and Biochemical Parisitology*, 66 (1994) 175–179.

[18] Conditt, S.A., Ribble, D.O. Social organization of *Neotoma micropus*, the southern plains woodrat. *Am. Midl. Nat.*, 137 (1996) 290–297.

[19] Cruz-Reyes, A., Pickering-López, J.M. Chagas disease in Mexico: and analysis of geographical distribution during the past 76 years. *Mem Inst Oswaldo Cruz.* 101 (2006) 345–354.

[20] Curto de Casas, S.I., Carcavalla, R.U., Galíndez Girón, I., Burgos, J.J. Bioclimatic factors and zones of life. *Atlas of Chagas' Disease Vectors in the Americas.* Rio de Janeiro 1999.

[21] Curto de Casas, S.I., Carcavallo, R.U. Climate change and vector-borne diseases distribution. *Social Science and Medicine*, 40 (1995) 1437–1440.

[22] Davis, W.B., Schmidley, D.J. *The Mammals of Texas Online Edition* Texas Tech University, 1997

[23] Devillers, H., Lobry, J.R., Menu, F. An agent-based model for predicting the prevalence of *Trypanosoma cruzi* I and II in their host and vector populations. *Journal of Theoretical Biology*, 255 (2008) 307–315.

[24] Diekmann, O., Heesterbeek, J. A. P., Metz, J.A., On the definition and the computation of the basic reproduction ratio, $R_0$ in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28 (1990) 365–380.

[25] Dorn, P., Buekens, P., Hanford, E. Whac-a-mole: Future trends in Chagas disease transmission and a global perspective on disease control. *Future Microbiology*, 2 (2007) 365–367.

[26] Dorn, P.L., Monroy, C., Curtis, A. *Triatoma dimidiata* (Latreille, 1811): A Review of its Diversity Across its Geographic Range and the Relationship Among Populations. *Infection, Genetics and Evolution*, 7 (2007) 343–352.

[27] Eads, R.B., Trevino, H.A., Campos, E.G. Triatoma (Hemiptera:Reduviidae) infected with *Trypanosoma Cruzi* in south Texas wood rat dens. *The Southwestern Naturalist*, 8 (1963) 38–42.

[28] Fisher, R.A., The wave of advance of advantageous genes. *Annals of Eugenics*, 7 (1937) 355–369.

[29] Fritzell, E.K., Haroldson, K.J., Urocyon cinereoargenteus. *Mammalian Species*, 189 (1982) 1–8.

[30] Fu, S., Milne, G. Epidemic modelling using cellular automata. *Proceedings of the Austrailian Conference on Artificial Life*, 2003.

[31] Gehrt, S.D., Fritzell, E.K. Duration of familial bonds and dispersal patterns for raccoons in south Texas. *J. Mammalogy*, 79 (1998) 859–872.

[32] Gourbière, S., Dumonteil, E., Rabinovich, J.E., Minkoue, R., Menu, F., Demographic and dispersal constraints for domestic infestation by non-domicilated Chagas disease vectors in the Yucatan Peninsula, Mexico. *American Journal of Tropical Medicine and Hygiene*. 78 (2008) 133–139.

[33] Guerenstein, P.G., Lazzari, C.R., Host-seeking: How triatomines acquire and make use of information to find blood. *Acto Tropica*, 110 (2009) 148–158.

[34] Gurevitz, J.M., Ceballos, L.A., Kitron, U., Gürtler, R.E. Flight initiation of *Triatoma Infestans* (Hemiptera: Reduviidae) under natural climatic conditions. *Journal of Medical Entomology*, 43 (2006) 143–150.

[35] Hanford, E.J., Zhan, F.B., Lu, Y., Giordano, A., Chagas disease in Texas: recognizing the significance and implications of evidence in the literature. *Social Science and Medicine*, 65 (2007) 60–79.

[36] Harry, M., Lema, F., Romaña, C.A. Chagas' disease challenge. *The Lancet*, 355 (2000) 236.

[37] Henner C.M., Chamberlain, M.J., Leopold, B.D., Wes Burger Jr., L. A multi-resolution assessment of raccoon den selection. *The Journal of Wildlife Management*, 68 (2004) 179–187.

[38] Huang, W., Han, M., Liu, K. Dynamics of an SIS reaction-diffusion epidemic model for disease transmission. *Mathematical Biosciences and Engineering*, 7 (2010) 51–66.

[39] Ibarra-Cerdeña, C.N., Sánchez-Cordero, V., Peterson, A.T., Ramsey, J.M. Ecology of north american triatominae. *Acta Tropica*, 110 (2009) 178–186.

[40] Ikenga, J.O., Richerson, J.V. *Trypanosoma cruzi* (Chagas) (Protozoa: Kinetoplastida: Trypanosomatidae) in invertebrate and vertebrate hosts from Brewster County in Trans-Pecos Texas. *Journal of Economic Entomology*. 77 (1984) 126–129.

[41] Kjos, S.A., Snowden, K.F., Olson, J. Biogeography and *Trypanosoma cruzi* infection prevalence of Chagas disease vectors in Texas, USA. *Vector-Borne and Zoonotic Diseases*, 9 (2009) 41–50

[42] Kot, M., Lewis, M.A., van den Driessche, P. Dispersal data and the spread of invading organisms. *Ecology*. 77 (1996) 2027–2042.

[43] Kribs Zaleta, C., Alternative transmission modes for *Trypanosoma cruzi*. *Mathematical Bioscences and Engeenering*, 7 (2010) 657–673.

[44] Kribs Zaleta, C. Estimating contact process saturation in sylvatic transmission of *Trypanosoma cruzi* in the U.S. *PLOS Neglected Tropical Diseases*, 44:e656 (2010) 1–14.

[45] Kribs Zaleta, C., Vector consumption and contact process saturation in sylvatic transmission of *T. cruzi*. *Mathematical Population Studies*, 13 (2006) 132–152.

[46] Lambert, R.C., Kolivras, K.N., Resler, L.M., Brewster, C.C., Paulson, S.L., The potential for emergence of Chagas disease in the United States. *Geospatial Health*, 2 (2008) 227–239.

[47] Lehane, M.J., McEwen, P.K., Whitaker, C.J., Schofield, C.J. The role of temperature and nutritional status in flight initiation by *Triatoma Infestans*. *Acta Tropica*, 52 (1992) 27–38.

[48] Lehane, M.J., Schofield, C.J., Measurement of speed and duration of *triatomine* vectors. *The Tranactions of the Royal Society of Tropical Medicine and Hygiene*, 72 (1978) 438.

[49] Lewis, M., Renclawowicz, J., van den Driessche, P. Traveling waves and spread rates for a West Nile virus model. *Bulletin of Mathematical Biology*. 68 (2006) 3–23.

[50] Linzey, A.V., Timm, R. Álvarez-Castañeda, S.T., Castro-Arellano, I., Lacher, T. *Neotoma micropus*. In: IUCN 2009. IUCN Red List of Threatened Species. Version 2009.2. http://www.iucnredlist.org. Accessed 10 March 2010.

[51] Macdonald, G. *The epidemiology and control of malaria*. Oxford: Oxford University Press, 1957.

[52] Maidana, N.A., Mo Yang, H., Describing the geographic spread of dengue disease by traveling waves. *Mathematical Biosciences*. 215 (2008) 64–77.

[53] McBee, K., Baker, R.J., Dasypus novemcinctus. *Mammalian Species*, 162 (1982) 1–9.

[54] Medlock, J., Kot, M. Spreading disease: integro-differential equations old and new. *Mathematical Biosciences*, 184 (2003) 201–222.

[55] Merkelz, R. Kerr, S.F. Demographics, den use, movements,and absence of *Leishmania Mexicana* in southern plains woodrats *Neotoma micropus*. *The Southwestern Naturalist*, 47 (2002) 70–77.

[56] Milei, J., Guerri-Guttenberg, R.A., Rodolfo Grana, D., Storino, R. Prognostic Impact of Chagas Disease in the United States. *American Heart Journal*, 159 (2009) 22–29.

[57] Mollison, D. Possible velocities for a simple epidemic. *Advances in Applied Probability*. 4 (1972) 233–258.

[58] Mollison, D. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society: Series B*, 39 (1977) 283–326.

[59] Moreno, E.Z., Rivera, I.M., Moreno, S.C., Alarcón, M.E. Lugo-Yarbuh, A., Vertical transmission of *Trypanosoma cruzi* in Wistar rats during the acute phase of infection. *Investigación clínica* (Maracaibo) 44 (2003).

[60] Murray, J.D., Stanely, E.A., Brown, D.L. On the spatial spread of rabies among foxes. *Precedings of the Royal Society of London*, 229 (1986) 111–1150.

[61] Native Prairies Association of Texas. http://texasprairie.org/ Accessed 27 Feb 2010.

[62] U.S. Environmental Protection Agency (EPA). Omernik's Level III Ecoregions of the Continental United States, in National Atlas of the United States, January 2005. http://nationalatlas.gov Accessed 15 Feb 2010.

[63] Peterson, A.T., Sánchez-Cordero, V., Beard, C.B., Ramsey, J.M. Ecological niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerging Infectious Diseases*. 8 (2002) 662–667.

[64] Pietzrak, S.M., Pung, O.J. Trypanosomiasis in Raccoons From Georgia. *Journal of Wildlife Diseases*. 34 (1998) 132–136.

[65] Pippin, W.F., The biology and vector capability of *Triatoma Sanguisuga Texana Usinger* and *Triatoma Gerstaeckeri* (Stål)(Hemiptera: Triatominae). *Journal of Medical Entomology*, 7 (1970) 30–45.

[66] Pung, O.J., Banks, C.W., Jones, D.N., Krissinger, M.W., *Trypanosoma cruzi* in wild raccoons, opossums, and triatomine bugs in southeast Georgia, USA. *Journal of Parasitology*. 81 (1995) 583–587.

[67] Raun, G.G. A population of woodrats (*Neotoma micropus*) in southwest Texas. *Bull. Texas Mem. Mus.*. 11(1966) 1–62.

[68] Raymond, R.W. McHugh, C.P., Witt, L.R. Kerr, S.F. Temporal and spatial distribution of *Leishmania mexicana* infections in a population of *Neotoma micropus*. *Mem. de Inst. Oswaldo Cruz*. 98 (2003) 171–180.

[69] Roellig, D.M., Brown, E.L., Barnabé, C., Tibayrenc, M., Steurer, F.J., Yabsley M.J. Molecular typing of *Trypanosoma cruzi* isolates, United States. *Emerging Infectious Diseases*. 14 (2008) 1123–1125.

[70] Roellig, D.M., Ellis, A.E., Yabsley, M.J. Genetically different isolates of *Trypanosoma cruzi* elicit different infection dynamics in raccoons (*Procyon lotor*) and Virginia opossums (*Didelphis virginiana*). *International Journal for Parasitology*. 39 (2009) 1603–1610.

[71] Ross, R., *The Prevention of Malaria* (2nd edition). London: Murray 1911.

[72] Sarkar, S., Strutz, S.E., Frank, D.M., Rivaldi, C-L., Sissel, B., Sànchez-Cordero, V. Chagas disease risk in Texas. *PloS Neglected Tropical Diseases*. 4 (10) (2010): e836 1–14.

[73] Schoefield, C.J., Lehane, M.J., McEwen, P., Catala, S.S., Gorla, D.E. Dispersive Flight by *Triatoma Infestans* Under Natural Climatic Conditions in Argentia. *Medical and Veterinary Entomology*, 6 (1992) 51–56.

[74] Schofield, C.J., Lehane, M.J., McEwen, P., Catala, S.S., Gorla, D.E. Dispersive flight by *Triatoma infestans* under natural climatic conditions in Argentina. *Medical and Veterinary Entomology*, 6 (1992) 51–56.

[75] Schofield, C.J., Lehane, M.J., McEwan, P., Catala, S.S., Gorla, D.E. Dispersive flight by *Triatoma sordida*. *The Tranactions of the Royal Society of Tropical Medicine and Hygiene*, 85 (1991) 676–678.

[76] Schönfisch, B. Propagation of fronts in cellular automata. *Physica D*, 80 (1995) 433–450.

[77] Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecological Modelling*, 133 (2000) 209–223.

[78] Skellam, J.G. Random dispersal in theoretical populations. *Biometrika*, 38 (1951) 196–218.

[79] Slimi, R., El Yacoubi, S., Dumonteil, E., Gourbière, S. A cellular automata model for Chagas disease. *Applied Mathematical Modelling*, 33 (2009) 1072–1085.

[80] Sun, GQ., Liu, QX, Jin, Z., Chakraborty, A., Li, BL. Influence of infection rate and migration on extinction of disease in spatial epidemics. *Journal of Theoretical Biology*, 624 (2010) 95–103.

[81] Thieme, H.R. Convergence results and a Poincaré-Bendixson trichotomy for asymptotically autonomous differential equations. *Journal of Mathematical Biology*. 30 (1992) 755–763.

[82] Thieme, H.R. Asymptotically autonomous differential equations in the plane. *Rocky Mountain Journal of Mathematics*. 24(1) Winter 1994 351–380.

[83] Thies, K.M., Thies, M.L., Caire, W. House construction by the southern plains woodrat (*Neotoma micropus*) in southwestern Oklahoma. *The Southwestern Naturalist*, 41 (1996) 116–122.

[84] Thies, M., Caire, W. Nearest-neighbor analysis of the spatial distribution of houses *Neotoma micropus* in southwestern Oklahoma. *The Southwestern Naturalist*, 36 (1991) 233–262.

[85] Van den Driessche, P., Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*. 180 (2002) 29–48.

[86] Vazquez-Prokopec, G.M., Ceballos, L.A., Kitron, U., and Grtler, R.E. Active dispersal of natural populations of *Triatoma infestans* (Hemiptera: Reduviidae) in rural northwestern Argentina. *Journal of Medical Entomology*. 41 (2004) 614–621.

[87] Vazquez-Prokopec, G.M., Ceballos, L.A., Marcet, P.L., Cecere, M.C., Cardinal, M.V., Kitron, U., Grtler, R.E. Seasonal variations in active dispersal of natural populations of *Triatoma infestans* in rural north-western Argentina. *Medical and Veterinary Entomology*. 20 (2006) 273–279.

[88] Velasco-Hernàndez, J.X. An epidemiological model for the dynamics of Chagas' disease. *Biosystems*. 26 (1991) 127–134.

[89] Villagrán, M.E., Marin, C., Hurtado, A., Sánchez-Moreno, M., Antonio de Diego, J. Natural infection and distribution of Triatomines (Hemiptera: Reduviidae) in the state of Querétaro, Mexico. *The Royal Society of Tropical Medicine and Hygiene*, 102 (2008) 833–838.

[90] Wade-Smith, J., Verts, B.J., Mephitis mephitis. *Mammalian Species*, 173 (1982) 1–7.

[91] White, S. , Martín del Rey, A., Rodríguez Sánchez, G. Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186 (2007) 193–202.

[92] World Health Organization (2010) *Chagas disease: American Trypanosomiasis*. Retrieved from http://www.who.int/mediacentre/factsheets/fs340/en

[93] Yabsley, M.J. Pittman Noblet, G. Biological and molecular characterization of a raccoon isolate of *Trypanosoma cruzi* From South Carolina. *Journal of Parasitology*. 88 (2002) 1273–1276.

[94] Yabsley, M., Barnes, J., Ellise, A., Kjos, S., Roxanne, C., Southern plains woodrats (*Neotoma micropus*) from southern Texas are important reservoirs of two genotypes of *Trypanosoma cruzi* and a host of a putative novel *Trypanosoma* species. *Vector-Borne and Zoonotic Diseases*, In Press, 2012.

[95] Zeledón, R., Rabinovich, J.E. Chagas' Disease: An Ecological appraisal with special emphasis on its insect vectors. *Annual Review of Entomology*, 26 (1981) 101–133.

[96] Zingales,B., Andrade, S.G., Briones, M.R.S., Campbell, D.A., Chiari, E., Fernandes, O., Guhl, F., Lages-Silva, E., Macedo, A.M., Machado, C.R., Miles, M.A., Romanha, A.J., Sturm, N.R., Tibayrenc, M., and Schijman, A.G. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz*. 104 (2009) 1051–1054.

## BIOGRAPHICAL STATEMENT

Britnee A. Crawford was born in Plano, Texas, USA, in 1983. She received her B.S. degree in Mathematics from Dallas Baptist University, USA in 2005, her M.S. and Ph.D. degrees in Mathematics from The University of Texas at Arlington in 2008 and 2012, respectively She has taught high school mathematics courses at Grace Preparatory Academy for 7 years, and served as an adjunct professor at Dallas Baptist University for 2 years. She is married to Christopher Crawford, and together they have a one year old son, Easton. Her current research interest is in the area of mathematical biology, particularly epidemiology.