SPATIOTEMPORAL MODELING AND ANALYSIS OF DISEASE SPREAD IN WILDLIFE

A DISSERTATION IN
Mathematics
and
Physics

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
the requirements for the degree

DOCTOR OF PHILOSOPHY

by
GERALD WALKER BAYGENTS

B.S., Georgia Southern University, 1993
M.S., University of South Carolina, 1996

Kansas City, Missouri
2018

SPATIOTEMPORAL MODELING AND ANALYSIS OF DISEASE SPREAD IN WILDLIFE

Gerald Walker Baygents, Candidate for the Doctor of Philosophy Degree

University of Missouri-Kansas City, 2018

ABSTRACT

Wildlife and wildlife diseases have been frequent topics in mathematical epidemiology. However, due to the complexity of real-world systems and the varying degree of randomness in the behavior of any one individual organism, it can be difficult to obtain reliable and accurate spatiotemporal results with any given methodology. In this work, we look at hemorrhagic diseases (HD) in white-tailed deer as a case study to explore statistical and mathematical modeling techniques for analyzing disease spread in wildlife. We concentrate on two modeling approaches to evaluate their capabilities and usefulness in predicting and analyzing the dynamics of wildlife diseases. Statistical modeling implemented with SaTScan enables us to identify significant clusters of disease activity, clusters that are significant with respect to geography or time or both. The spatial clusters of years 1980, 1988, 2007, 2012, and 2013 suggest patterns of outbreaks every six to eight years, with the next potential outbreak during 2018 - 2020. Using mathematical modeling with ordinary differential equations (ODE), we derive a model for the dynamics of the disease that includes the migration of the host. We also derive the basic reproduction number R_0 of this model to uncover the

conditions that lead to an outbreak of the disease. In addition, we also apply several techniques using MATLAB to estimate the parameters of such a set of ODE which are useful when the available data set is limited in size.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Graduate Studies, have examined a dissertation titled "Spatiotemporal Modeling and Analysis of Disease Spread in Wildlife", presented by Gerald Walker Baygents, candidate for the Doctor of Philosophy degree, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Majid Bani-Yaghoub, Ph.D., Committee Chair
Department of Mathematics and Statistics

Noah Rhee, Ph.D.
Department of Mathematics and Statistics

Naveen Vaidya, Ph.D.
Department of Mathematics and Statistics, San Diego State University

Mark Brodwin, Ph.D.
Department of Physics and Astronomy

Paul Rulis, Ph.D.
Department of Physics and Astronomy

CONTENTS

ILLUSTRATIONS

TABLES

ACKNOWLEDGMENTS

I would like to thank Dr. Majid Bani-Yaghoub for his guidance and patience, constantly reminding me that the end was in sight and attainable. I would also like to thank the Missouri Department of Conservation for creating the spark for this project and providing the data used within.

DEDICATION

To everyone who said I'd get through this through during all the times I was convinced I wouldn't.

CHAPTER 1

INTRODUCTION

Hemorrhagic disease (HD) is a fatal disease of white-tailed deer (*Odocoileus virginianus*). It is the collective term used for epizootic hemorrhagic disease (EHD) and bluetongue disease (genus *Orbivirus*). Because symptoms caused by EHD and bluetongue are nearly indistinguishable, they are frequently grouped together and referred to as hemorrhagic disease (HD) [38, 42], and the first suspected outbreak of HD in the USA occurred in the 1890s [19]. Clinical signs include swelling in the head, tongue, neck, and lungs due to fluid accumulation [6, 19]. Other symptoms include swelling and hemorrhage throughout the body, sloughing of hooves, and may also include sores or ulcers to form on the deers tongue or on portions of the stomach and on the roof of the mouth, high fever, and loss of fear of humans [11, 50]. Hemorrhagic diseases are expressed in three different forms: peracute, acute, and chronic. The peracute form is the quickest of the HD forms. It progresses quickly and it can cause death within a week. The acute HD causes death within one to two weeks [13], and both the peracute and acute forms can cause a deer to become sluggish and to develop a high fever, and this may be why infected deer die in or close to sources of water. The chronic form of HD consists of nearly 15% of the cases, in which the infected deer will survive with some degree of tissue damage [11]. Secondary infections may lead to death; but in those female deer that do survive, antibodies to the HD virus will be passed to her offspring [12, 38]. It is possible for a deer to survive, but it is

rare. In addition to white-tailed deer, HD can be transmitted to other wild ruminants and domestic animals, most commonly hoof stock, but it rarely causes disease. The infection does not affect humans or non-ruminant animals [11].

Deer Society

White-tailed deer live in forested areas and around the borders of clearings. The size of their home range varies, largely determined by the abundance of food. Females may use 35 to 320 acres (less than a square mile), while males may use 180 to 1200 acres (just under 2 square miles). This range increases during the mating season in October through December. Fawns are born in late May and early June, and their average life span in the wild is three to five years (though it may be as high as ten or fifteen years). In captivity, this increases to twenty to twenty five years. (MDC field guide).

Severinghaus and Cheatum[49] define social organization in white-tailed deer as limited to the family group - an older doe with her fawns, sometimes with her previous years offspring. Hawkins et al. [16] broaden this to include any group involving does and fawns that are spatially and socially related (frequency of association between all members is 50% or more) over a substantial period of time (usually several months). The two most common grouping are an adult (typically doe)/yearling doe/two fawns and a doe/two sibling fawns. Buck groups are not common. Large groups of 25 to 30 consisting of intermingling of family and buck groups are somewhat common in late winter and early spring, but these groups are temporary, and the deer do not move together as a herd [16]. Fawns break off from their mother when two or three years old, some regrouping later in life. Caton [7] reports that yearling bucks

leave the family group after one year.

White-tailed deer tend to not travel very far as this exposes them to danger. Within a season, when resources in one area become sparse, average distance to new area is three miles [16]. Even during winter migration, groups tend to stay within a radius of a few miles of their home range. They may travel up to twelve miles to reach winter ranges, depending on weather and terrain [57]. Fawns accompany their mother to the winter range, and several social units may use the same wintering area. Groups may return to the same area (yard) from year to year.

The Biting Midge

The vector that spreads HD is small biting midge (*Culicoides Ceratopogonidae*), also known as a sand gnat, sandfly, or no-see-um. These midges are tiny, blood-sucking flies that are merely pests to humans, but they are the vectors in the spread of the disease in deer and livestock. The female midge requires a blood meal in order to produce eggs. The virus is transferred to the deer (or from infected deer to midge) during this blood meal. Midges lay their eggs in muddy areas; the eggs hatch and live as larvae in shallow water until they mature into their adult form as a winged insect which leaves the water to reproduce. The insect's entire life cycle is one to three months, but can vary based on temperature. As this cycle is repeated during the warm months (late spring and summer), midge populations increase rapidly and this coincides with the largest outbreaks of HD when midge populations have reached their peak [11]. *Culicoides* typically feed at dusk and through the night and are most active in late spring and early summer. They are most abundant near potential breeding sites, but will disperse for food and to mate. They are weak fliers and typ-

ically disperse no more than about a mile from the site of larval development, with females flying farther than males [45]. As they are weak fliers, their flying activity is greatly reduced in windy conditions. They may fly as far as six miles or more, but this is very rare [48]. Freezing temperatures kill the adult midges and brings an end to outbreaks. Warm winter weather seems to intensify outbreaks of HD the following summer, which implies that more viruses are able to remain viable over a warm winter, though the reservoir of infection is uncertain. HD viruses cannot survive outside a host or biting fly vector. In fact, when an infected deer dies, the virus will quickly deteriorate, and samples of the virus are rarely obtained beyond twenty four hours after death.

Weather has an effect on both the midge population and the life cycle of the HD virus. Midge populations thrive in damper areas, and in 2012, there was an above average amount of rain in the late winter / early spring, filling ponds and other water bodies. In addition, record warm temperatures in that spring and summer caused midges to become more active sooner than normal. Next, the high temperatures caused water sources to dry up, and not only did the resulting mud flats become ideal breeding areas for subsequent generations of midges, but also caused deer to visit water sources more frequently due to lower water content in the plants they ate as part of their diet. These same high temperatures also cause female midges to lay more eggs, and Wittmann and Baylis also revealed that higher temperatures decrease the extrinsic incubation period of the HD virus within the midges [62]. Thus, the virus develops faster and allows a midge to infect more deer during its life span. Periods of biting midge abundance coincide with the seasonal occurrence of

HD. Lower temperatures in fall and winter bring an end to HD outbreaks, but how the virus survives through the winter is not clear. It is known that some ruminants may carry the virus for several weeks. Some believe that in some areas with a milder climate, midge populations may remain at least partially active enough to sustain a year-round transmission. Others believe that the virus could over-winter though the colder conditions. Unfortunately, while there are several theories, none of them have been proven.

HD Mortality and Prevention

HD occurs often, but its range and toll on the deer population varies greatly and regional mortality rates differ. In the Southeast United States, the infection is mild with little mortality; however, mortality rates are high in the Midwestern US and northeastern regions. Usually death losses during an outbreak are below 25%, but a few instances of 50% or more have been recorded. High density herds may have greater mortality rates, but the correlation between deer density in an area and the severity of HD is not clear. Other contributing factors to mortality rates may include the number of deer that are immune, the amount of livestock (as reservoirs of infection) in the area, and the abundance of midge vectors.

Currently, there is little that can be done to prevent HD. Risks are minimized in smaller herds, and the most practical means of regulating herd population is through sport hunting. Deer populations in an area will fluctuate for a variety of reasons. In addition to natural birth and death rates, deer population may decline due to disease. In Missouri, large-scale outbreaks of HD are sporadic (occurring roughly every six to eight years) and are frequently localized even at the county level

- and usually have no long-term effects. The availability of food affects the population in multiple ways. In addition to increased mortality due to lower food levels, deer population will decline as a herd has to search farther and more frequently for food. This search may lead to areas that leave them more exposed and vulnerable to predators and humans. In 2012, the central region of Missouri suffered its third HD outbreak in five years in addition to a poor acorn crop, and deer populations over the next few years was below average. As a prey species, deer and deer population are linked with local predators. In Missouri, the coyote is one such predator. Some coyote predator studies have been done, but these are admittedly outdated. However, it has been shown that deer make up a portion of a coyotes diet and that large increases or decreases in predator populations may influence deer mortality rates. Hunting is a popular activity in Missouri, and regulations are put in place to ensure that hunters are allowed to harvest deer while at the same time leaving enough deer in the population to keep it stable. Regulations include not only limiting the total number of permits available, but also limiting the amount harvested. The MDC recommends harvesting 20% to 25% of the population to maintain stability. Even further, when deer population is low, restrictions may be placed on what kind of deer may be harvested. Antler point restricts can be put in place to limit the number of bucks removed from the population, allowing them to mature and breed. Then, as more bucks survive into the older age classes, buck harvest will be allowed to increase.

To lessen the chances of an HD outbreak, the MDC makes a few recommendations. People should promote vegetation around water sources to reduce midge breeding areas, encouraging the survival of other insects which compete with the

midge for resources, protecting natural predators of the midge such as fish and amphibians (especially in midge-breeding locations), restricting access of deer to water with fencing, and removing salt blocks from locations near water sources where the midge may breed. Salt blocks increases deer activity around where they are placed. Moreover, as salt blocks near water can decrease other insect population, this reduces and removes competition for midges.

Because of its very high mortality rate, HD can have a significant effect upon the deer population in a given area, reducing its numbers drastically. A common observation in outbreaks involving large numbers of deer is that they are single epizootics which do not recur. Die-offs involving small numbers of deer occur almost annually, and the disease appears to be enzootic in these areas. Hemorrhagic disease can be transmitted to other wild ruminants. The HD virus can also infect domestic animals, most commonly hoof stock, but rarely causes disease.

Up to this point, other than gathering data on the number of reported cases of HD in Missouri, very little mathematical or statistical analysis has been done in the state to attempt to study the dynamics of the disease. How the virus survives the winter is not fully understood, and there is no way to predict how the disease moves and when major outbreaks may occur. Few models have been constructed to analyze the dynamics of HD in white-tailed deer populations and dairy farms. Park et al. [42] studied these dynamics by first fitting a statistical model to predict HD incidents as a function of seroprevalence (i.e., the number of individuals in a population who tested positive for HD). Then, using ordinary differential equations (ODE), they formulated a mechanistic model to support the theory that there is a

correlation between the number of HD cases and the number of deer in a population with the virus. Their study suggests that the maximum number of cases occurs at intermediate levels of this seroprevalence. By constructing a realistic model, we will be able to analyze and simulate the dynamics of HD. A better understanding of HD dynamics gives epidemiologists and biologists the capacity to control and predict future epidemics in white-tailed deer populations. The present work is the first step toward realistic modeling of HD dynamics with a focus on migrating effects of white-tailed deer population.

Statement of the Problem

In this dissertation, we will explore how we can use mathematical and statistical techniques to model the spread of disease in wildlife, using HD as a case study. It is (relatively) easy enough to collect data on the number of cases of a disease, but those values don't tell the whole story. Deeper statistical analysis is needed to look for trends in the number and location of incidents. We will be using SaTScan to accomplish this. SaTScan is able to identify significant clusters of HD incidents, clusters significant in geography, significant in time, or both. Using this cluster analysis, we will be able to identify which areas of Missouri are likely to the locations of smaller outbreaks. We will also use temporal analysis to determine the time and frequency of major outbreaks throughout the entire state and discover which areas of the state are starting to see trends of HD increase.

In addition, we need to be able pair this analysis mathematical techniques - using delay and ordinary differential equations to attempt to model how populations of hosts and vectors change over time. We then need to be able to make adjustments

to that model to adapt to the ever-changing dynamics of real life. Furthermore, we must also be able to adapt to smaller data sets. When information is scarce (or missing), the modeling process becomes more difficult, yet we must still be able to provide results that not only mimic what actually happens in wildlife but is also able to make reasonable predictions on future population behaviors. Thus, we will propose a model that includes migration effect to study the dynamics of HD, discover what conditions may lead to an outbreak, and estimate the parameters within this ecological system.

The rest of this work is organized as follows. Chapter 2 contains an analysis of HD within Missouri by using SaTScan to identify significant clusters of HD with data provided by the Missouri Department of Conservation. In Chapter 3, we propose and analyze a delay differential system of equations to mathematically model the dynamics of the disease. In Chapter 4, we use MATLAB to estimate some of the parameters used in the mathematical model. Finally, in Chapter 5, we discuss the results found, note some of the major limitations of the work to this point, and suggest some ideas and topics for future study.

CHAPTER 2

SATSCAN AND STATISTICAL ANALYSIS

Overview

Outbreaks of deer hemorrhagic disease (HD) have been documented in the USA for many decades. In the year 2012, there was a severe HD outbreak in Missouri with mortalities reaching approximately 6.9 per thousand. Using the data of suspected HD occurrence in Missouri, the primary goal of this chapter is to statistically determine if HD in Missouris white-tailed deer occurs in spatial clusters. As shown in this chapter, the spatial clusters of years 1980, 1988, 2007, 2012, and 2013 suggest patterns of outbreaks every six to eight years, with a potential outbreak in years 2018-2020. Moreover, these spatial clusters were more frequent in the central and southern counties. We will use SaTScan (version 9.4) to look for clusters of significant HD incidents, significant in both geography and in time. The clustering analysis we will employ have potential applications for improving surveillance programs and designing early warning systems for effective deer population management and potentially reducing the number of HD cases.

The information and results in this chapter have been submitted to BioMed Central (BMC) Ecology for publication and is currently undergoing review and revisions.

Methods

The MDC provided data on deer population and suspected HD occurrence. Estimated instances of HD, by county, were available for the years 1980, 1988, 2005 2014. Estimates of deer population were available for all years except 1980 and 1988. A summary of the data for the years 2005 - 2014 is in Appendix A. In order to apply Kulldorff's spatial and space-time scan statistics to the data, we used SaTScan version 9.4.2 [29] over the 33-year study period (though primarily concentrating on the 2005 - 2014). The geographic center (centroid) of each county was used to represent the location of the presence of (or absence of) HD in the county.

Spatial and temporal patterns of HD have been described in the southeast United States by using the space-time K function and Martin Kuldorffs scan statistic [9, 31, 51, 22]. Significant clusters were most evident in Alabama, North Carolina, and South Carolina between 1980 and 2013. Other studies have applied Kulldorffs space and space-time scan statistic to several geographical regions affected by various disease outbreaks [41, 59, 32, 14, 36]. Over 43% of all 2012 reported EHD cases in captive white-tailed deer belonged to the State of Missouri (see Table 3 of [54]), and, in a previous study, Beringer [6] noted that the HD exposure rate could be as high as 24% within Missouris white-tailed deer population. Moreover, there have been four major HD outbreaks in Missouris white-tailed deer population in years 1988, 2007 and 2012. Therefore, there is a need to further investigate the HD dynamics in Missouri. The most severe outbreak was in the year 2012 when every county in Missouri reported at least one case of HD with more than 10,000 cases of mortality.

Kulldorffs space and space-time scan statistics [28, 30] use a theoretical cylin-

drical window with a circular (or elliptical) base. The base is geographic and, in turn, is centered on each of several possible grid points throughout the area of study. For each grid point, the radius of the window varies continuously in size from zero to a user-specified upper limit based on distance and/or percentage of population. The height of the cylinder corresponds to a period of time within the study period. These cylindrical windows vary in space and/or in time. Thus, for each possible geographic location, it considers multiple-sized circles around the location and multiple possible time frames. For each location and scanning window, the program computes a likelihood ratio based on the number of observed cases versus the number of expected cases both inside and outside the window, using different probability models depending on the data. This expected value is determined by a user-defined number of replications of the data. The number of incidents remains the same, but their distribution in the region is random. The program determines the significance of a cluster based on the actual number of incidents in each window in comparison to the expected number of incidents based on all the replications. With the discrete Poisson model, the program and analysis assumes that the number of cases at each location follows a Poisson distribution and that the expected number of cases in each location is proportional to its population size. The space-time permutation model requires only case data and the number of observed cases in a cluster is compared to what would have been expected if all cases were independent of each other in both space and time as if there were no space-time interaction. Under the null hypothesis of no significant clusters in the window, the window with the largest likelihood statistic is the most likely cluster. The program also identifies all secondary clusters with a P-value less than 0.05. I used

three different scans within SaTScan version 9.4. First, for the spatial scan statistic, I used the annual data to locate clusters in each year and to observe how these clusters changed across years. Second, the space-time scan statistic was used. The space-time permutation model is ideal because it requires only case data, with information about the spatial location and time for each case. Moreover, it has the potential of identifying clusters that may not have been significant for any one specific year but are over spans of multiple years. Third, the spatial scan with temporal trends using the Poisson model was applied to all cases over the study period to locate clusters with more significant variations in the percentage change in the number of cases per year. As part of the scan analysis, I chose elliptical scanning windows. I set the maximum spatial window to 50 percent of the total population, the maximum temporal window (when needed) to 50 percent of the study period with a one-calendar year time aggregation, and the number of random Monte Carlo replications to 4,999. For the years when population data was not available, SaTScan estimated the population through linear interpolation. No additional information about controls or background population at risk is necessary.

Results

Data Analysis

There were 16,853 cases of suspected HD reports over all counties during the study period. If we count the number of times each county reported at least one case, there were 406 times a county reported at least one case (out of 912 potential reporting times). During all years represented, 2012 had the largest number of cases (10,177) with all counties reporting at least one case and the estimated prevalence of

6.9 deer per thousand. Table 1 provides a summary of deer population, HD incidents, the number of counties affected and prevalence per thousand.

Table 1. A summary of the estimated deer population in Missouri, the number of counties (out of 114) reporting suspected HD cases, the number of suspected HD incidents, and the prevalence of suspected HD in thousands.

Year	1980	1988	2005	2006	2007	2008
Harvest	53,298	149,064	286,027	321,828	298,360	280,054
Population[1]	NA	NA	1,550,106	1,575,757	1,494,703	1,508,662
Counties	41	71	21	13	91	0
Incidents[2]	315	1410	772	484	3095	0
Prevalence	NA	NA	0.517	0.321	2.179	0
Year	2009	2010	2011	2012	2013	2014
Harvest	294,346	272,774	287,361	307,979	250,135	255,529
Population[1]	1,539,097	1,548,465	1,599,525	1,406,514	1,245,359	1,177,235
Counties	0	2	22	114	43	76
Incidents[2]	0	150	197	10,117	2696	612
Prevalence	0	0.105	0.123	6.899	2.165	0.520

Notes: [1] *Estimated deer population, and therefore the prevalence, was not available for the years 1980 and 1988.* [2] *The estimated prevalence is per thousand.*

Spatial Clusters By Individual Years

Table 2 provides the locations of the most significant cluster in each year that HD data was available. In Figure 2, counties are shaded based on the number of years in which SaTScan identified them as part of any cluster during the entire study period. The darker the shading, the more frequently it was identified. We observe that SaTScan identified clusters in central to southwestern Missouri more frequently. Figure 2 shows primary and secondary clusters over the study period. Although there is a gap between 1988 and 2005 data, we can see that the outbreaks have occurred in cycles of six to eight years.

Spatiotemporal Clusters

Four significant spatiotemporal clusters were detected, where the primary cluster consists of 32 counties in the eastern and southeastern portions of Missouri. Figure 2 shows the locations of the significant spatio-temporal primary and secondary clusters. The three secondary clusters were located in the southwest (cluster 2), a small portion in the northeast (cluster 3), and a small cluster in the center of the state (cluster 4). See Table 3 for a summary of the significant clusters and the number of counties affected.

Temporal Trends In Spatial Clusters

Using the Poisson model, a trend of 19% annual increase was detected over the study period. There were no instances where a cluster had a significant annual decrease, and Figure 2 shows where the annual increase was the most significant. The primary cluster is the northernmost third of Missouri. In the cases where a secondary

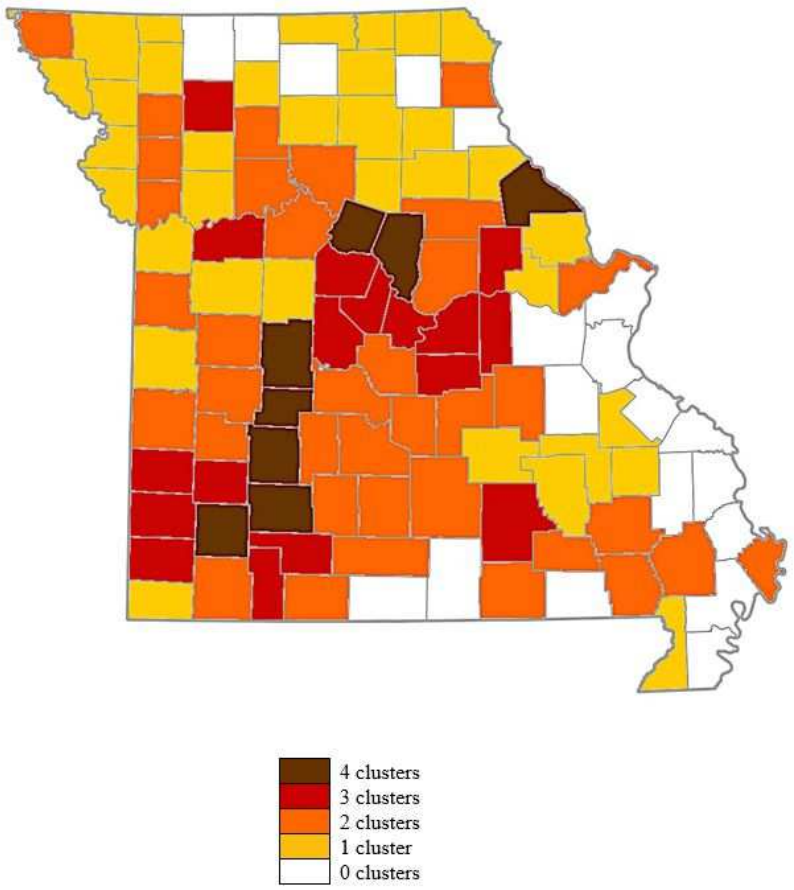


Figure 1. Frequency (in number of years) of HD cluster occurrence for each county during the study period. The darker the shading, the more frequently the county was identified in a cluster.

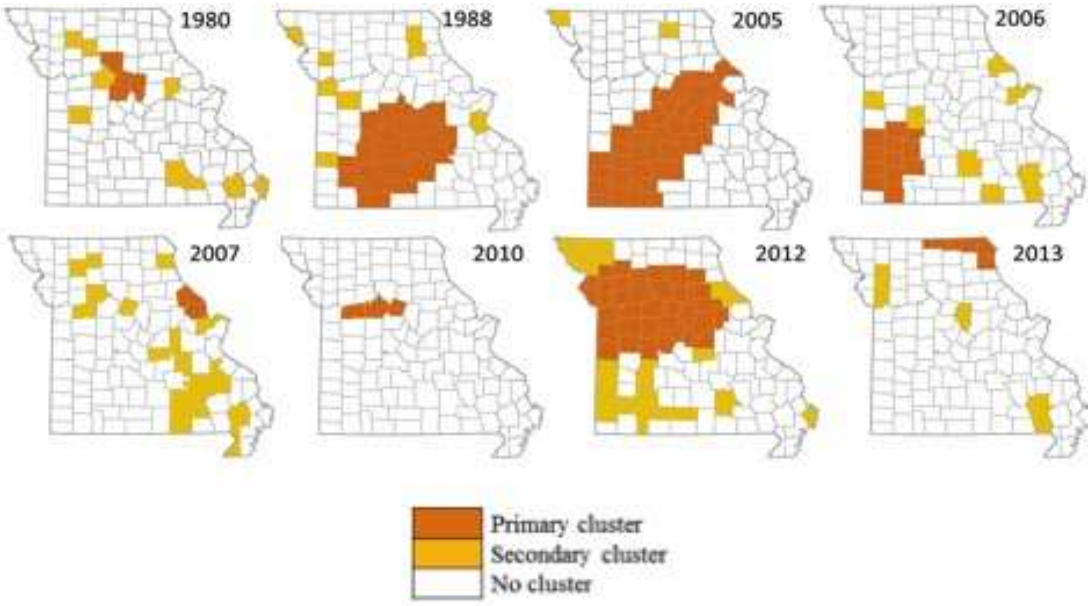


Figure 2. Spatial cluster of years 1980-2013 suggests presence of 6-8 years cycles of HD outbreaks in Missouri. An HD outbreak is anticipated for during 2018-2020.

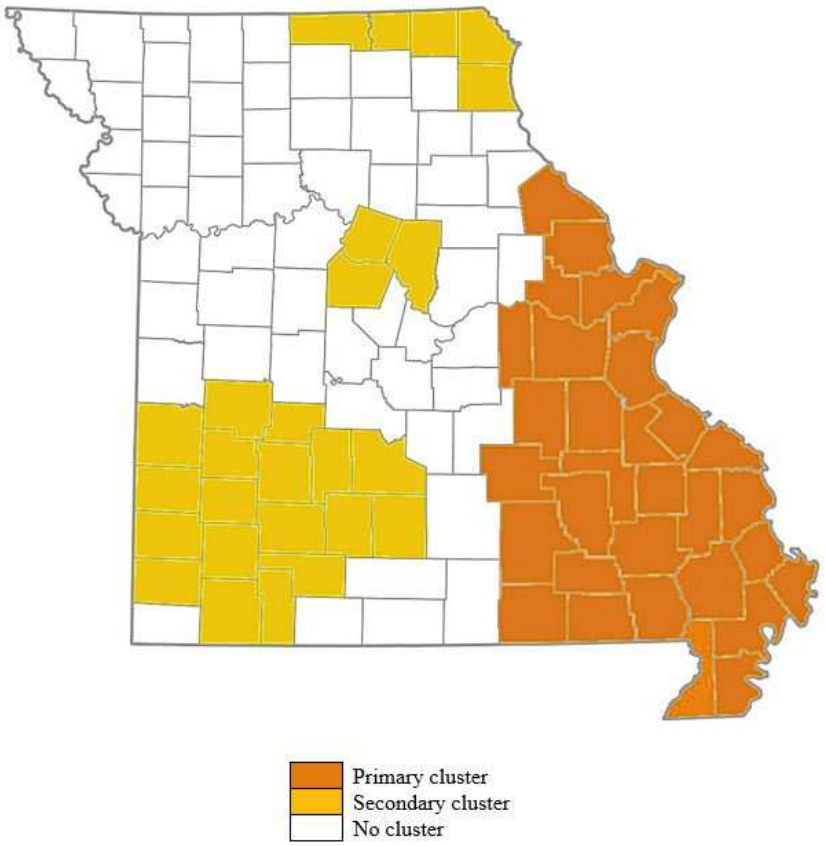


Figure 3. Significant spatiotemporal cluster of HD in white-tailed deer. Primary and secondary clusters of HD presence are displayed as orange and yellow, respectively.

Table 2. The most significant spatial clusters of HD (by individual year) in white-tailed deer during the period 1980 - 2013 with a maximum spatial window = 50% of the total population.

Year	Location(s)*	Observed	Expected
1980	Central (Howard)	189	15.52
1988	Central, South (Laclede)	762	337.18
2005	Central, Southwest (Dallas)	711	263.86
2006	Southwest (Dade)	246	42.64
2007	East (Lincoln)	843	91.31
2010	Central (Saline)	150	3.11
2012	Central, West (Lafayette)	5358	3223.70
2013	Northeast (Clark)	270	20.94

** Denotes county of approximate center of each cluster. Observed: the number of incidents in the most significant cluster. Expected: the expected number of incidents in the most significant cluster based on the random replications. The significance value in each case was* $P < 0.0001$*. The number of estimated HD cases was available only for the years presented here.*

cluster overlaps the primary cluster, the counties in the overlap are grouped within the primary cluster. Table 4 gives the proportion of cases in each cluster and its trend of annual increase. The highest trend of annual increase belongs to Howell County in southern Missouri. However, the five counties (Audrain, Callaway, Osage, Maries, and Phelps) in central Missouri have the highest number of annual cases (57.6 per 100,000).

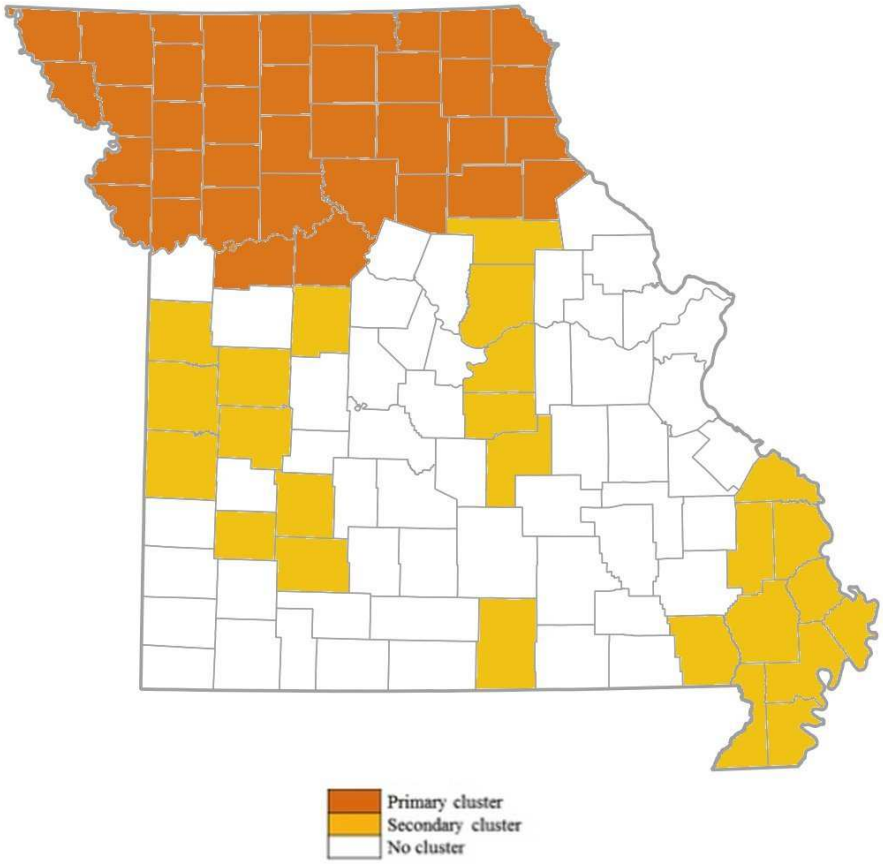


Figure 4. Significant temporal trends (annual increases) of HD in white-tailed deer. Primary and secondary clusters of HD presence are displayed as orange and yellow, respectively. In the cases where a secondary cluster overlaps the primary cluster, the counties in the overlap are grouped within the primary cluster.

Table 3. Significant spatiotemporal clusters of HD in white-tailed deer during 1980 - 2013 with maximum spatial window = 50% of the total number of HD cases and maximum temporal window = 50% of the entire study period.

	Cluster	Counties	Observed	Expected	Period
Primary	1	32	1993	785.95	2006 - 2007
Secondary	2	18	657	168.66	2005 - 2006
	3	5	270	20.69	2013
	4	3	160	11.45	1980

Cluster: cluster ID. Counties: the number of counties in each cluster. Observed: the number of incidents in each cluster. Expected: the expected number of incidents in the cluster based on the random replications. The significance value in each case was $P < 0.0001$.

Discussion And Conclusions

In summary, using the statistical models and the available data, we identified the significant spatial and the spatiotemporal clusters of HD in white-tailed deer population residing in Missouri. The significant temporal trends and the spatiotemporal clusters of HD were identified in northern and southeastern counties of Missouri, respectively (See Figures 1 and 2). These trends and clusters are in agreement with the density of captive white-tailed deer EHD cases during the most severe outbreak in 2012 (see Figure 3 of [54]). However, as shown in Figure 3, the frequencies of significant spatial clusters are mainly located in the central and southwestern counties. Thus, there is a greater likelihood of outbreaks in the central and southwestern counties. Moreover, the spatial clusters shown in Figure 2 suggest that there might

Table 4. Significant temporal trends of HD in white-tailed deer during 1980 - 2013 with maximum spatial window = 50% of the total number of HD cases and maximum temporal window = 50% of the entire study period.

	Cluster	Counties	Location	Annual cases (per 100,000)	Trend of annual increase	P-value
Primary	1	37	North	33.1	32.1%	0.0002
Secondary	2	5	Central	57.6	30.9%	0.0002
	3	4	West	46.6	31.1%	0.0002
	4	10	Southeast	15.4	33.2%	0.0002
	5	3	West	21.0	31.3%	0.0002
	6	2	Southwest	21.6	38.5%	0.0006
	7	1	South	2.7	102.1%	0.0018

Cluster: cluster ID. Counties: the number of counties in each cluster.

be patterns of HD outbreaks. Xu et al. [63] identified similar cycles of six to eight years in an independent study of HD outbreak in the southeastern USA. Therefore, we speculate that there will be an HD outbreak in Missouris white-tailed deer population between the years 2018 - 2020.

It is important to note that HD occurs seasonally and nearly all reported cases occur during late summer and fall. This seasonal occurrence could be related to high abundance of *Culicoides* biting midges during late summer and fall as they transmit the disease. In particular, it is likely that HD outbreaks are more prevalent when weather conditions during the late summer and fall cause an abundance of muddy areas where midges breed. This could be due to high summer temperatures that cause bodies of water to recede and leave mud flats or by overly rainy and wet conditions

in late spring. Those very rare HD cases that are in late fall and winter represent the chronic form of HD.

As outlined below, this study carries a number of limitations related to the data. In general, data availability in wildlife is often an issue. Populations are not enclosed nor controlled, and getting accurate population counts is impossible. Counting the number of HD occurrences depends on observations of harvested deer. Variations in deer population density, regulations on who may harvest the deer, regulation on how many deer may be harvested, and other factors affect this count. Indirect reports from the public may not be verifiable, and some regions may be restricted to hunters and the public at large. So, in actuality, these reports are only estimations and suspected reports. Furthermore, in years when there is not a significant known outbreak, results were reported to the MDC in January of the following year (if at all), and because of this time lag, there is some concern over the accuracy of the reports. Regardless, information of the spatiotemporal clustering may improve or design local surveillance and early warning systems [52, 37]. In particular, areas with spatial and spatiotemporal HD clusters can be targets of more frequent surveillance. These programs can serve as a sentinel to reduce number of HD cases in local farms and to sustain free-living deer population.

Currently there are no effective wildlife management tools or strategies to control or prevent the hemorrhagic diseases in wildlife [11]. However, fencing off livestock and captive white-tailed deer from ponds can reduce the probability of encountering midges. Thus, conservationists and wildlife managers may be able to use the outcomes of the clustering analysis to establish an early warning system to reduce

the number of HD cases in livestock and captive white-tailed deer. An early warning system is also necessary for correct management of the free-living deer population. In particular, an early detection of HD outbreak can critically help the MDC officials to reduce the number of hunting permits in order to sustain the deer population in subsequent seasons. The outcomes of the clustering analysis provided in this study reveals the significant magnitudes and directions of the HD spread in Missouri in the past three decades. In conclusion, cluster analysis can improve our understanding of the epidemiology of hemorrhagic diseases and it can lead to designing effective surveillance and early warning programs.

CHAPTER 3

THE HD MODEL WITH MIGRATION EFFECTS

Overview

In this chapter, we propose a vector-borne disease model which takes into account migrating effects of deer population using distributed delay terms. The model is employed to analyze the effects of deer migration and HD spread. This is carried out in three steps. First, we linearize the model about the disease free equilibrium and stability conditions are derived. Second, using the method of the Next Generation Matrix, we derive the basic reproduction expression R_0 from the model. Third, using the R_0 expression and its numerical simulations, we illustrate that the severity of an HD outbreak is directly influenced by the migration rates of infected and susceptible deer. Using the method of chain trick, we reduce the proposed model with distributed delay to a system of ordinary differential equations where we numerically explore the convergence of the system to endemic and diseases free equilibriums.

The information and results in this chapter have been published in the Journal of Applied Mathematics and Physics[3].

Background

In this work, we build a mathematical model to investigate the dynamics of HD. The amount of literature dedicated to the mathematical modeling of vector-borne diseases is extensive (See for example [46, 34, 33, 58]). The model by Nobel Prize winner Ronald Ross [46] is at the cornerstone of such models, and he used his model

to investigate the spread of malaria. Over four decades later, George Macdonald developed it further [34]. In fact, there have been several extensions to the Ross-Macdonald model. For instance, Lou and Zhou [33] included advection and diffusion terms to take the spatial movements of individuals into account. Reaction-diffusion models have also been used for investigating dynamics of vector-borne diseases such as dengue fever [58] and Zika [10]. Using a deterministic modeling approach, the main objective of this chapter is to have a better understanding of the possible effects of deer-midge interactions and deer migrations on HD dynamics in a deer population.

In recent years, more realistic models have been constructed which take into account dispersion time and host movements. A key article is the work by Neubert et al. [39], which argues that dispersion in Lotka-Volterra predator-prey models is unrealistic as individuals leaving an area (i.e., a patch) immediately appear in another. In nature, an individual requires a finite amount of time to complete a trip from one patch to another or to complete a round trip leaving and returning to the same patch. During this time, the migrating individuals are not interacting with other predators or prey in this patch. Thus, Neubert and his co-authors [39, 25] demonstrate that models that incorporate explicit travel-time are often more stable.

The Migration Model

In the attempt to create a mathematical model of HD outbreak in a population of white-tailed deer, we make certain assumptions based on the ecology of deer and midge populations and the characteristics of HD. The deer (host) and midge (vector) populations are divided into susceptible and infected classes. At time t, there are $D_S(t)$ susceptible deer, $D_I(t)$ infected deer, $M_S(t)$ susceptible midges, and $M_I(t)$ in-

fected midges. The total deer population at time t is $D_N(t) = D_S(t) + D_I(t)$, and the total midge population is $M_N(t) = M_S(t) + M_I(t)$. Susceptible deer become infected through bites of infected midges; susceptible midges become infected when they feed on the blood of an infected deer. As observed in the wild, deer will migrate (disperse) out of and back into a region (i.e., a patch) due to seasonal variations, availability of food, or predators; midges, however, will not. They are weak fliers and typically disperse no more than about a mile from the site of larval development, with females flying farther than males [45]. We therefore consider the following assumptions in the model construction:

1. All newborns are susceptible in both populations of deer and midges (i.e., no inherited infection or vertical transmission is considered).
2. Susceptible deer become infected only by adequate contact with infected midges and cannot become infected via contact with an infected deer.
3. Once infected, a deer will die from the disease. (Note, in actuality, there are cases where a deer survives the infection, but it is rare.)
4. Individuals in both populations will die naturally by both density independent and density dependent factors.
5. By the law of mass action, we assume that infection transmission is proportional to the population densities of deer and midges.
6. Deer will frequently travel out of and into a geographic area (a patch), but midges do not (as the amount of dispersal in midge populations is negligible).

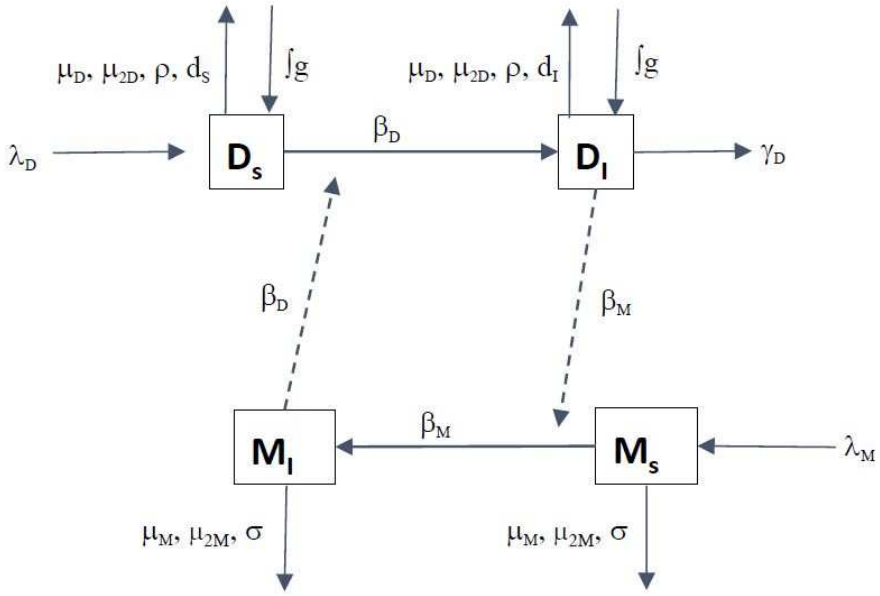


Figure 5. A compartmental diagram of the HD model (3.1) with population dispersal. Dashed lines represent the HD transmission between the vector and host. Deer migration into the patch is denoted by $\int g$ and migration out of the patch is denoted by d_S and d_I. See Table 5 for a summary of the parameters and variables.

Table 5. Summary of the variables and parameters used in the delayed HD model (3.1)

Symbol	Description
$D_S(t)$	Number of susceptible deer at time t
$D_I(t)$	Number of infected deer at time t
$D_N(t)$	Total deer population at time t
β_D	Infection rate (deer)
λ_D	Birth rate (deer)
ρ	Harvest rate
μ_D	Death rate (deer), density independent
μ_{2D}	Death rate (deer), density dependent
d_S	Net flux rate, susceptible deer
d_I	Net flux rate, infected deer
γ_D	Pathogenic induced death rate (deer)
δ_S	Probability of death per unit of time of a susceptible deer during migration
δ_I	Probability of death per unit of time of an infected deer during migration
$M_S(t)$	Number of susceptible midges at time t
$M_I(t)$	Number of infected midges at time t
$M_N(t)$	Total midge population at time t
β_M	Infection rate (midges)
λ_M	Birth rate (midges)
σ	Efficacy rate of midge control measures
μ_M	Death rate (midges), density independent
μ_{2M}	Death rate (midges), density dependent

A compartmental diagram of the proposed HD model is seen in Figure 3, and a summary of parameters and variables is given in Table 5. All parameters are assumed to be non-negative. Given the above-mentioned assumptions and the model diagram, the set of delayed differential equations representing the model is given by

$$\begin{aligned}
\frac{dD_S}{dt} &= \lambda_D D_N - \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + d_S + \mu_{2D} D_N) D_S + \\
&\quad d_S \int_0^\infty g(z) e^{-\delta_S z} D_S(t-z) dz \\
\frac{dD_I}{dt} &= \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + \gamma_D + d_I + \mu_{2D} D_N) D_I + \\
&\quad d_I \int_0^\infty g(z) e^{-\delta_I z} D_I(t-z) dz \\
\frac{dM_S}{dt} &= \lambda_M M_N - \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_S \\
\frac{dM_I}{dt} &= \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_I.
\end{aligned} \tag{3.1}$$

In absence of the disease, population growths of deer and midges are formulated with logistic growth models. These are the terms that include λ_D, λ_M, μ_{2D}, and μ_{2M} in model (3.1). Similar to [40], the carrying capacity for the deer population exists and must be positive. Hence, it is required that

$$H_1 : \lambda_D > \mu_D + \rho + d_S \tag{3.2}$$

and

$$H_2 : \lambda_D > \mu_D + \rho + \gamma_D + d_I. \tag{3.3}$$

Also, the carrying capacity for midges exists and is positive. Thus,

$$H_3 : \lambda_M > \mu_M + \sigma. \tag{3.4}$$

Individual deer immigrate from the patch at a constant per capita rate (d_S and d_I) and return z units of time after their departure. The integrals in the first two equations of model (3.1) are distributed delay terms representing the influx of susceptible and infected deer, respectively, from all points in time in the past up to and including the present time [27]. The function $g(z)$ in the integrals is a probability density function for the time it takes for a deer to disperse given that the deer survives the trip, and $g(z)dz$ is the probability that a successfully dispersing deer departing at time t completes the trip between time $t + z$ and $t + z + dz$. As $g(z)$ is a probability density function, it is normalized so that $\int_0^\infty g(z)dz = 1$. The functions $e^{-\delta_S z}$ and $e^{-\delta_I z}$ in the integrals are the probabilities of a deer surviving a trip of duration z given constant probabilities per unit of time δ_S and δ_I for the mortality during travel of susceptible and infected deer, respectively. All deer migrating back into this single patch originated in the patch; in other words, there are no new deer entering the patching that originated from somewhere else. Hence, we are studying a herd of deer concentrated within a patch with the ability of migrating in and out of it.

Linear Stability Analysis

In this section, we provide a formal procedure of linear stability analysis which leads to the characteristic equation and the stability conditions for the equilibrium solutions. Specifically, Disease Free Equilibrium (DFE) (i.e., $D_I^* = 0$ and $M_I^* =$

0) and Endemic Equilibrium (EE) are the constant solutions of model (3.1). In epidemiology, a stable DFE is always desired whereas a stable EE can be of great concern. The first two equations of model (3.1) have an integral influx term that may be simplified by the following method. Letting

$$f^{[1]}(D_S, D_I, M_S, M_I) = \lambda_D D_N - \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + d_S + \mu_{2D} D_N) D_S, \tag{3.5}$$

we rewrite the first equation as

$$\frac{dD_S}{dt} = f^{[1]}(D_S, D_I, M_S, M_I) + d_S \int_0^\infty g(z) e^{-\delta_S z} D_S(t-z) dz. \tag{3.6}$$

Similarly, we rewrite the second equation as

$$\frac{dD_I}{dt} = f^{[2]}(D_S, D_I, M_S, M_I) + d_I \int_0^\infty g(z) e^{-\delta_I z} D_I(t-z) dz, \tag{3.7}$$

where

$$f^{[2]}(D_S, D_I, M_S, M_I) = \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + \gamma_D + d_I + \mu_{2D} D_N) D_I. \tag{3.8}$$

As the bottom two equations of model (3.1) have no integral term, we let $f^{[3]}$ and $f^{[4]}$ equal the right-hand side of the third and fourth equations in model (3.1), respectively. Let a solution $\big(D_S(t), D_I(t), M_S(t), M_I(t)\big)$ of model (3.1) nearby an equilibrium solution $E = (D_S^*, D_I^*, M_S^*, M_I^*)$ be in the form of

$$D_S(t) = D_S^* + \widetilde{D}_S(t), D_I(t) = D_I^* + \widetilde{D}_I(t), M_S(t) = M_S^* + \widetilde{M_S}(t), M_I(t) = M_I^* + \widetilde{M_I}(t) \tag{3.9}$$

for some $\widetilde{D}_S(t)$, $\widetilde{D}_I(t)$, $\widetilde{M}_S(t)$, and $\widetilde{M}_I(t)$. Using the Taylor expansion, we linearize the first equation in model (3.1) about equilibrium E by substituting (3.9) into (3.6) and dropping the nonlinear terms. Thus, the first equation of model (3.1) is linearized as follows.

$$\begin{aligned}
\frac{dD_S}{dt} &= \frac{d\widetilde{D}_S}{dt} \\
&= f^{[1]}(D_S, D_I, M_S, M_I) + d_S \int_0^\infty g(z)e^{-\delta_S z} D_S(t-z)dz \\
&= f^{[1]}(D_S^*, D_I^*, M_S^*, M_I^*) + \frac{\partial f^{[1]}}{\partial D_S}(E)\cdot \widetilde{D}_S + \frac{\partial f^{[1]}}{\partial D_I}(E)\cdot \widetilde{D}_I + \frac{\partial f^{[1]}}{\partial M_S}(E)\cdot \widetilde{M}_S + \\
&\quad \frac{\partial f^{[1]}}{\partial M_I}(E)\cdot \widetilde{M}_I + d_S \int_0^\infty g(z)e^{-\delta_S z}\big(D_S^* + \widetilde{D}_S(t-z)\big)dz \\
&= f^{[1]}(D_S^*, D_I^*, M_S^*, M_I^*) + \frac{\partial f^{[1]}}{\partial D_S}(E)\cdot \widetilde{D}_S + \frac{\partial f^{[1]}}{\partial D_I}(E)\cdot \widetilde{D}_I + \frac{\partial f^{[1]}}{\partial M_S}(E)\cdot \widetilde{M}_S + \\
&\quad \frac{\partial f^{[1]}}{\partial M_I}(E)\cdot \widetilde{M}_I + d_S D_S^* \int_0^\infty g(z)e^{-\delta_S z}dz + d_S \int_0^\infty g(z)e^{-\delta_S z}\widetilde{D}_S(t-z)dz.
\end{aligned} \tag{3.10}$$

We know that equilibrium E satisfies the first equation of model (3.1), hence

$$f^{[1]}(D_S^*, D_I^*, M_S^*, M_I^*) + d_S D_S^* \int_0^\infty g(z)e^{-\delta_S z}dz = 0, \tag{3.11}$$

and thus

$$d_S D_S^* \int_0^\infty g(z)e^{-\delta_S z}dz = -f^{[1]}(D_S^*, D_I^*, M_S^*, M_I^*). \tag{3.12}$$

Substituting (3.12) into (3.10) yields

$$\begin{aligned}\frac{d\widetilde{D}_S}{dt} &= \frac{\partial f^{[1]}}{\partial D_S}(E)\cdot\widetilde{D}_S + \frac{\partial f^{[1]}}{\partial D_I}(E)\cdot\widetilde{D}_I + \frac{\partial f^{[1]}}{\partial M_S}(E)\cdot\widetilde{M}_S + \frac{\partial f^{[1]}}{\partial M_I}(E)\cdot\widetilde{M}_I \\ &\quad + d_S\int_0^\infty g(z)e^{-\delta_S z}\widetilde{D}_S(t-z)dz.\end{aligned} \tag{3.13}$$

Applying the same procedure to equation (3.7), we get that the second equation of model (3.1) is linearized by

$$\begin{aligned}\frac{d\widetilde{D}_I}{dt} &= \frac{\partial f^{[2]}}{\partial D_S}(E)\cdot\widetilde{D}_S + \frac{\partial f^{[2]}}{\partial D_I}(E)\cdot\widetilde{D}_I + \frac{\partial f^{[2]}}{\partial M_S}(E)\cdot\widetilde{M}_S + \frac{\partial f^{[2]}}{\partial M_I}(E)\cdot\widetilde{M}_I \\ &\quad + d_I\int_0^\infty g(z)e^{-\delta_I z}\tilde{D}_I(t-z)dz.\end{aligned} \tag{3.14}$$

Using equations (3.9) - (3.14), model (3.1) is linearized about equilibrium E and takes the form

$$Y'(t) = AY(t), \tag{3.15}$$

where $Y(t) = [\widetilde{D}_S(t), \widetilde{D}_I(t), \widetilde{M}_S(t), \widetilde{M}_I(t)]^T$ and A is the Jacobian matrix evaluated at E. However, the specific form of matrix A cannot be extracted due to the presence of the integral terms in (3.13) and (3.14). To bypass this issue, we use the Fundamental Theorem of linear systems of differential equations [44] and look for exponential solutions of the form

$$\begin{bmatrix}\widetilde{D}_S(t)\\ \widetilde{D}_I(t)\\ \widetilde{M}_S(t)\\ \widetilde{M}_I(t)\end{bmatrix} = \begin{bmatrix}r_1\\ r_2\\ r_3\\ r_4\end{bmatrix} e^{\lambda t} = Re^{\lambda t}. \tag{3.16}$$

I also let $\tilde{g}$ be the (one-sided) Laplace transform of the travel-time distribution $g(z)$.

That is,

$$\tilde{g}(x) \equiv \int_0^\infty g(z)e^{-xz}dz. \tag{3.17}$$

We have the following Lemma.

LEMMA 3.1. *The Laplace transform $\tilde{g}$ is a positive, decreasing function that is bounded above by 1 for all non-negative values of x.*

PROOF. Let $g(z)$ be a probability density function as described above. Because the function e^{-xz} is positive for all real x and fixed z, $e^{-xz} = 1$ when $x = 0$, and e^{-xz} decreases for all $x > 0$. Therefore, it must be the case that $0 < g(z)e^{-xz} \leq 1$ and $g(z)e^{-xz}$ decreases for all non-negative x. Thus, $\tilde{g}(x) \equiv \int_0^\infty g(z)e^{-xz}dz$ is a positive decreasing function bounded above by 1. □

By substituting (3.16) into (3.15) and simplifying the terms, we get the specific form of matrix A, and 3.15 is rewritten as

$$\begin{bmatrix} A_1 & \frac{\partial f^{[1]}(E)}{\partial D_I} & \frac{\partial f^{[1]}(E)}{\partial M_S} & \frac{\partial f^{[1]}(E)}{\partial M_I} \\ \frac{\partial f^{[2]}(E)}{\partial D_S} & A_2 & \frac{\partial f^{[2]}(E)}{\partial M_S} & \frac{\partial f^{[2]}(E)}{\partial M_I} \\ \frac{\partial f^{[3]}(E)}{\partial D_S} & \frac{\partial f^{[3]}(E)}{\partial D_I} & A_3 & \frac{\partial f^{[3]}(E)}{\partial M_I} \\ \frac{\partial f^{[4]}(E)}{\partial D_S} & \frac{\partial f^{[4]}(E)}{\partial D_I} & \frac{\partial f^{[4]}(E)}{\partial M_S} & A_4 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{3.18}$$

where $A_1 = \frac{\partial f^{[1]}(E)}{\partial D_S} + d_S\tilde{g}(\lambda+\delta_S) - \lambda$, $A_2 = \frac{\partial f^{[2]}(E)}{\partial D_I} + d_I\tilde{g}(\lambda+\delta_I) - \lambda$, $A_3 = \frac{\partial f^{[3]}(E)}{\partial M_S} - \lambda$, and $A_4 = \frac{\partial f^{[4]}(E)}{\partial M_I} - \lambda$.

The linear system in (3.18) has a nontrivial solution if and only if the deter-

minant of the matrix is zero. This leads to the characteristic equation corresponding to model (3.1) linearized about E. Before deriving the characteristic equation, we prove the existence of DFE.

Proposition 1. *The disease free equilibrium of model (3.1) exists if and only if* $\lambda_D > \mu_D + \rho + d_S(1 - \tilde{g}(\delta_S))$ *and* $\lambda_M > \mu_M + \sigma$ *are satisfied.*

PROOF. Noting that $D_I^* = 0$, $D_N = D_S^*$, and $\frac{dD_S}{dt} = 0$ at the DFE, the first equation in model (3.1) gives us $D_S^* = \frac{\lambda_D - (\mu_D + \rho) - d_S(1-\tilde{g}(\delta_S))}{\mu_{2D}}$. Similarly, $M_I^* = 0$ and $M_N = M_S^*$, and the third equation of model (3.1) gives rise to $M_S^* = \frac{\lambda_M - (\mu_M + \sigma)}{\mu_{2M}}$. As $D_S^* > 0$ and $M_S^* > 0$ by parameter assumptions, the disease free equilibrium exists. □

Remark 1. *The inequalities (3.2) and (3.4) and Lemma 3.1 imply that the conditions of Proposition 1 are always satisfied. Hence, the DFE always exists and it is given by*

$$\begin{aligned} D_S^* &= \frac{\lambda_D - (\mu_D + \rho) - d_S(1 - \tilde{g}(\delta_S))}{\mu_{2D}} \\ D_I^* &= 0 \\ M_S^* &= \frac{\lambda_M - (\mu_M + \sigma)}{\mu_{2M}} \\ M_I^* &= 0. \end{aligned} \tag{3.19}$$

By linearizing model (3.1) about the DFE, we get the characteristic equation

$$det\big(J(\lambda)\big) = 0, \tag{3.20}$$

where $J(\lambda)$ is the matrix in (3.18) evaluated at $E = DFE$, and it simplifies to

$$J(\lambda) = \begin{bmatrix} J_1(\lambda) & \lambda_D - \mu_{2D} D_S^* & 0 & -\beta_D \\ 0 & J_2(\lambda) & 0 & \beta_D \\ 0 & -\frac{\beta_M M_S^*}{D_S^*} & J_3(\lambda) & \lambda_M - \mu_{2M} M_S^* \\ 0 & \frac{\beta_M M_S^*}{D_S^*} & 0 & J_4(\lambda) \end{bmatrix} \tag{3.21}$$

such that

$$J_1(\lambda) = \lambda_D - \mu_D - \rho - d_S - 2\mu_{2D} D_S^* + d_S \tilde{g}(\lambda + \delta_S) - \lambda, \tag{3.22}$$

$$J_2(\lambda) = -\mu_D - \rho - \gamma_D - d_I - \mu_{2D} D_S^* + d_I \tilde{g}(\lambda + \delta_I) - \lambda, \tag{3.23}$$

$$J_3(\lambda) = \lambda_M - \mu_M - \sigma - 2\mu_{2M} M_S^* - \lambda, \tag{3.24}$$

and

$$J_4(\lambda) = -\mu_M - \sigma - \mu_{2M} M_S^* - \lambda. \tag{3.25}$$

Hence, the characteristic equation (3.20) is rewritten

$$J_1(\lambda) J_3(\lambda) \left[J_2(\lambda) J_4(\lambda) - \frac{\beta_D \beta_M M_S^*}{D_S^*} \right] = 0. \tag{3.26}$$

Since $J_1(\lambda)$ and $J_2(\lambda)$ are not polynomials, the Routh-Hurwitz criteria [47] is not applicable for determining stability. However, with a specific form of $g(z)$, we may compute the roots of the characteristic equation and determine the necessary and sufficient conditions for the stability of the DFE.

Basic Reproduction Number

The basic reproduction number R_0 is defined as the expected number of secondary infections produced by a single case of an infection introduced to a completely susceptible population [8]. When $R_0 > 1$, the infection will spread as the number of infected individuals increases. When $R_0 < 1$, the infection will die out in the long run. Thus, we seek conditions and parameter values so that $R_0 < 1$.

The magnitude of R_0 determines the severity of infection. Larger values of $R_0 > 1$ lead to faster disease spread, whereas smaller values of $R_0 < 1$ lead to the disease dying out more rapidly. Using the Next Generation Matrix (NGM) approach [2, 56], the expression for R_0 can be derived. Specifically, the next generation matrix is given by $K = FV^{-1}$, and the spectral radius of K is equal to R_0. The elements of matrix F, using the extended definition of the matrix F [21], represent new infections, where the entry (i, j) of F represents the rate at which secondary individuals appear in class i per individual of type j. The elements of matrix V are the transition of infections.

In order to calculate the R_0 expression, we make some simplifying assumptions in our model. In particular, we assume the integral terms in the first and second equations of model (3.1) are simplified to

$$\int_0^\infty g(z)e^{-\delta_S z}D_S(t-z)dz = \tilde{g}(\delta_S)D_S(t) \tag{3.27}$$

and

$$\int_0^\infty g(z)e^{-\delta_I z}D_I(t-z)dz = \tilde{g}(\delta_I)D_I(t) \tag{3.28}$$

respectively.

Remark 2. *The assumptions in (3.27) and (3.28) result in a positive outflow of deer out of the patch. The first equation of model (3.1) contains the expression* $-d_S D_S(t) + d_S \int_0^\infty g(z)e^{-\delta_S z} D_S(t-z)dz$. *Using (3.27), this simplifies to* $d_S(\tilde{g}(\delta_S)-1)D_S(t)$ *which is negative by the above Lemma. In other words, there are more susceptible deer leaving the patch than entering it. The same is true for the infected deer as concluded from the second equation of model (3.1) and assumption (3.28).*

Using the assumptions in (3.27) and (3.28), we get that

$$F = \begin{bmatrix} d_I \tilde{g}(\delta_I) & \beta_D \\ \frac{\beta_M M_S^*}{D_S^*} & 0 \end{bmatrix}, \tag{3.29}$$

$$V = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}, \tag{3.30}$$

and

$$FV^{-1} = \begin{bmatrix} \frac{d_I \tilde{g}(\delta_I)}{V_1} & \frac{\beta_D}{V_2} \\ \frac{\beta_M M_S^*}{V_1 D_S^*} & 0 \end{bmatrix}, \tag{3.31}$$

where

$$V_1 = \mu_D + \rho + \gamma_D + d_I + \mu_{2D} D_S^* \tag{3.32}$$

and

$$V_2 = \mu_M + \sigma + \mu_{2M} M_S^*. \tag{3.33}$$

As mentioned earlier, the basic reproduction number R_0 is the spectral ra-

dius of FV^{-1}. Since FV^{-1} is a positive definite matrix, R_0 is equal to the largest eigenvalue of FV^{-1}. After simplifying, the expression for R_0 can be written as

$$R_0 = \frac{1}{2}\Big(R_0^{[1]} + \sqrt{\left(R_0^{[1]}\right)^2 + 4R_0^{[2]}}\Big), \tag{3.34}$$

where

$$R_0^{[1]} = \frac{d_I \tilde{g}(\delta_I)}{V_1}, \tag{3.35}$$

representing the contribution of deer migration to disease outbreaks, and

$$R_0^{[2]} = \frac{\beta_D \beta_M M_S^*}{V_1 V_2 D_S^*}, \tag{3.36}$$

representing the effects of the deer-midge interactions on disease outbreaks. Therefore, the migration effects of infected deer and the effects of deer-midge interactions within the patch on HD outbreaks can be studied separately.

1. **Pure migration effects** ($R_0^{[2]} = 0$). This occurs when either β_D or β_M is zero, and thus there is no transmission of the disease between the midges and the deer (or vice-versa) within the patch. Using equation (3.34), $R_0^{[2]} = 0$ implies $R_0 = R_0^{[1]}$. In reality, this can effectively occur when the midge population in the patch is negligible. It can be seen that $R_0^{[1]}$ is a concave down increasing function of d_I. Thus, the flux rate of infected deer d_I may increase $R_0^{[1]}$. From equation (3.32), we get that $\lim_{d_I \to \infty} R_0^{[1]} = \tilde{g}(\delta_I)$. Using Lemma 3.1, $\tilde{g}(\delta_I) \leq 1$. Therefore, d_I alone cannot cause an outbreak even though it increases the $R_0^{[1]}$ value. In fact, using equations (3.32) and (3.35), it can be easily shown that $R_0^{[1]} < 1$ for all parameter values of the model. Hence,

assumptions (3.27) and (3.28) are underestimating the migration effects of deer population on disease outbreak.

2. **Residential effects** ($R_0^{[1]} = 0$). This occurs when $d_I = 0$, which means that infected deer have limited mobility and cannot leave or enter the patch due to illness. In this case, $R_0^{[1]} = 0$ implies $R_0 = \sqrt{R_0^{[2]}}$. In this case, an epidemic may be prevented if $R_0^{[2]} < 1$. This, in fact, may be possible as the harvest rate, ρ, is a part of the expression of $R_0^{[2]}$. On the other hand, small values of V_2 (i.e., low mortality of midges) may result in an outbreak.

The following proposition indicates the effects of parameter values on R_0 in general.

Proposition 2. *The basic reproduction number R_0 is defined in equation (3.34) and it has the following properties:*

i . R_0 is an increasing function of δ_S and d_S.

ii . R_0 is a decreasing function of δ_I.

iii . R_0 is an increasing function of d_I if d_S or the product $\beta_D\beta_M$ is sufficiently small.

iv . R_0 is a decreasing function of d_I if d_S or the product $\beta_D\beta_M$ is sufficiently large.

PROOF. Part (i): As shown below, the partial derivative of R_0 with respect to δ_S is positive.

$$\frac{\partial R_0}{\partial \delta_S} = \frac{-\beta_D \beta_M d_S M_S^* \tilde{g}'(\delta_S)}{\mu_{2D} V_1 V_2 (D_S^*)^2 \sqrt{\left(R_0^{[1]}\right)^2 + 4R_0^{[2]}}} > 0. \tag{3.37}$$

Note that $\tilde{g}'(\delta_S) < 0$ because $\tilde{g}(\delta_S)$ is a decreasing function (See Lemma 3.1). Similarly, the partial derivative of R_0 with respect to d_S is positive.

$$\begin{aligned}\frac{\partial R_0}{\partial d_S} &= \frac{d_I \tilde{g}(\delta_I)(1-\tilde{g}(\delta_S))}{2V_1^2} + \\ &\frac{1}{\sqrt{\left(R_0^{[1]}\right)^2 + 4R_0^{[2]}}} \left[\frac{(d_I \tilde{g}(\delta_I))^2 (1-\tilde{g}(\delta_S))}{2V_1^3} + \frac{\beta_D \beta_M M_S^*}{\mu_{2D} V_2 (V_1 D_S^*)^2 (1-\tilde{g}(\delta_S))} \left(\mu_{2D} D_S^* + V_1\right) \right] \\ &> 0.\end{aligned} \tag{3.38}$$

Part (ii): The partial derivative of R_0 with respect to δ_I is negative.

$$\frac{\partial R_0}{\partial \delta_I} = \frac{d_I \tilde{g}'(\delta_I)}{2} \left(\frac{1}{V_1} + \frac{d_I \tilde{g}(\delta_I)}{V_1^2 \sqrt{\left(R_0^{[1]}\right)^2 + 4R_0^{[2]}}} \right) < 0. \tag{3.39}$$

To prove statements (iii) and (iv), note that the partial derivative of R_0 with respect to d_I is given by

$$\frac{\partial R_0}{\partial d_I} = \frac{\tilde{g}(\delta_I)(V_1 - d_I)}{2V_1^2} + \frac{1}{V_1^2 \sqrt{\left(R_0^{[1]}\right)^2 + 4R_0^{[2]}}} \left[\frac{d_I (V_1 - d_I)(\tilde{g}(\delta_I))^2}{2V_1} - \frac{\beta_D \beta_M M_S^*}{V_2 D_S^*} \right]. \tag{3.40}$$

Also note that $V_1 - d_I = \mu_D + \rho + \gamma_D + \mu_{2D} D_S^* > 0$. The expression

$$\frac{d_I (V_1 - d_I)(\tilde{g}(\delta_I))^2}{2V_1} - \frac{\beta_D \beta_M M_S^*}{V_2 D_S^*} > 0 \tag{3.41}$$

is equivalent to

$$d_I V_2 D_S^*(V_1 - d_I)(\tilde{g}(\delta_I))^2 - 2V_1\beta_D\beta_M M_S^* > 0. \tag{3.42}$$

Recall that $D_S^* = \frac{\lambda_D - (\mu_D + \rho) - d_S(1-\tilde{g}(\delta_S))}{\mu_{2D}}$. When d_S is sufficiently small, $d_I V_2 D_S^*(V_1 - d_I)(\tilde{g}(\delta_I))^2$ will be sufficiently large and the inequality holds. When $\beta_D\beta_M$ is sufficiently small, $2V_1\beta_D\beta_M M_S^*$ will be sufficiently small and the inequality holds. Thus $\frac{\partial R_0}{\partial d_I} > 0$. Similarly, when either d_S or the product $\beta_D\beta_M$ is sufficiently large, $\frac{\partial R_0}{\partial d_I} < 0$.

□

Remark 3. *Proposition 2 implies that the flux rate d_I of infected deer can have two opposing effects based on the value of d_S or the product $\beta_D\beta_M$. Because the directional behavior of R_0 changes due to the value of these, there must be critical values ($d_S^{[c]}$ and $(\beta_D\beta_M)^{[c]}$) such that R_0 is an increasing function of d_I when d_S or $\beta_D\beta_M$ are below either of the critical values and R_0 is a decreasing function of d_I when d_S or $\beta_D\beta_M$ are above either of them.*

The following Lemma is associated with the structure of the R_0 expression in equation (3.34).

LEMMA 3.2. *For $a, b \geq 0$, $a + b < 1$ if and only if $\frac{1}{2}\left(a + \sqrt{a^2 + 4b}\right) < 1$.*

PROOF. ($\Longrightarrow$) If $a + b < 1$, then $b < 1 - a$. Also, as $0 \leq a < 1$, $|a - 2| = 2 - a$. Thus,

$$\begin{aligned}
\frac{1}{2}\big(a+\sqrt{a^2+4b}\big) &< \frac{1}{2}\big(a+\sqrt{a^2+4(1-a)}\big)\\
&= \frac{1}{2}\big(a+\sqrt{a^2-4a+4}\big)\\
&= \frac{1}{2}(a+|a-2|)\\
&= \frac{1}{2}(a+2-a)\\
&= 1.
\end{aligned} \tag{3.43}$$

($\Longleftarrow$)

$$\begin{aligned}
\frac{1}{2}\big(a+\sqrt{a^2+4b}\big) &< 1\\
a+\sqrt{a^2+4b} &< 2\\
\sqrt{a^2+4b} &< 2-a\\
a^2+4b &< 4-4a+a^2\\
4a+4b &< 4\\
a+b &< 1.
\end{aligned} \tag{3.44}$$

□

Remark 4. *Let $a = R_0^{[1]}$ and $b = R_0^{[2]}$. Using Lemma 3.2, we get that $R_0 < 1$ is equivalent to $R_0^{[1]} + R_0^{[2]} < 1$. As indicated in [17, 18], the expression $R_0^{[1]} + R_0^{[2]}$ is known as a Type-Reduction number which can be more accurate than R_0 to calculate the minimum disease eradication efforts.*

Proposition 3. *Under the assumptions (3.27) and (3.28), the DFE of model (3.1) is locally asymptotically stable if and only if $R_0^{[1]} + R_0^{[2]} < 1$ or, equivalently, $R_0 < 1$.*

PROOF. ($\Longleftarrow$) We determine stability conditions at the DFE by using the Jacobian of the system of equations. The DFE is locally asymptotically stable if the real parts of all eigenvalues of the Jacobian matrix are negative as explained in Section 3.1. Using assumptions (3.27) and (3.28), the Jacobian matrix evaluated at the DFE is given by:

$$A = \begin{bmatrix} A_1 & \lambda_D - \mu_{2D} D_S^* & 0 & -\beta_D \\ 0 & A_2 & 0 & \beta_D \\ 0 & -\frac{\beta_M M_S^*}{D_S^*} & A_3 & \lambda_M - \mu_{2M} M_S^* \\ 0 & \frac{\beta_M M_S^*}{D_S^*} & 0 & A_4 \end{bmatrix}, \tag{3.45}$$

where $A_1 = \lambda_D - \mu_D - \rho - (d_S(1 - \tilde{g}(\delta_S)) + 2\mu_{2D} D_S^*)$, $A_2 = d_I \tilde{g}(\delta_I) - V_1$, $A_3 = \lambda_M - \mu_M - \sigma - s\mu_{2M} M_S^*$, and $A_4 = -V_2$. The characteristic equation of this matrix, using Λ for the eigenvalues, is

$$f(\Lambda) = (A_1 - \Lambda)(A_3 - \Lambda)\left[(A_2 - \Lambda)(A_4 - \Lambda) - \frac{\beta_D \beta_M M_S^*}{D_S^*}\right]. \tag{3.46}$$

For the first eigenvalue A_1, we note that since the DFE must satisfy $D_S' = 0$, we can show that $\lambda_D = \mu_D + \rho + d_S + \mu_{2D} D_S^* - d_S \tilde{g}(\delta_S) = V_1 - d_S \tilde{g}(\delta_S)$. Therefore,

$$\begin{aligned} A_1 &= \lambda_D - \mu_D - \rho - (d_S(1 - \tilde{g}(\delta_S)) + 2\mu_{2D} D_S^*) \\ &= (\mu_D + \rho + d_S + \mu_{2D} D_S^* - d_S \tilde{g}(\delta_S)) - \mu_D - \rho - (d_S(1 - \tilde{g}(\delta_S)) + 2\mu_{2D} D_S^*) \\ &= -\mu_{2D} D_S^* \\ &< 0. \end{aligned} \tag{3.47}$$

Similarly, for the second eigenvalue, given that the DFE must satisfy $M_S' = 0$, we can

show $\lambda_M = \mu_M + \sigma + \mu_{2M} M_S^*$, and thus $A_3 = \lambda_M - \mu_M - \sigma - 2\mu_{2M} M_S^* = -\mu_{2M} M_S^* < 0$.

For the remaining two eigenvalues, we rewrite the part of the characteristic equation in brackets as

$$\Lambda^2 - (A_2 + A_4)\Lambda + A_2 A_4 - \frac{\beta_D \beta_M M_S^*}{D_S^*} = 0. \tag{3.48}$$

This is a quadratic of the form $\Lambda^2 + b\Lambda + c$. According to the Routh-Hurwitz criteria [47], the roots of a quadratic will have negative real parts if the linear coefficient and the constant term are positive. The linear coefficient is $-(A_2 + A_4)$ and is positive as shown below.

$$\begin{aligned} A_2 + A_4 &= d_I \tilde{g}(\delta_I) - V_1 - V_2 \\ &= d_I \tilde{g}(\delta_I) - \mu_D - \rho - \gamma_D - d_I - \mu_{2D} D_S^* - V_2 \\ &= -d_I(1 - \tilde{g}(\delta_I)) - \mu_D - \rho - \gamma_D - d_I - \mu_{2D} D_S^* - V_2 \\ &< 0. \end{aligned} \tag{3.49}$$

If $R_0^{[1]} + R_0^{[2]} < 1$, then

$$\begin{aligned} \frac{d_I \tilde{g}(\delta_I)}{V_1} + \frac{\beta_D \beta_M M_S^*}{V_1 V_2 D_S^*} &< 1 \\ d_I \tilde{g}(\delta_I) V_2 + \frac{\beta_D \beta_M M_S^*}{D_S^*} &< V_1 V_2 \\ V_1 V_2 - d_I \tilde{g}(\delta_I) V_2 - \frac{\beta_D \beta_M M_S^*}{D_S^*} &> 0 \end{aligned} \tag{3.50}$$

Hence, the constant term of the characteristic equation

$$\begin{aligned} A_2A_4 - \tfrac{\beta_D\beta_M M_S^*}{D_S^*} &= -(d_I\tilde{g}(\delta_I) - V_1)V_2 - \frac{\beta_D\beta_M M_S^*}{D_S^*} \\ &= V_1V_2 - d_I\tilde{g}(\delta_I)V_2 - \frac{\beta_D\beta_M M_S^*}{D_S^*} \\ &> 0. \end{aligned} \tag{3.51}$$

Therefore, both roots of the quadratic (i.e. the two eigenvalues) must have negative real parts. Thus, under the given conditions, the system is stable at DFE.

($\Longrightarrow$) If the DFE of model (3.1) is locally asymptotically stable, then by Theorem 8.12.*iii* of [24], the real parts of all eigenvalues of the Jacobian matrix A are negative. By (3.50), this occurs when $V_1V_2 - d_I\tilde{g}(\delta_I)V_2 - \frac{\beta_D\beta_M M_S^*}{D_S^*} > 0$ which is the same as $R_0^{[1]} + R_0^{[2]} < 1$. □

We must now prove the existence of an endemic equilibrium solution in the proposed model. However, this is difficult as two of the variables, D_S and D_I, are contained within the integral dispersion terms. Therefore, we utilize a technique called the chain trick [27] to reduce model (3.1) to an ODE model.

Reduction to ODE Model

Using the chain trick method [27], we can rewrite the first two equations as

$$\begin{aligned} \frac{dD_S}{dt} &= \lambda_D D_N - \frac{\beta_D M_I D_S}{D_N} - ((\mu_D + \rho) + d_S + \mu_{2D}D_N)D_S + \overline{D}_S \\ \frac{dD_I}{dt} &= \frac{\beta_D M_I D_S}{D_N} - ((\mu_D + \rho) + \gamma_D + d_I + \mu_{2D}D_N)D_I + \overline{D}_I \end{aligned} \tag{3.52}$$

where

$$\overline{D}_S = d_S \int_0^\infty g(z) e^{-\delta_S z} D_S(t-z) dz \tag{3.53}$$

and

$$\overline{D}_I = d_I \int_0^\infty g(z) e^{-\delta_I z} D_I(t-z) dz. \tag{3.54}$$

These quantities are treated as new model variables, so we may now differentiate both of them and amend the existing set of equations.

In time delay models, there are two distributions that are commonly used. The first is a uniform distribution with mean τ given by

$$g(u) = \begin{cases} \frac{1}{\tau\rho}, & \text{for } \tau\left(1-\frac{\rho}{2}\right) \leq u \leq \tau\left(1+\frac{\rho}{2}\right) \\ 0, & \text{elsewhere.} \end{cases} \tag{3.55}$$

The second is the gamma distribution given by

$$g(u) = \frac{u^{p-1}\alpha^p e^{-\alpha u}}{\Gamma(p)}, \tag{3.56}$$

where $\alpha, p \geq 0$ are parameters which determine the shape of the distribution and the mean of the distribution is p/α. In the case when $p = 1$, the result is the exponential distribution, $g(z) = \alpha e^{-\alpha z}$. Using (3.56) with $p = 1$, the expression for $\frac{d\overline{D}_S}{dt}$ is computed to be:

$$\begin{aligned}
\overline{D}_S &= \int_0^\infty g(z)e^{-\delta_S z}D_S(t-z)dz \\
&= \int_{-\infty}^t g(t-u)e^{-\delta_S(t-u)}D_S(u)du \\
&= e^{-\delta_S t}\int_{-\infty}^t g(t-u)e^{-\delta_S u}D_S(u)du \\
&= e^{-\delta_S t}\int_{-\infty}^t \alpha e^{-\alpha(t-u)}e^{-\delta_S u}D_S(u)du \\
&= \alpha e^{-(\delta_S+\alpha)t}\int_{-\infty}^t e^{(\delta_S+\alpha)u}D_S(u)du
\end{aligned} \tag{3.57}$$

Thus, by the product rule for differentiation,

$$\begin{aligned}
\frac{d\overline{D}_S}{dt} &= \alpha(-(\delta_S+\alpha))e^{-(\delta_S+\alpha)t}\int_{-\infty}^t e^{(\delta_S+\alpha)u}D_S(u)du + d_S\alpha e^{-(\delta_S+\alpha)t}e^{(\delta_S+\alpha)u}D_S(t) \\
&= -(\delta_S+\alpha)\overline{D}_S + \alpha D_S
\end{aligned} \tag{3.58}$$

The simplification is the same for $\overline{D}_I$, and so the delayed model in (3.1) is reduced to the ODE model formulated by

$$\begin{aligned}
\frac{dD_S}{dt} &= \lambda_D D_N - \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + d_S + \mu_{2D} D_N) D_S + d_S \overline{D}_S \\
\frac{dD_I}{dt} &= \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + \gamma_D + d_I + \mu_{2D} D_N) D_I + d_I \overline{D}_I \\
\frac{dM_S}{dt} &= \lambda_M M_N - \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_S \\
\frac{dM_I}{dt} &= \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_I \\
\frac{d\overline{D}_S}{dt} &= -(\delta_S + \alpha)\overline{D}_S + \alpha D_S \\
\frac{d\overline{D}_I}{dt} &= -(\delta_I + \alpha)\overline{D}_I + \alpha D_I.
\end{aligned} \tag{3.59}$$

The disease free equilibrium (DFE) is computed to be

$$\begin{aligned}
D_S^* &= \frac{\alpha + (\delta_S + \alpha)(\lambda_D - (\mu_D + \rho) - d_S)}{(\delta_S + \alpha)\mu_{2D}} \\
D_I^* &= 0 \\
M_S^* &= \frac{\lambda_M - (\mu_M + \sigma)}{\mu_{2M}} \\
M_I^* &= 0 \\
\overline{D}_S^* &= \frac{\alpha[d_S \alpha + (\delta_D + \alpha)(\lambda_D - (\mu_D + \rho) - d_S)]}{(\delta_S + \alpha)^2 \mu_{2D}} \\
\overline{D}_I^* &= 0.
\end{aligned} \tag{3.60}$$

In the next section, we provide the numerical simulations of the ODE model (3.59) and the R_0 expression (3.34).

Numerical Simulations

Using MATLAB 9.1, we generated the surface plots of R_0 values based on the model parameters $R_0^{[1]}$ and $R_0^{[2]}$ (See Figure 6). As proven in Proposition 2, Figure 6a shows that R_0 is an increasing function with respect to d_S. The influx of

additional, susceptible deer into a patch leads to an increased number of potential interactions with infected midges and thus an increase in the number of infections overall. Figure 6c shows that R_0 is an increasing function with respect to δ_S and a decreasing function with respect to δ_I. Figures 6a and 6b demonstrate the behavior of R_0 with respect to the influx of infected deer, d_I. For smaller values of d_S or $\beta_D\beta_M$, R_0 is an increasing function with respect to d_I; for larger values of d_S or $\beta_D\beta_M$, it is a decreasing function with respect to d_I. Thus, there must be a critical value ($d_S^{[c]}$ or $(\beta_D\beta_M)^{[c]}$) where the behavior changes.

If we consider R_0 as a function of the deer-midge interactions, then R_0 is essentially a linear function of $R_0^{[1]}$ and a function of the square root of $R_0^{[2]}$. The graph of R_0 would be increasing and concave down with respect to an increase in $R_0^{[2]}$ (See Figure 6d). This is consistent with what we would expect to happen. As the amount of interaction increases, so does the number of potential new infections with a greater chance of an outbreak occurring. Plus, as a greater proportion of the deer population becomes infected, the rate of increase of R_0 must decrease as the number of uninfected deer will consequently drop.

We also demonstrate numerically that the solutions of model (3.1) converge to the endemic equilibrium if $R_0 > 1$ and achieves a disease free equilibrium if $R_0 < 1$. To do this, a MATLAB code was written utilizing the ODE45 solver, and the results were verified against the computed R_0 value for a given set of parameters. At time $t = 0$, we have the following initial values: $D_S(0) = 30$, $D_I(0) = 10$, $M_S(0) = 20$, $M_I(0) = 5$, $\overline{D}_S(0) = 10$, and $\overline{D}_I(0) = 1$. See Table 6 for the specific parameter values used for the numerical simulations.

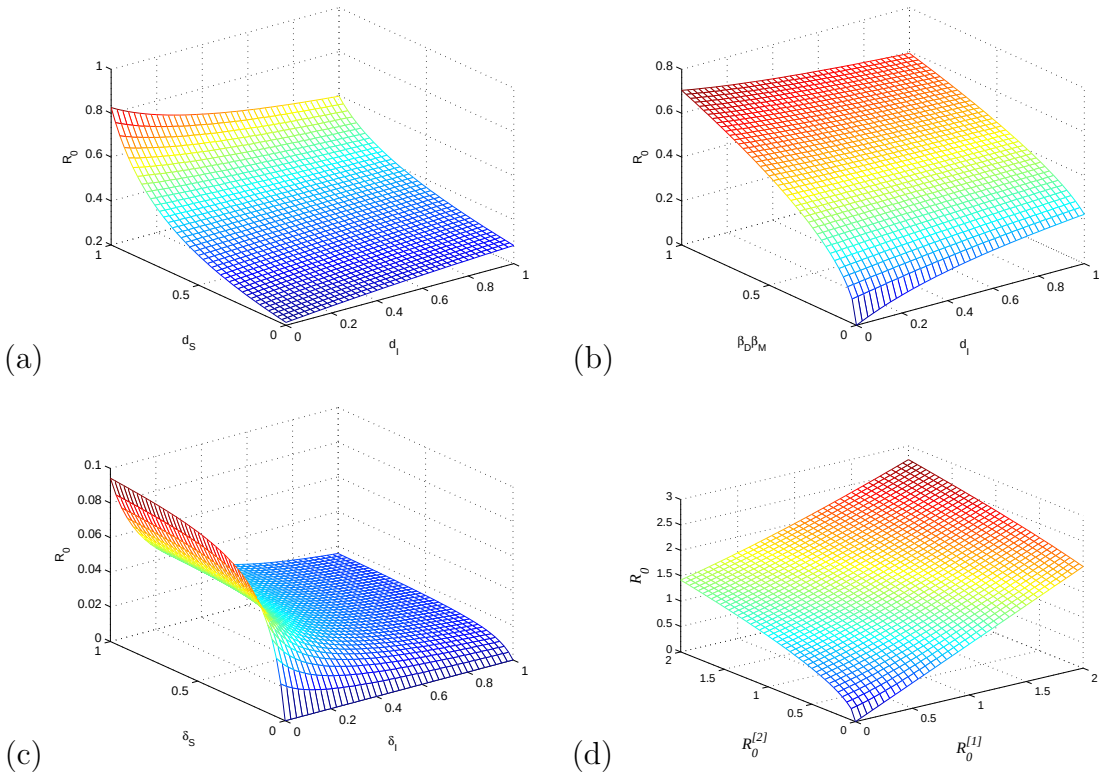


Figure 6. Numerical simulations of R_0 as a function of the selected model parameters. (a) R_0 values increase with d_I provided d_S values are small. When d_S values are large, R_0 decreases with d_I. (b) R_0 increases both with $\beta_D\beta_M$ and d_I. (c) R_0 increases with δ_S and decreases with δ_I. (d) R_0 increases linearly with $R_0^{[1]}$ and increases parabolically with $R_0^{[2]}$.

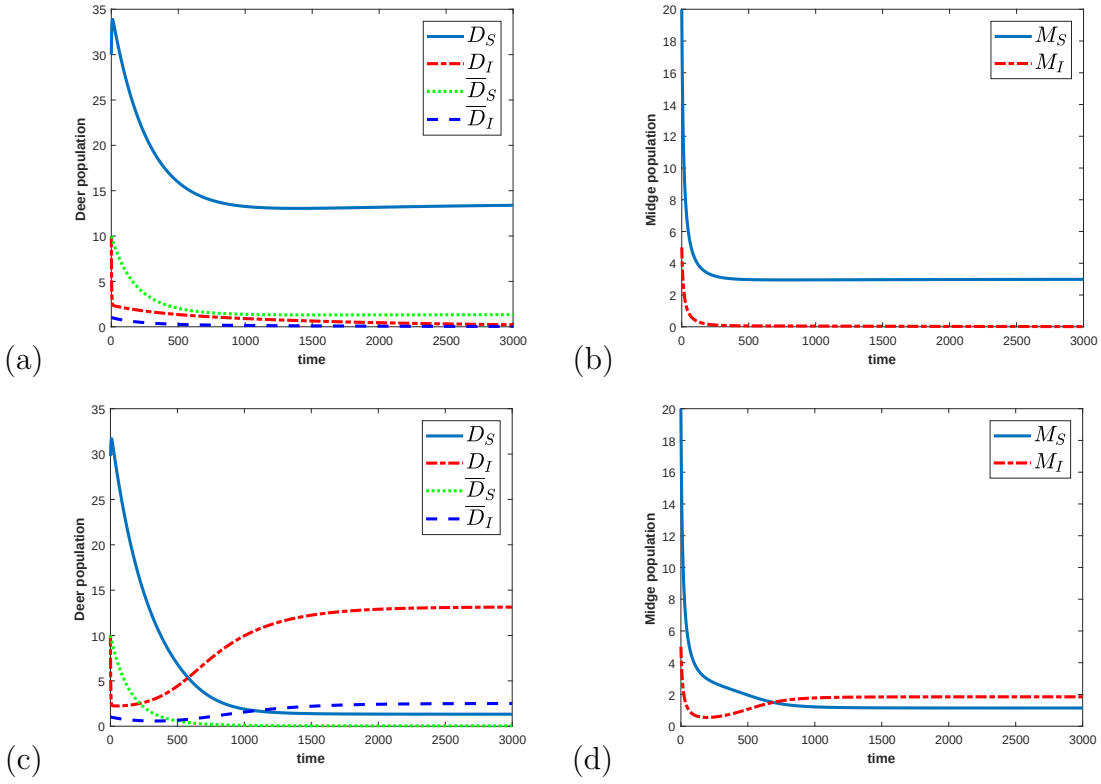


Figure 7. (a), (b) When the basic reproduction number $R_0 < 1$, the system stabilizes to its disease free equilibrium and the number of infected deer, the number of dispersing infected deer, and the number of infected midges tends to zero as t increases. (c), (d) When the basic reproduction number $R_0 > 1$, the system stabilizes to its endemic equilibrium. See Table 6 for the specific values used and the corresponding values of R_0.

Table 6. Parameter values used in model simulation and the calculated R_0 values.

Parameter	Value when $R_0 = .40$	Value when $R_0 = 2.19$	Parameter	Value when $R_0 = .40$	Value when $R_0 = 2.19$
β_D	0.2	1.1	γ_D	0.35	0.1
λ_D	0.9	0.9	β_M	0.2	1.6
ρ	0.2	0.2	λ_M	0.9	0.9
μ_D	0.1	0.1	σ	0.2	0.2
μ_{2D}	0.2	0.5	μ_M	0.05	0.05
d_S	0.3	0.1	μ_{2M}	0.05	0.05
d_I	0.3	0.42			

Note: The R_0 values are consistent with the numerical simulations shown in Figure 7. Similar results were obtained using different sets of parameter values.

Figures 7a and 7c show the long-term behavior of the four classes of deer populations - total susceptible, total infected, susceptible influx, and infected influx - plotted on the same graph, while figures 7b and 7d show the long-term behavior of the susceptible and infected midge populations. Figures 7a and 7b indicate that when $R_0 < 1$, the system will stabilize to its disease free equilibrium. Figures 7c and 7d show that when $R_0 > 1$, the system will stabilize to an endemic equilibrium. These outcomes are robust for large sets of initial values and parameter values.

Discussion

In this chapter, we have developed a distributed delay model for transmission dynamics of HD in a deer population. Though mathematical models for disease and HD specifically are established, we chose to focus on how the dynamics are affected by the dispersion (migration) of deer specifically and how the basic reproduction number is affected by these dispersion rates (i.e., d_S and d_I). The results show that there are critical values for the interaction parameters $(\beta_D\beta_M)^{[c]}$ and rates of susceptible deer dispersion $d_S^{[c]}$. Hence, possible outbreaks could be avoided by controlling how and where these deer move. The R_0 expression provides insights into the effects of deer movement on the spread of disease.

One of the primary limitations at this stage is the lack of actual parameter values. Although the qualitative behavior of model (3.1) remains fairly distinctive, (i.e., convergence to DFE or EE) for large sets of parameter values, some of the values were chosen based on reasonable assumptions. It is our goal to estimate some of the parameter values using data from the Missouri Department of Conservation concerning the prevalence of HD in Missouri's white-tailed deer. Nevertheless, the graphs presented in Figures 6 and 7 show consistent tendencies in the behavior in the model. We also have not considered behavior in a multi-patch system, where migrating individuals leave one patch and eventually enter a neighboring patch, nor did we consider a delay in the traveling time. Holt [20] and Weisser et al. [60] extended their results to a system of multiple patches joined through a pool of dispersing individuals. Moreover, the proposed model (3.1) does not include the effect of predators on the population of white-tailed deer. As a prey species, deer are linked with local

predators. In Missouri, the coyote is one such predator. Some coyote predator studies have been done, but these are admittedly outdated. However, deer make up a portion of a coyote's diet and that large increases or decreases in predator populations may influence deer mortality rates [12]. Finally, our model assumed only one vector for the transmission of HD. With the species richness of the *Culicoides* genus, we may reasonably expect more and different interaction rates and different levels of control efficacy [43]. We also note that weather has an effect on both the midge population and the life cycle of the HD virus [50, 35]. Midge populations thrive in damper areas, and in 2012, there was an above average amount of rain in the late winter/early spring, filling ponds and other water bodies in Missouri [12]. In addition, record warm temperatures in that spring and summer may cause midges to become more active sooner than normal [12]. Next, the high temperatures caused water sources to dry up, and not only did the resulting mud flats become ideal breeding areas for subsequent generations of midges, but also caused deer to visit water sources more frequently due to lower water content in the plants they ate as part of their diet. These same high temperatures also cause female midges to lay more eggs, and Wittmann et al. also revealed that higher temperatures decrease the extrinsic incubation period of the HD virus within the midges [62]. Thus, the virus develops faster and allows a midge to infect more deer during its life span. None of these factors have been considered in the model (3.1). Instead, the main focus has been on migration effects of deer population on overall HD dynamics within a patch.

The above mentioned limitations demand model extensions to study the effectiveness of control and preventive strategies. Deer species are important members

of the ecosystem as they feed on brush and grass in a given area and keep them in check. This is a first step towards inclusion of migration effects of deer population modeling of HD dynamics. ODE models have been effective tools in modeling real-world dynamics as solutions to these equations (specific solutions or simulation) mimic these dynamics. The primary limitation here is how many different aspects we wish to include in a model. Simulations will still prove insights into behaviors, but as a model becomes more complicated, justifying its mathematical validity becomes more and more difficult. Thus we must be able to strike a balance between a perfect model and a useable one.

CHAPTER 4

PARAMETER ESTIMATES

Overview

In order to estimate the parameters in model (3.1), we use data supplied by the Missouri Department of Conservation. Some of them may be estimated from the data itself using the estimated deer population, estimated number of HD incidents, and number of deer harvested for each year. Some of the remaining parameters may be estimated using the MATLAB curve-fitting app for some and three scripts for the remaining. To simplify computations, a sequence of adjustments are made to the model. Also, due to the spiked nature of the data (as breakouts occur at specific times), we make adjustments to the model to emphasize those specific times and to model years of major outbreaks. When MATLAB fails to successfully provide results due to computational issues, we can further simplify the model by combining the first two equations of model (3.1) and deriving a function that closely models the number of infected deer, D_I. From this point, MATLAB is finally able to provide a list of parameter values.

The information and results in this chapter will be published to fill the current gap in the existing literature about the specific parameter values derived[5].

Methods

Three of the parameters (ρ, γ_D, and μ_{2D}) are calculated directly from the data. The computation of ρ requires only the data itself, and the other two may be

estimated using MATLAB's curve-fitting app. For most of the remaining parameters, we use MATLAB to simulated their values using a variety of techniques and model adjustments. As no migration information is available, we remove any parameters that deal with dispersion. Still, because of the large number of parameters to estimate, it is ideal to have as many data values as possible. However, summary values are only available on a yearly basis, and there most of the data collected has only been over the last ten years or so. Thus, we can increase the number of incident values by changing the time scale from a yearly one to a monthly one, and then assuming the number of incidents per year follows a normal distribution over the summer months. The number of incidents remains the same each year; they are just divided up over different points in time during the year.The remainder of the parameters we are able to estimate are simulated with three MATLAB scripts. See Appendix B. One of the scripts is the model itself and its evolving variants. A second script takes a given set of parameter values, uses them in the model, and returns the sum-of-squared errors (SSE) between the simulated values and the actual data values over the study period. The third script repeatedly runs the second, making slight variations to the parameters each time in search for set with the lowest SSE. Upper and lower boundaries for the parameters may also be applied. From this optimized set of parameter values, the numerical simulation is graphed against the MDC data. Using these scripts, and a derived function for D_I, we obtain estimates for λ_D, μ_D, β_M, λ_M, μ_M, and μ_{2M} within a simplified model.

Table 7. Summary of the parameter estimates in the delayed HD model (3.1)

Parameter	Value	Bounds [1],[2]
λ_D	196.315	$(-1.042, 1.071)^{[1]}$
ρ	0.1958	$(0.1851, 0.2064)^{[1]}$
μ_D	0.4385	$(0, 7.95 \times 10^{-4})^{[2]}$
μ_{2D}	2.608×10^{-11}	$(-1.596 \times 10^{13}, 1.597 \times 10^{13})^{[1]}$
γ_D	0.2243	NA[3]
β_M	460.2656	$(0, 1500)^{[2]}$
λ_M	0.6980	$(0, 1.6)^{[2]}$
μ_M	0.9363	$(0, 2)^{[2]}$
μ_{2M}	2.29×10^{-6}	$(0, 1)^{[2]}$

Notes: [1] *95% Confidence interval (when available).* [2] *Upper and lower bounds used within MATLAB as part of the fmincon routine.* [3] *The value for* γ_D *was based on the ratio of two estimated values in the best-fit logistic function using MATLAB's curve fitting app, and thus a 95% confidence interval was not available.*

Estimating the Parameters

In order to estimate the parameters in model (3.1), I used data provided by the MDC, which included the estimated population, the total harvest, and the total number of reported HD cases from 2005 to 2014. The estimate for ρ was the most straightforward. For each year between 2005 and 2014, the ratio of the deer harvested from the total population was computed, and ρ is the average of those values.

Estimates for the remaining parameters were found utilizing MATLAB. The set of parameters found represent the dynamics for the entire state of Missouri as one single patch. As such, and because of the limited number of data values, I assume there is no dispersion. The simplified model becomes

$$\begin{aligned}
\frac{dD_S}{dt} &= \lambda_D D_N - \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + \mu_{2D} D_N) D_S + \\
\frac{dD_I}{dt} &= \frac{\beta_D M_I D_S}{D_N} - (\mu_D + \rho + \gamma_D + \mu_{2D} D_N) D_I + \\
\frac{dM_S}{dt} &= \lambda_M M_N - \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_S \\
\frac{dM_I}{dt} &= \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_I.
\end{aligned} \tag{4.1}$$

Thus, there is no need to find d_S, d_I, δ_S and δ_I. To get the most representative values for λ_D and μ_{2D}, only the time interval between 2007 and 2012 was used as those two years were years of significant HD outbreaks, and it seems a reasonable assumption that the population will behave the most normally and predictably during this time. Curve fitting was obtained by using MATLAB's curve fitting app using the estimated population data (See figure 8). Without the presence of the disease,

the model predicts that the population will grow logistically. Thus, the custom fit equation $f(x) = \frac{D_0 K}{D_0+(K-D_0)e^{-rx}}$ was used. The deer population growth rate, λ_D, is equivalent to r in this function, and the death rate, μ_{2D}, is equal to r/K. In addition, to force the app to fit the regression curve, reasonable bounds on each of the parameters in the custom equation were used (such as each value is non-negative) and adjusted to minimize the mean-square error. The resulting estimated value of λ_D was 0.01467 with a 95% confidence interval of $(-1.042, 1.071)$. The resulting estimated value of μ_{2D} was 2.608×10^{-11} with a 95% confidence interval of $(-1.596 \times 10^{13}, 1.597 \times 10^{13})$. I note here that the confidence intervals are quite large, most likely due to estimating two data values with only five data points where the data values (the population of deer) are all around 1.5 million.

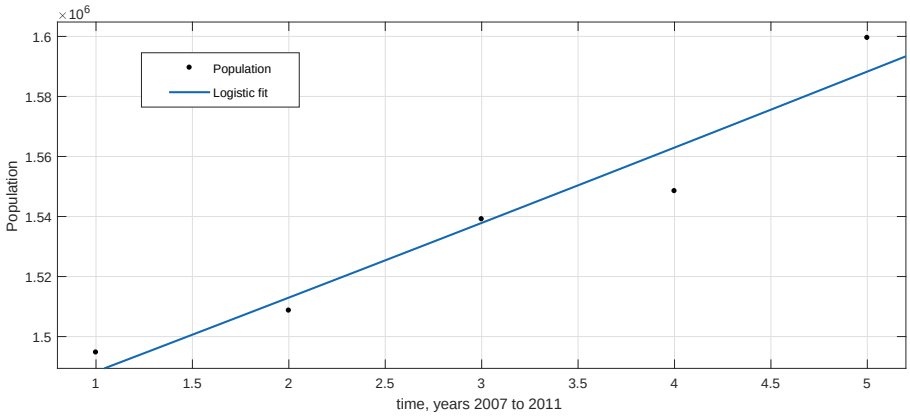


Figure 8. The estimated population of deer for 2007 - 2011 and the best-fit logistic curve via MATLAB's curve fitting app

To estimate γ_D, MATLAB's curve fitting app was once again used. If we isolate the γ_D term in the second equation of the model, the result is

$$\frac{dD_I}{dt} = \gamma_D D_I$$

which has a solution of

$$D_I(t) = D_0 e^{\gamma_D t}$$

for some D_0. Applying the natural logarithm to the equation yields

$$lnD_I(t) = lnD_0 + \gamma_D t.$$

Therefore, I applied a linear fit on the natural log of the number of HD cases reported over all years between 2005 and 2014, and the coefficient of the linear term is the desired value for γ_D. Thus, $\gamma_D = 0.2243$ with a 95% confidence interval of (-0.353, 0.8016). In any instances when there were zero cases reported, the value was changed to one as this would have an insignificant effect on the overall estimation of the parameter and while still allowing the natural log to be applied (See figure 9).

I used MATLAB to estimate the remaining parameters via three scripts. One script was the original model, the second computed the total sum of squared errors (SSE) between the actual data values and the simulated data values given a set of parameters, and the third varied the parameters within a given set of bounds. Each time this third script generated a new, proposed set of parameter values, it would send these values to the second script to compute the SSE. This process was repeated until either the change in the SSE fell within a certain tolerance lever or until the

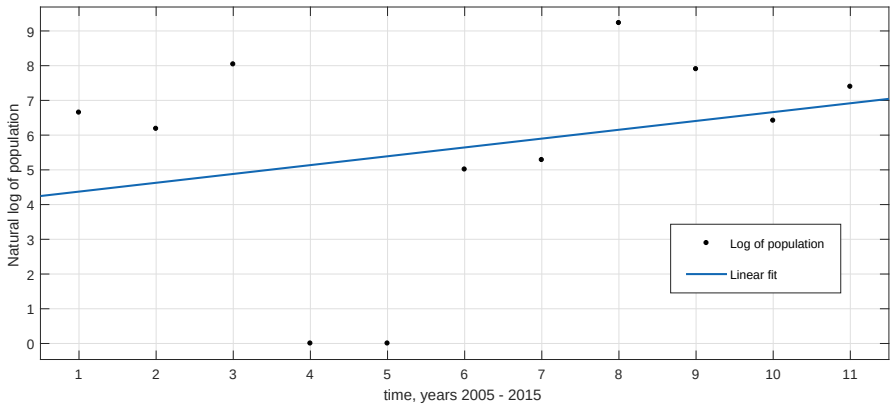


Figure 9. The natural logarithm of the number of reported incidents of HD for 2005 - 2015 and the best-fit linear curve via MATLAB's curve fitting app

process ran through a user-defined number of iterations.

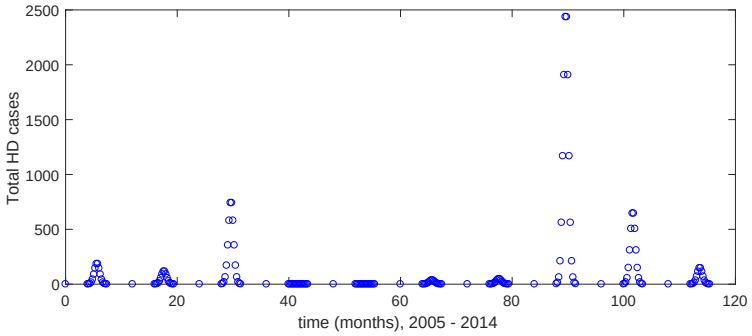
The first attempts to estimate the parameters used the simplified model without dispersion 4.1. The values for the total deer population were those provided by the MDC. As HD is fatal in roughly 80% of cases, the total number of reported cases in each year was divided by 0.8 to reflect the true number of cases (but not necessarily fatal cases). Initial code runs produced numerical results that somewhat reflected the susceptible population but not the infected population. To force a better fit with the infected population, a penalty factor was multiplied to the computation of the error of infected. Even when the penalty factor was 100 or more, the code failed to approximate the infected population, and, at the same, increased the error with the susceptible population. I also attempted to scale the susceptible population size so those values were closer to the infected population and to shorten the run time of the code. This also failed to have any noticeable improvement on the infected solution

curve.

Next, the timescale for the approximation was changed from years to months in the hopes that more iterations of the code running might improve results. As part of this change, the number of infected deer needed to be represented by more than one single value in a year. To simulate the number of infected deer at any time t, the following method was employed. As nearly all cases of HD occur during the late spring and summer, I assume this takes place over a sixteen week period roughly centered on July 1st. I also assume that the total number of infected deer during this time follows a normal distribution whose mean occurs at July 1st. I only use the portion of the normal distribution that is within four standard deviations of the mean which corresponds to 99.9% of all values. Thus, the standard deviation corresponds to two weeks of time. It is then a simple task to compute the percentage of that values that fall within each two week period. For example, the percentage of values that fall between week six and week eight corresponds to the probability that z, the number of standard deviations units away from the mean, is in between -1 and 0. Here, $P(-1 < z < 0) = 0.3413$, and so 34.13% of the HD cases in each year occurred between weeks 6 and 8 of the infection season. I also assumed that all deer infected with HD within a two week period also died within that same two week period. Though only 85% of deer that are infected die within the first two weeks of contracting the disease, the extra 15% that should have been included in the subsequent time period were ignored as this value was much smaller compared to the the total number of cases per year and even yet smaller when compared to the total deer population as a whole (See Figure 10a). In addition to this, the total number

on non-HD and non-natural related deaths (for example, when a deer is hit by a car) was also assumed to have the same normal distribution and was apportioned in the same manner. Finally, to simulate the total population during the infection period, the population size was incrementally reduced by the sum of the number of HD and non-HD, non-natural deaths (See Figure 10b).

(a)



(b)

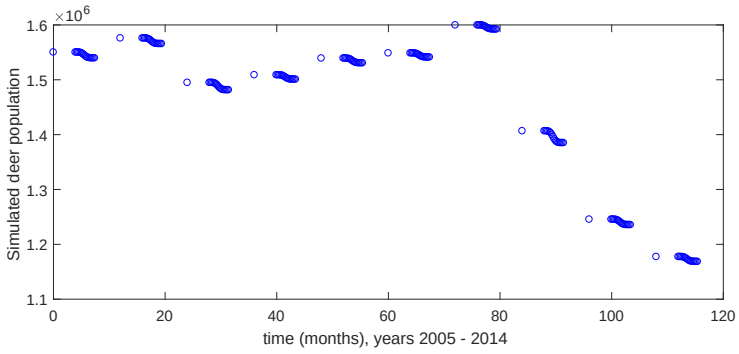


Figure 10. Simulation of (a) the number of HD incidents and (b) the total deer population per two-week period

Because I was considering the populations in two week increments, I added in time values that corresponded to these periods. For example, April 1st, 2005

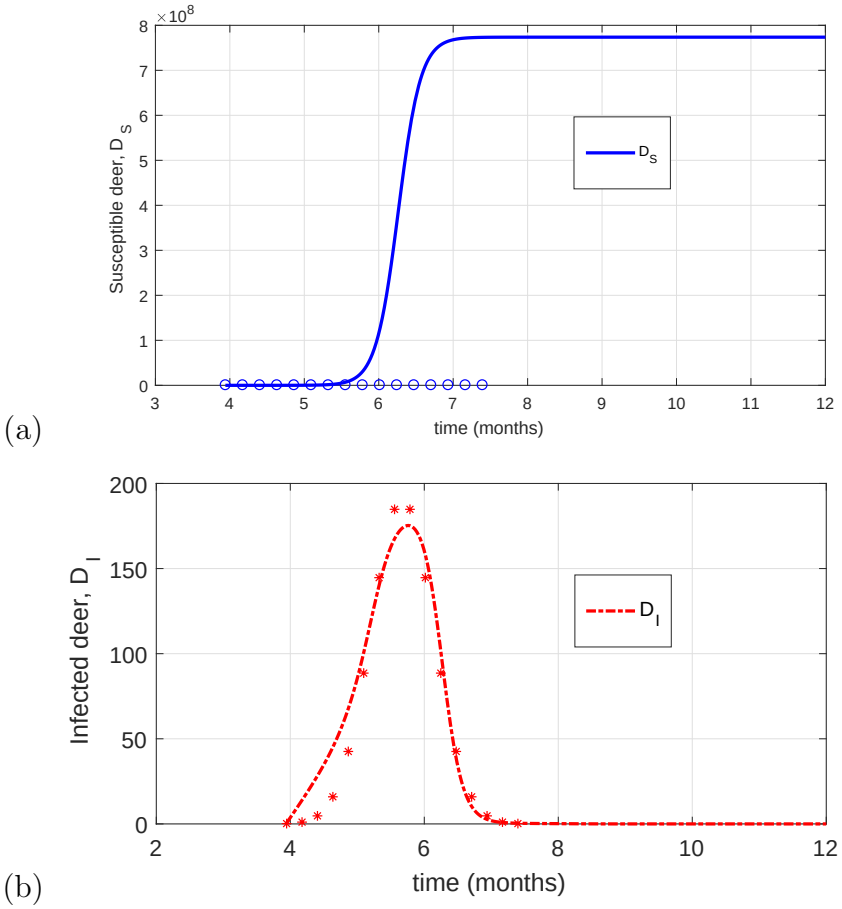
corresponded to $t = 120/365$, April 15th corresponded to $t = 134/365$, etc. Even with the simulated set of population and infected values, MATLAB either suffered from excessive run times (and was forced to stop) or produced solutions that was clearly not a good fit to the data as it showed an over trend evenly during the years rather than producing spikes during the infection season. To emphasize this time period, the third third equation in the model (corresponding to the birth rate of susceptible midges) was altered so that there would be an increase in the midge population during this time,

$$\frac{dM_S}{dt} = (1 + \epsilon \cos(\alpha t))\lambda_M M_N - \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N) M_S.$$

The introduction of cosine forces the infection period to be periodic, and the addition of 1 makes sure that the midge population remains positive. Results improved, but not enough.

To simplify matters, I decided to see if MATLAB could provide a good simulation of the number of infected deer during just the first year (See Figure 11), and then during three years rather than over the entire ten-year period.

The simulations were unable to simulate the outbreak years (2007 in the case of the first three year period), so a winter effect was added to the model. In the winter months before the larger outbreaks, temperatures were higher than normal, and it is theorized that these higher temperatures kept more midges alive and kept the HD virus more active. Therefore, the model needed an outbreak effect to allow for this increased midge activity to occur in certain years. To that end, the third equation of



(a)

(b)

Figure 11. Using MATLAB to attempt to numerically estimate the parameters in the simplified version of model (3.1), without dispersion. The graphs show the numerical solutions of the simplified model for (a) the number of susceptible deer and (b) the number of infected deer during 2005. This corresponds to the first "'hump"' in Figure 10

the model was updated to be a piecewise function.

$$\frac{dM_S}{dt} = \begin{cases} \dfrac{dM_{S1}}{dt} = (1 + \epsilon\cos(\alpha t))\lambda_M M_N - \dfrac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N)M_S, \\ \text{for } t \leq 24 \\ \dfrac{dM_{S2}}{dt} = (w + 1 + \epsilon\cos(\alpha t))\lambda_M M_N - \dfrac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma + \mu_{2M} M_N)M_S, \\ \text{for } t > 24 \end{cases} \tag{4.2}$$

The presence of w in $\frac{dM_{S2}}{dt}$ allows for the presence of more midges during 2007, and thus a greater number of infections. Here, w is treated as an additional parameter to be estimated with the hopes that a value for it will be determined so that the outbreak behavior will be captured.

For the next sequence of attempts, the winter-effect with outbreak model was slightly adjusted. Rather than have this effect applied to the first term of the equation for M_{S2}, it was applied to the third term in the following manner:

$$\frac{dM_{S2}}{dt} = (1 + \epsilon\cos(\alpha t))\lambda_M M_N - \frac{\beta_M D_I M_S}{D_N} - (\mu_M + \sigma - w + \mu_{2M} M_N)M_S.$$

Here, the thought was that there were more infections during this year not because midges were experiencing an increased growth rate. Instead, there were more infections because not as many midges were dying during the preceding winter.

At this point, I decided that the overall trends in the number of HD cases per year was more important than the number of susceptible and infected deer at specific times during the year. Plus, MATLAB was never able to produce a result accurate

to that size time scale. Therefore, I decided to use the original data corresponding to the number of cases reported each year and not the simulated values during each two-week period.

Because the initial estimate of λ_D was not consistent with (3.3), I adjusted the code to treat it as a parameter to estimate and set its lower bound to that of ρ. To reduce the number of computations MATLAB had to execute, more adjustments to the model were made. First, the first two equations of model (3.1) were combined to have one equation modeling the total population of deer, D_N, both susceptible and infected.

$$\frac{dD_N}{dt} = \lambda_D D_N - (\mu_D + \rho + \mu_{2D} D_N) D_N - \gamma_D D_I.$$

This also had the effect of eliminating β_D, which meant there was one less parameter the program had to estimate. In addition, σ was eliminated from the code as its role in the model is identical to μ_M. Thus, μ_M now simulated the death of midges both by natural causes and by control measures.

I also replaced D_I with an approximate function of t. None of the available functions in the curve fitting app gave a good fit to the infected deer data, but by removing the three largest outbreaks during this time period (2007, 2012, and 2013), the function

$$f_1(t) = 2422sin(0.0347t + 2.756) + 548sin(0.4754t + 2.758)$$

provided a good fit for the remaining data values. This also eliminated the need for α. To simulate the outbreaks of those three years, a combination of Gaussian

functions was generated. Time values corresponding to two weeks before and after the outbreaks were computed. For example, if the first outbreak occurs when $t = 3$, then $t = 3 \pm 2/52$ represent those two-week time values. Six integer values between 1 and 6 were generated, and these were used as values for D_I immediately before and after the time value of each outbreak. This is not to say we are adding to the total number of infected; this was done only to generate a Gaussian curve with a small standard deviation. The resulting Gaussian function generated is

$$f_2(t) = (1.021 \times 10^4)e^{(-\frac{t-8.002}{0.03872})^2} + 3099e^{(-\frac{t-2.998}{0.04318})^2} + 2696e^{(-\frac{t-9}{0.01103})^2},$$

and thus $D_I(t) = f_1(t) + f_2(t)$ (See Figure 12). The function was substituted into the updated model for all instance of D_I. Finally, now that the outbreaks were incorporated into the simulation via this $D_I(t)$, I was able to eliminate the need to estimate ϵ and w. However, the code still failed to execute fully as simulated population values were to large for MATLAB to run with them.

The last adjustment to the code was to insure that the simulation produced logistic growth instead of decay. To that end, λ_D was written as a function of ρ and μ_D,

$$\lambda_D = \rho + \mu_D + \epsilon$$

via (3.3), and the value of ϵ was set at a small value ($\epsilon = 7.6 \times 10^{-5}$) to control the simulated deer population values so that MATLAB would be able to run the code in its entirety. Here again, this also reduced the number of parameters MATLAB had to

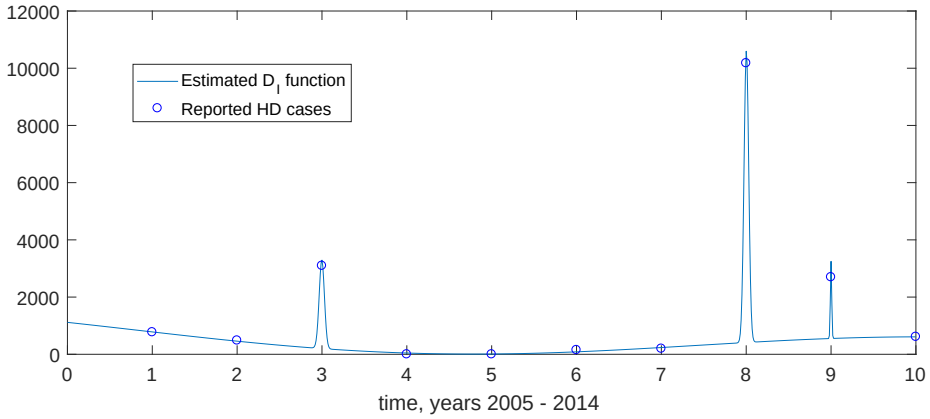


Figure 12. The simulated curve for $D_I(t)$ against the reported cases of HD by year, 2005 - 2014

simulate. In addition, a timescale of 10^{-3} factor was incorporated into the equation for D_N to help the program run and allow enough iterations for the model to exhibit logistic behavior. To compensate for this, the estimated values for parameters in this equation, λ_D and μ_D, were divided by this factor to give their true value. As before, the program was run many times, adjusting the lower and upper bounds of the parameter values. After each run, it was noted as an improvement only if the SSE was reduced and the resulting graph was a reasonable approximation of the number of infected deer over the time period. The final set of parameter values is given at the beginning of this section in table 7. Note that the output values of λ_D and μ_D were multiplied by 10^3 to compensate for the timescale factor in the code. Figure 13 shows the approximation for D_I. The simulated graph shows three vertical drops, and these correspond to the years of a major HD outbreak, and the rise in the middle of the graph is consistent with the logistic growth of the model. We must note that the

population data is an estimated population for each year, and thus our estimations could have errors. The code for this final version can be found in Appendix B.

Though MATLAB was able to produce results, the values for at least two of the parameters (λ_D and β_M) seem excessively large. This may be due to the lack of data in two respects. First because we only have a summary of yearly data. Second, there are two years (2008 and 2009) where the MDC data said there were zero cases reported. However, it is not clear if there were in fact, no cases or if data was not collected that year at all. Clearly we need more data, and more robust data, to obtain better results.

Using the parameter values derived in this chapter and, when needed, those assumed in Table 6, we determine that $R_0 = 0.0366$.

Table 8. Summary of the values used to compute R_0 as in 3.34

Parameter	Value
β_D	10.1
λ_D	196.315
ρ	0.1958
μ_D	0.4385
μ_{2D}	2.608×10^{-11}
d_S	0.3
d_I	0.3
γ_D	0.2243
δ_S	0.1
δ_I	0.1
β_M	460.2656
λ_M	0.6980
σ	0.2
μ_M	0.9363
μ_{2M}	2.29×10^{-6}

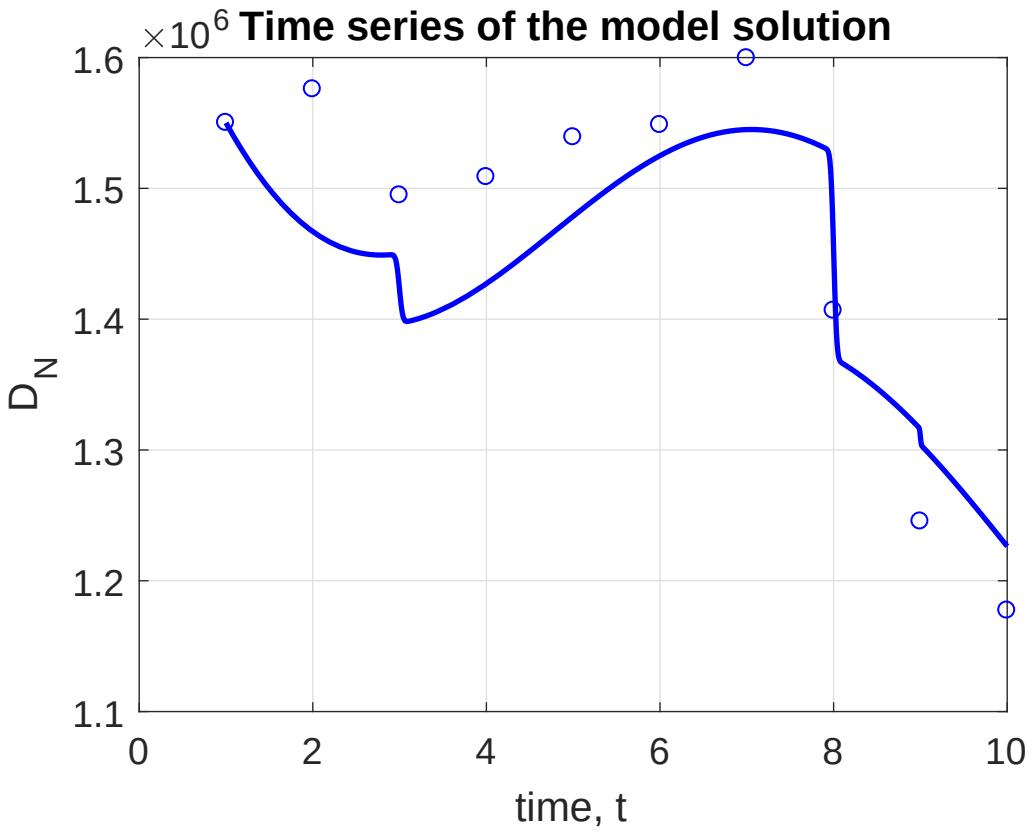


Figure 13. The estimated solution curve of the total deer population D_N against the actual population. The vertical drops in the simulated graph correspond to years of a major HD outbreak.

CHAPTER 5

MODEL EXTENSIONS, CONCLUSIONS, AND FUTURE WORK

In the last chapter, we developed a distributed delay model for transmission dynamics of HD in a deer population. Though mathematical models for disease and HD specifically are established, we chose to focus on how the dynamics are affected by the dispersion (migration) of deer specifically and how the basic reproduction number is affected by these dispersion rates (i.e., d_S and d_I). The results show that there are critical values for the interaction parameters $(\beta_D\beta_M)^{[c]}$ and rates of susceptible deer dispersion $d_S^{[c]}$. Hence, possible outbreaks could be avoided by controlling how and where these deer move.

Extension of the ODE Model to a Multi-Patch Model

From this point, there is a clear path to follow for subsequent study. When looking for the endemic equilibrium for the ODE model in (3.59), the Jacobian of the system is

$$B = \begin{bmatrix} B_1 & \lambda_D - \mu_{2D}D_S^* & 0 & -\beta_D & 1 & 0 \\ 0 & B_2 & 0 & \beta_D & 0 & 1 \\ 0 & \frac{\beta_M M_S^*}{D_S^*} & B_3 & \lambda_M - \mu_{2M}M_S^* & 0 & 0 \\ 0 & \frac{\beta_M M_S^*}{D_S^*} & 0 & B_4 & 0 & 0 \\ d_S\alpha & 0 & 0 & 0 & -(\delta_S+\alpha) & 0 \\ 0 & d_I\alpha & 0 & 0 & 0 & -(\delta_I+\alpha) \end{bmatrix}, \quad (5.1)$$

where $B_1 = \lambda_D - (\mu_D + \rho) - d_S - 2\mu_{2D}D_S^*$, $B_2 = -(\mu_D + \rho) - \gamma_D - d_I - \mu_{2D}D_S^*$, $B_3 = \lambda_M - (\mu_M + \sigma) - 2\mu_{2M}M_S^*$, and $B_4 = -(\mu_M + \sigma) - \mu_{2M}M_S^*$. As before, we attempt to find the eigenvalues of B via the characteristic equation $\det(B - \Lambda I) = 0$. The result is a polynomial in Λ of sixth degree. Using MATLAB, this polynomial can be factored as the product of three polynomials of degrees 1, 2 and 3, respectively. As the roots of the third degree polynomial cannot be determined easily, we can apply the Routh-Hurwitz criteria to determine when the eigenvalues will have negative real parts. However, the resulting needed criteria are too complicated to provide any useful information. So, it will be easier to analyze the model in terms of proportions of susceptible and infectious individuals, so we make the change of variables

$$u = \frac{D_S}{D_N^*}, \quad v = \frac{D_I}{D_N^*}, \quad w = \frac{M_S}{M_N^*}, \quad x = \frac{M_I}{M_N^*}, \quad y = \frac{\overline{D}_S}{D_N^*}, \quad z = \frac{\overline{D}_I}{D_N^*}, \tag{5.2}$$

so that

$$u + v = 1 \Rightarrow v = 1 - u, \quad w + x = 1 \Rightarrow x = 1 - w \tag{5.3}$$

We scale time t with the quantity $1/\mu_M$ by setting $\tau = \mu_M t$. We also scale the total populations by their respective carrying capacities by setting $D_N = ((\lambda_D - (\mu_D + \rho))/\mu_{2D})D_N^*$ and $M_N = ((\lambda_M - (\mu_M + \sigma))/\mu_{2M})M_N^*$. Hence, we introduce the following dimensionless parameters

$$\begin{gathered}\tau = \mu_M t, \quad \lambda = \frac{\lambda_D}{\mu_M}, \quad \gamma = \frac{\gamma_D}{\mu_M}, \quad \beta = \frac{\beta_M}{\mu_M}, \\ d_{s^*} = \frac{d_S}{\mu_M}, \quad d_{i^*} = \frac{d_I}{\mu_M}, \quad \delta_{s^*} = \frac{\delta_S + \alpha}{\mu_M}, \quad \delta_{i^*} = \frac{\delta_I + \alpha}{\mu_M} \\ \xi = \frac{\beta_D \mu_{2D}(\lambda_M - (\mu_M + \sigma))M_N^*}{\mu_{2M}(\lambda_D - (\mu_D + \rho))D_N^*}\end{gathered} \tag{5.4}$$

where the asterisks have been dropped. Thus, System 3.59 becomes

$$\begin{aligned}\frac{du}{d\tau} &= \lambda(1-u) + u(1-u)(\gamma - d_{i^*}) + \tfrac{w}{\mu_M} - \xi u w \\ \frac{dw}{d\tau} &= 1 - w - \beta(1-u)w \\ \frac{dy}{d\tau} &= -\delta_{s^*} y + d_{s^*}\alpha u \\ \frac{dz}{d\tau} &= -\delta_{i^*} z - d_{i^*}\alpha(1-u)\end{aligned} \tag{5.5}$$

From here, it is hoped that the subsequent analysis will yield results consistent with those in Chapter 3.

A Two-Patch Model

In actuality, a single-patch model is too simplistic. In the real world, animal populations may divide themselves into many sub-population groups and live within multiple habitats. A simple two-patch extension of the model in 3.1 is given by:

$$
\begin{aligned}
\frac{dD_S^{[1]}}{dt} &= \lambda_D D_N^{[1]} - \frac{\beta_D M_I^{[1]} D_S^{[1]}}{D_N^{[1]}} - ((\mu_D + \rho) + d_S + \mu_{2D} D_N^{[1]}) D_S^{[1]} - D_S^{[1]} d_S + \\
&\quad d_S \int_0^\infty g(z) e^{-\delta_S z} D_S^{[2]}(t-z) dz \\
\frac{dD_S^{[2]}}{dt} &= \lambda_D D_N^{[2]} - \frac{\beta_D M_I^{[2]} D_S^{[2]}}{D_N^{[2]}} - ((\mu_D + \rho) + d_S + \mu_{2D} D_N^{[2]}) D_S^{[2]} - D_S^{[2]} d_S + \\
&\quad d_S \int_0^\infty g(z) e^{-\delta_S z} D_S^{[1]}(t-z) dz \\
\frac{dD_I^{[1]}}{dt} &= \frac{\beta_D M_I^{[1]} D_S^{[1]}}{D_N^{[1]}} - ((\mu_D + \rho) + \gamma_D + d_I + \mu_{2D} D_N^{[1]}) D_I^{[1]} - D_I^{[1]} d_I + \\
&\quad d_I \int_0^\infty g(z) e^{-\delta_I z} D_I^{[2]}(t-z) dz \\
\frac{dD_I^{[2]}}{dt} &= \frac{\beta_D M_I^{[2]} D_S^{[2]}}{D_N^{[2]}} - ((\mu_D + \rho) + \gamma_D + d_I + \mu_{2D} D_N^{[2]}) D_I^{[2]} - D_I^{[2]} d_I + \\
&\quad d_I \int_0^\infty g(z) e^{-\delta_I z} D_I^{[1]}(t-z) dz \\
\frac{dM_S^{[1]}}{dt} &= \lambda_M M_N^{[1]} - \frac{\beta_M D_I^{[1]} M_S^{[1]}}{D_N^{[1]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[1]}) M_S^{[1]} \\
\frac{dM_S^{[2]}}{dt} &= \lambda_M M_N^{[2]} - \frac{\beta_M D_I^{[2]} M_S^{[2]}}{D_N^{[2]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[2]}) M_S^{[2]} \\
\frac{dM_I^{[1]}}{dt} &= \frac{\beta_M D_I^{[1]} M_S^{[1]}}{D_N^{[1]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[1]}) M_I^{[1]} \\
\frac{dM_I^{[2]}}{dt} &= \frac{\beta_M D_I^{[2]} M_S^{[2]}}{D_N^{[2]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[2]}) M_I^{[2]}.
\end{aligned}
\tag{5.6}
$$

The superscripts indicate the patch number. As with the single-patch case, we assume that the entirety of the deer population resides in these two sole habitat patches, but when an individual deer migrates, it will leave one patch and arrive at the other patch at some time z later.

A Multi-Patch Model

We now assume that we can divide any geographic area into a set of patches, and we will apply the migration model in each patch. Kouokam et al.[26] considered a two-patch epidemiological system, where local interactions were governed by the classical Susceptible-Infected-Recovered-Susceptible (SIRS) model, specifically considering the rate of migration between patches was much smaller than the time-scale corresponding to the infection. In this case, they show that the basic reproduction number R_0 is smaller in the multi-patch case than in the single-patch case. They further explore the effects within more than two patches. Stability of multi-patch systems has been described by Jansen and Lloyd[23]. For prey-predator models, Auger and Benoît[1] study the dynamics of slow-moving and fast-moving activity sequences of animals, breaking down large scale ecological systems into a simplified organization of smaller ones. With simulations of the whole system and for the reduced systems, they show that behavior for reduced systems are very close the those of the whole system when the magnitudes for the slow and fast dynamics are sufficiently different.

When deer disperse, they will migrate into a neighboring patch. We can consider different types of arrangements of these patches. Neubert et al.[39] consider a ring of $m \geq 3$ identical, evenly spaced patches. Each patch has exactly two neighboring patches, and, consequently, only two different directions in which to leave the patch. The number of deer dispersing into a patch would be the combined total of half of the number of deer dispersing out of each neighboring patch. We assume that all migrations are taking place within the ring and that no deer are emigrating from anywhere outside the patches in this ring. Thus, the total emigration (flow) of

deer within the network is zero and can be represented by a one-dimensional Laplace equation

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} = 0, \tag{5.7}$$

where $u(t)$ represents the population of deer in a patch at time t. If we consider three adjacent patches in the ring - labeled u_{i-1}, u_i, and u_{i+1} for convenience - then using the finite difference method, we may estimate the emigration with the discretization

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &\approx d(u_{i-1} - 2u_i + u_{i+1}) \\ &= 2d\Big(\tfrac{u_{i-1}}{2} - u_i + \tfrac{u_{i+1}}{2}\Big), \end{aligned} \tag{5.8}$$

where $d > 0$ is a constant. If we consider each node as a patch, we can see that one "unit" of deer will leave the central patch while one half of a deer "unit" will enter the patch from both the left patch and right patch.

Thus, the differential equation for the number of susceptible deer becomes

$$\begin{aligned} \frac{dD_S^{[i]}}{dt} = \; & \lambda_D D_N^{[i]} - \frac{\beta_D M_I^{[i]} D_S^{[i]}}{D_N^{[i]}} - ((\mu_D + \rho) + d_S + \mu_{2D} D_N^{[i]}) D_S^{[i]} - \\ & D_S^{[i]} d_S + \frac{d_S}{2} \int_0^\infty g(z) e^{-\delta_S z} D_S(t - z) dz \end{aligned} \tag{5.9}$$

Next, we may consider patches in a line. This is essentially the same arrangement as the ring except that exactly two of the patches (the patches at the ends of the line) have only one neighbor and not two. For these two patches, the number of deer dispersing into the patch must be adjusted as there is only one neighbor. Moreover, the number of deer migrating into the boundary patch's neighbor must

also be adjusted as these two patches are receiving half the dispersing deer from their inner neighbor and all of the dispersing deer from their boundary neighbor. Thus, the resulting system of three equations will be the same as above for all patches that are at least two patches away from the boundary. Four additional equations would have to be added for the described special cases.

When considering dispersion in two dimensions, we consider the Laplace equation

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \tag{5.10}$$

How we discretize the migration depends on how we tessellate a particular area. If we consider a conventional square grid as in Figure 14, we can use a central discrete scheme where

$$\begin{aligned} \nabla^2 u = \left(\tfrac{\partial^2}{\partial x^2} + \tfrac{\partial^2}{\partial y^2}\right) u \quad &\approx \quad d(u_N + u_E + u_S + u_W - 4u_{i,j} \\ &= \quad 4d * \left(\tfrac{1}{4}(u_N + u_E + u_S + u_W) - u_{i,j}\right), \end{aligned} \tag{5.11}$$

where $u_{i,j}$ represents the population at coordinates (i, j).

Many two-dimensional models use a square grid for simplicity reasons, and we would like to consider an arrangement of patches that perfectly tessellate a given geometric area. It is a known result that the only regular polygons that will tessellate are equilateral triangles, squares, and hexagons as the measurements of their interior angles are a factor of 360°, the number of degrees in a circle. (See proof: http://mathandmultimedia.com/2011/06/04/regular-tessellations/) For convenience and more flexibility in geometric arrangement, I will assume that the patches are laid

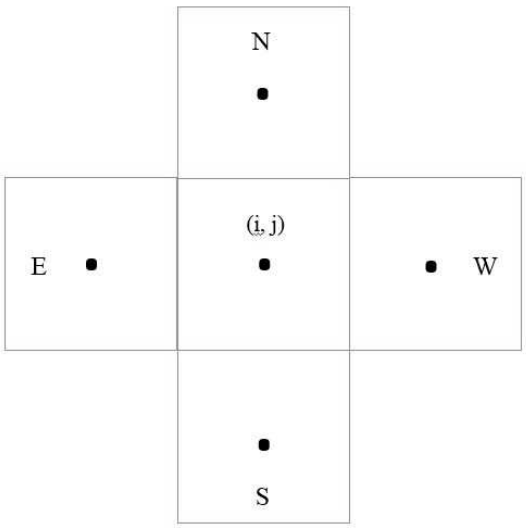


Figure 14. A patch with four neighbors

out in tessellated grid of identical hexagonal patches. Each patch would have exactly six neighbors as in Figure 15, and each patch would receive one-sixth of the deer migrating out of each neighboring patch.

$$\begin{aligned} \nabla^2 u = \left(\tfrac{\partial^2}{\partial x^2} + \tfrac{\partial^2}{\partial y^2}\right)u \quad &\approx \quad d\left(\tfrac{2}{3}(u_N + u_{NE} + u_{SE} + u_S + u_{SW} + u_{NW}) - 4u_{i,j}\right) \\ &= \quad 4d * \left(\tfrac{1}{6}(u_N + u_{NE} + u_{SE} + u_S + u_{SW} + u_{NW}) - u_{i,j}\right). \end{aligned} \tag{5.12}$$

When we consider dispersion in two dimensions, we will have a similar issue at the boundary as in the one-dimensional case. Hence, we may either consider an unbounded approach, or we may make similar adjustments to the patches on the boundary and their immediate neighbors. For the unbounded approach, the multi-patch model is given by:

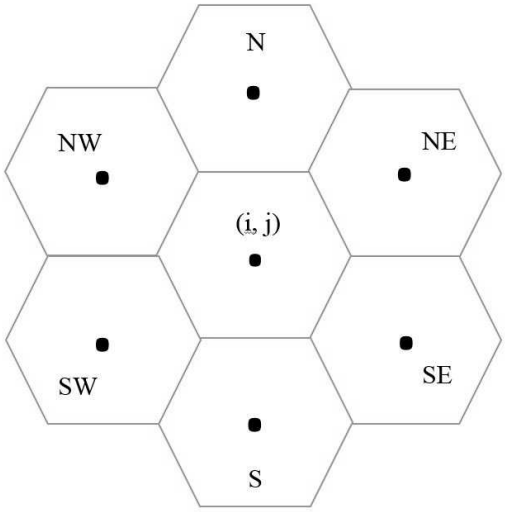


Figure 15. A patch with six neighbors

$$
\begin{aligned}
\frac{dD_S^{[i]}}{dt} &= \lambda_D D_N^{[i]} - \frac{\beta_D M_I^{[i]} D_S^{[i]}}{D_N^{[i]}} - ((\mu_D + \rho) + d_S + \mu_{2D} D_N^{[i]}) D_S^{[i]} - \\
&\quad D_S^{[i]} d_S + \tfrac{d_S}{n} \int_0^\infty g(z) e^{-\delta_S z} \sum_{j=1}^{n} D_S^{[j]}(t-z) dz \\
\frac{dD_I^{[i]}}{dt} &= \frac{\beta_D M_I^{[i]} D_S^{[i]}}{D_N^{[i]}} - ((\mu_D + \rho) + \gamma_D + d_I + \mu_{2D} D_N^{[i]}) D_I^{[i]} - \\
&\quad D_I^{[i]} d_I + \tfrac{d_I}{n} \int_0^\infty g(z) e^{-\delta_I z} \sum_{j=1}^{n} D_I^{[j]}(t-z) dz \\
\frac{dM_S^{[i]}}{dt} &= \lambda_M M_N^{[i]} - \frac{\beta_M D_I^{[i]} M_S^{[i]}}{D_N^{[i]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[i]}) M_S^{[i]} \\
\frac{dM_I^{[i]}}{dt} &= \frac{\beta_M D_I^{[i]} M_S^{[i]}}{D_N^{[i]}} - ((\mu_M + \sigma) + \mu_{2M} M_N^{[i]}) M_I^{[i]}
\end{aligned}
\tag{5.13}
$$

for $i = 1, ..., m$, where m is the number of patches and $n = 4$ or $n = 6$ is the number

of neighboring patches of patch i. It may also prove interesting to study what is the effect on the system and resulting model as n is changed (and perhaps increases to infinity).

When we consider dispersion in two dimensions, we will have a similar issue at the boundary as in the one-dimensional case. Hence, we may either consider an unbounded approach, or we may make similar adjustments to the patches on the boundary and their immediate neighbors.

Final Comments and Conclusions

Though mathematics and statistics can be extremely useful tools, we can rarely expect results using those tools to perfectly model the infinitely complex world we live in. However, that does not make the tools worthless. We are still able to make reasonably good approximations of complex behaviors, or at least break ground for the next person. Using SaTScan, we can see which areas in Missouri are most affected by HD. This may help lead efforts to control the disease to those specific parts of the state, and thus hopefully making any control measures more effective. We are also able to note trends in time as well, both in a small scale and a large scale. We learn where incidents are becoming more prevalent now and also observe cycles of large-scale outbreaks. Through parameter estimations, we gain insight on the dynamics of a host-vector disease system. These estimates can be used in future studies of these types of systems. They can also be the basis for use in a multi-patch model. Combining these values with an ODE model, we can predict a basic reproduction number, R_0. With this value, we learn two things. First, we are able to determine what conditions may lead to elimination of the disease ($R_0 < 1$) or may lead to its

spread ($R_0 > 1$). Second, in this latter case when the disease reaches its epidemic level, we can use its value to determine how severe the outbreak may be. Finally, analysis of this ODE model makes a contribution to mathematical studies as a whole. Through stability analysis and existence of equilibria, we provide a usable model for future studies.

Because of the real-world movement of deer in multiple directions, we note that one of the primary limitations of this study is the behavior in the multi-patch system, where migrating individuals leave one patch and eventually enter a neighboring patch. We also did not consider a delay in the traveling time. Holt [20] and Weisser et al. [60] extended their results to a system of multiple patches joined through a pool of dispersing individuals. Moreover, the proposed model (3.1) does not include the effect of predators on the population of white-tailed deer. As a prey species, deer are linked with local predators. In Missouri, the coyote is one such predator. Some coyote predator studies have been done, but these are admittedly outdated. However, deer make up a portion of a coyote's diet and that large increases or decreases in predator populations may influence deer mortality rates [12]. Finally, our model assumed only one vector for the transmission of HD. With the species richness of the *Culicoides* genus, we may reasonably expect more and different interaction rates and different levels of control efficacy [43]. We also note that weather has an effect on both the midge population and the life cycle of the HD virus [50, 35]. Midge populations thrive in damper areas, and in 2012, there was an above average amount of rain in the late winter/early spring, filling ponds and other water bodies in Missouri [12]. In addition, record warm temperatures in that spring and summer may cause

midges to become more active sooner than normal [12]. Next, the high temperatures caused water sources to dry up, and not only did the resulting mud flats become ideal breeding areas for subsequent generations of midges, but also caused deer to visit water sources more frequently due to lower water content in the plants they ate as part of their diet. These same high temperatures also cause female midges to lay more eggs, and Wittmann et al. also revealed that higher temperatures decrease the extrinsic incubation period of the HD virus within the midges [62]. Thus, the virus develops faster and allows a midge to infect more deer during its life span. None of these factors have been considered in the model (3.1). Instead, the main focus has been on migration effects of deer population on overall HD dynamics within a patch.

The above mentioned limitations demand model extensions to study the effectiveness of control and preventive strategies. Deer species are important members of the ecosystem as they feed on brush and grass in a given area and keep them in check. In conclusion, the present work is the first step towards inclusion of migration effects of deer population modeling of HD dynamics. The R_0 expression provides insights into the effects of deer movement on the spread of disease.

Future Work

From this point, there are some meaningful and important next steps to consider. First, we can continue the analysis of the two-patch and multi-patch models. The hope here is that results similar to those of the single, migration patch model: the model is stable, it has equilibrium values, and has a computable value for R_0. It would also be interesting to determine what happens to the dynamics of the system if the number of patches increases beyond two or four or six. We could have every

county be its own patch, or, if we allow the number of patches to approach infinity, we can (theoretically) study the dynamics at every point in the state. With such a model in place, and specific migration data, we could perform an in-depth analysis of the effects of migration on the dynamics of HD.

APPENDIX A:

Data Summary

The following tables summarize the data provided by the MDC. For the purpose of this paper, Missouri is divided into eight geographical regions as in Figure 16. For each region of Missouri, each set of tables gives yearly information on (a), the frequency of HD incidents, (b), the prevalence of HD incidents per 10,000 deer, (c), the estimated deer population, and (d), the number of deer harvested.

Figure 16. A map of Missouri divided into eight geographic regions

Missouri

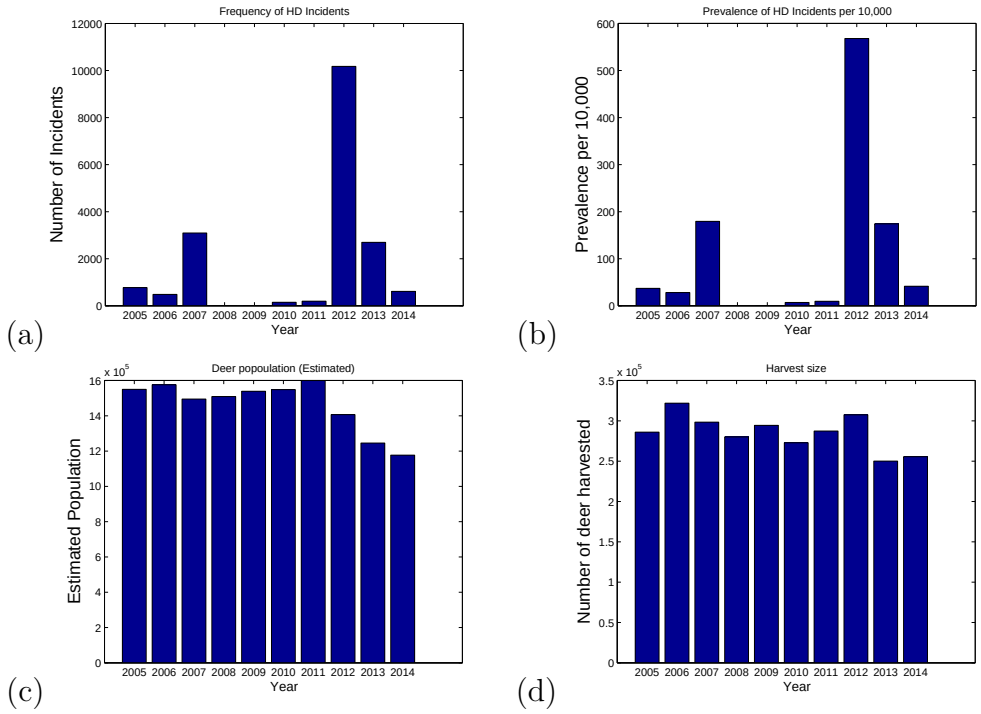


Figure 17. A summary of the data from Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Central

This region consists of fifteen counties in central Missouri: Audrain. Boone, Callaway, Camden, Cole, Cooper, Gasconade, Howard, Maries, Miller, Moniteau, Montgomery, Morgan, Osage, Saline.

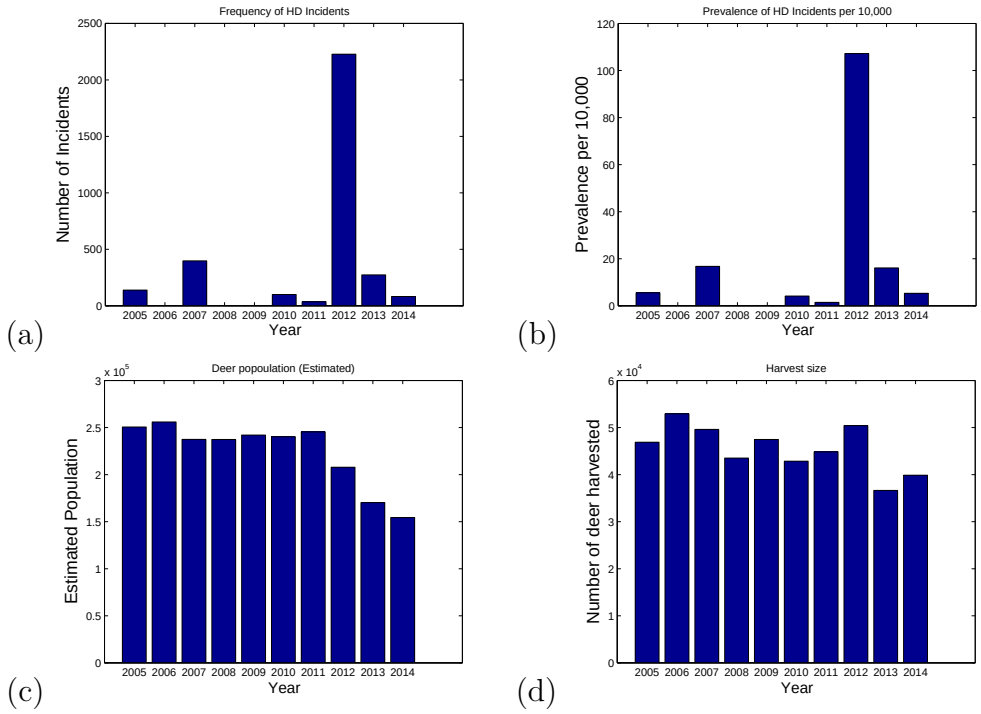


Figure 18. A summary of the data from central Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Northeast

This region consists of fifteen counties in northeastern Missouri: Adair, Clark, Knox, Lewis, Macon, Marion, Monroe, Pike, Putnam, Ralls, Randolf, Schuyler, Scotland, Shelby, Sullivan.

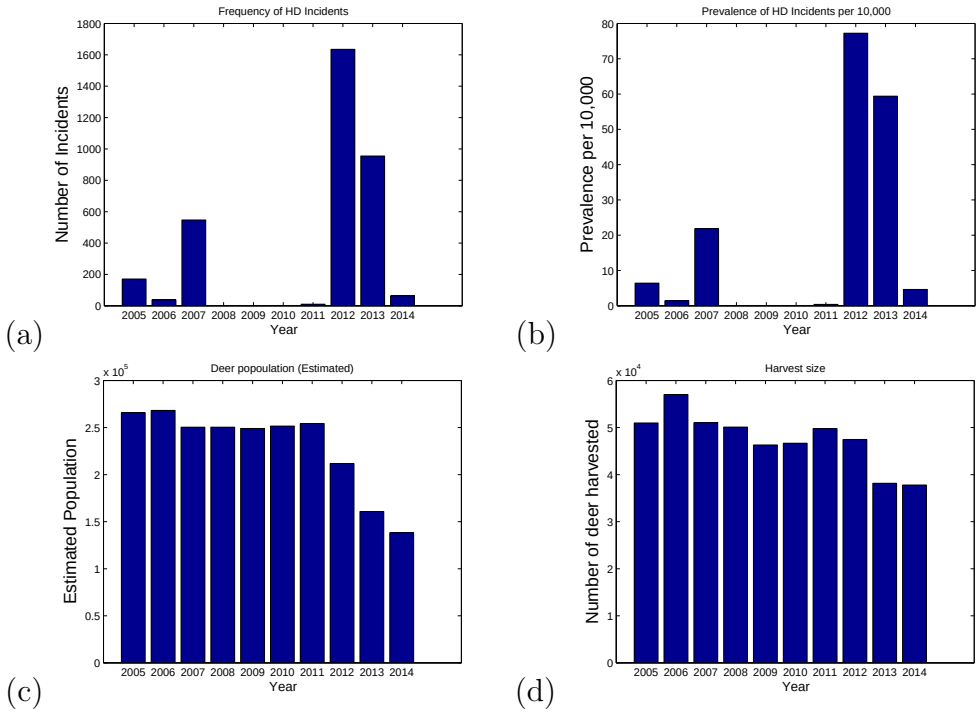


Figure 19. A summary of the data from northeast Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Northwest

This region consists of nineteen counties in northwest Missouri: Andrew, Atchinson, Caldwell, Carroll, Chariton, Clinton, Daviess, DeKalb, Gentry, Grundy, Harrison, Holt, Linn, Livingston, Mercer, Nodaway, Platte, Ray, Worth.

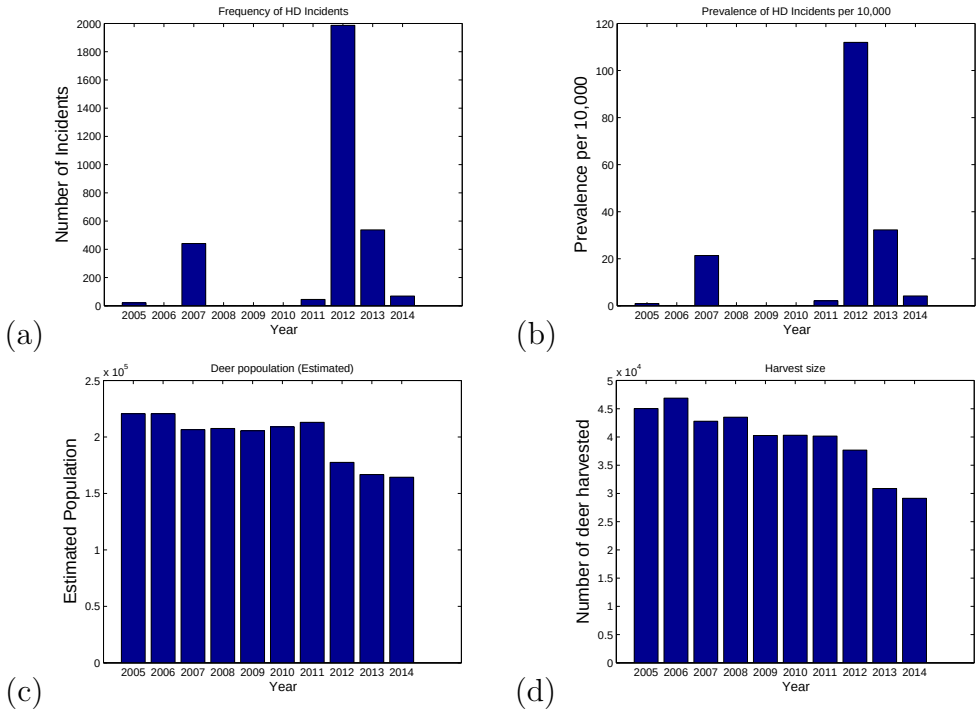


Figure 20. A summary of the data from northwest Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

St. Louis

This region consists of eight counties in eastern Missouri: Crawford, Franklin, Jefferson, Lincoln, St. Charles, St. Louis, Warren, Washington.

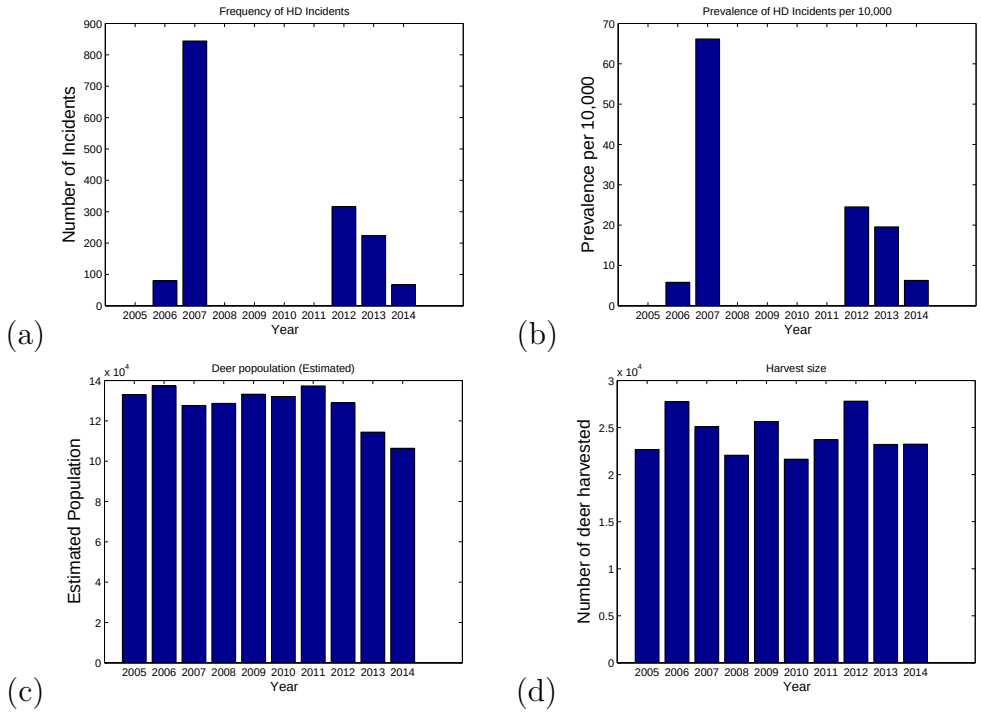


Figure 21. A summary of the data from east Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Kansas City

This region consists of twelve counties in western Missouri: Bates, Benton, Cass, Clay, Henry, Jackson, Johnson, Lafayette, Pettis, Platte, St. Claire, Vernon.

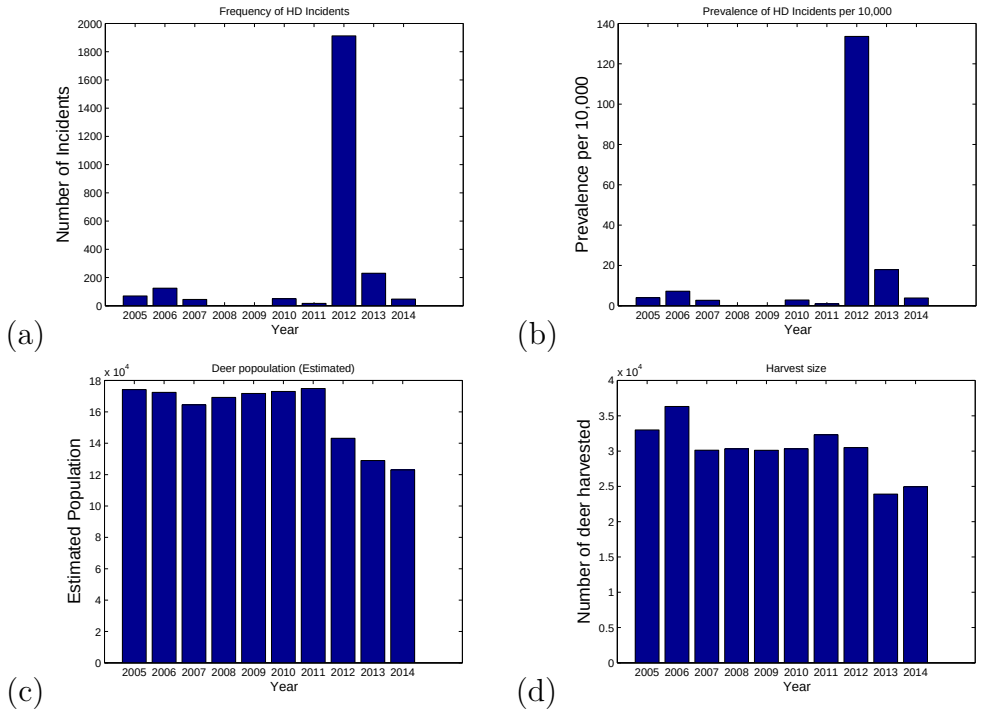


Figure 22. A summary of the data from west Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Southeast

This region consists of sixteen counties in southeast Missouri: Bollinger, Butler, Cape Girardeau, Dunklin, Iron, Madison, Mississippi, New Madrid, Pemiscot, Perry, Reynolds, Scott, St. Francois, Ste. Genevieve, Stoddard, Wayne.

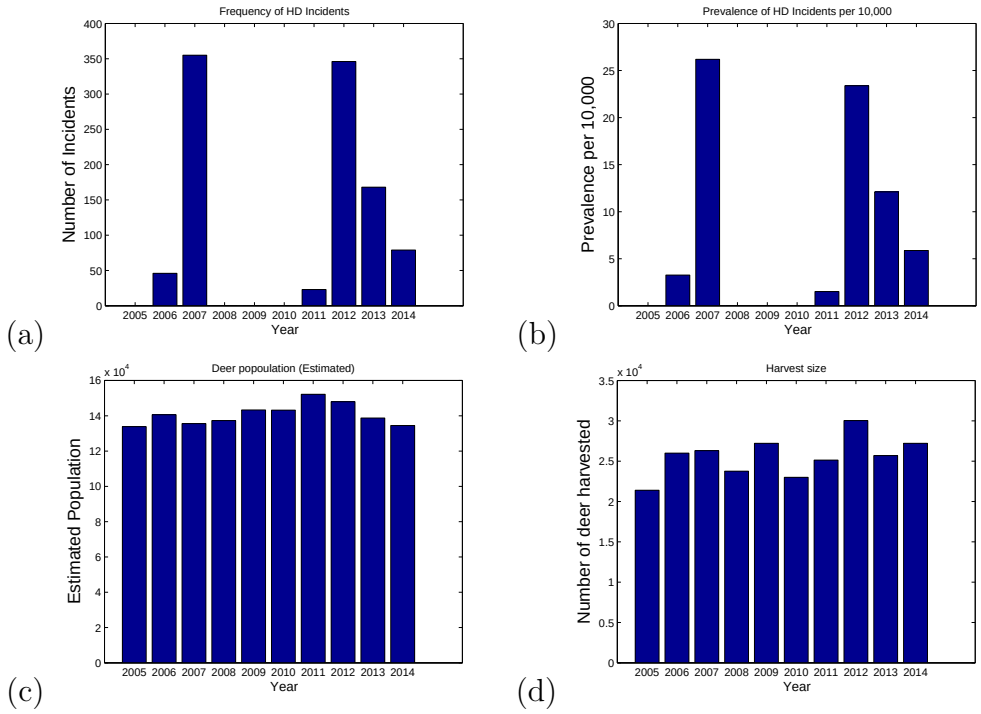


Figure 23. A summary of the data from southeast Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Ozark

This region consists of twelve counties in southern Missouri: Carter, Dent, Douglas, Howell, Oregon, Ozark, Phelps, Pulaski, Ripley, Shannon, Texas, Wright.

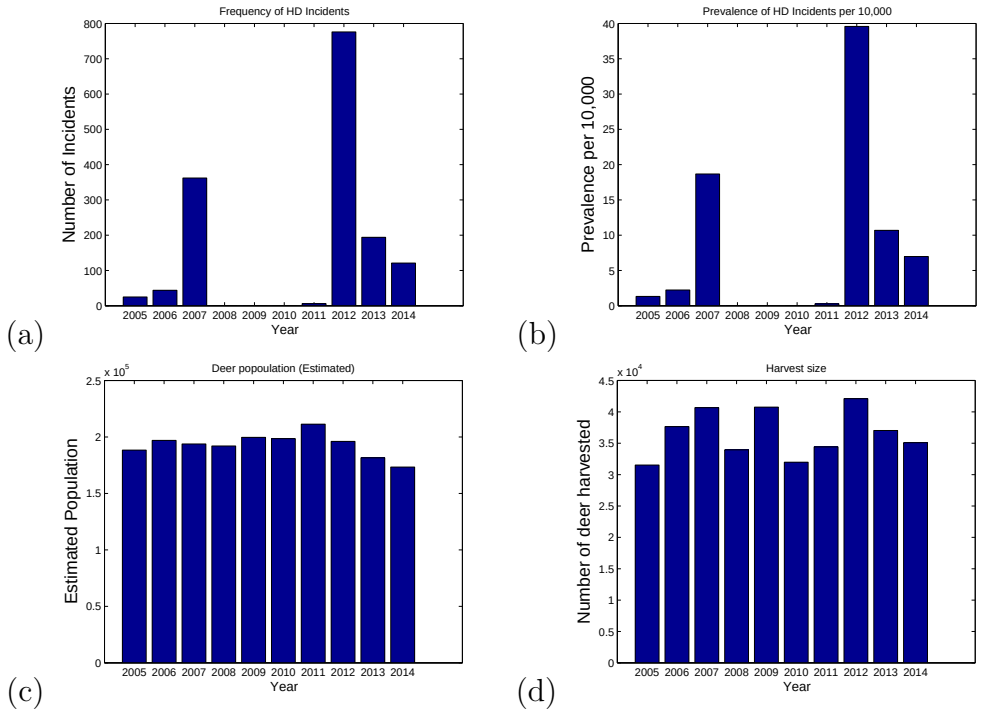


Figure 24. A summary of the data from south Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Southwest

This region consists of seventeen counties in southwest Missouri: Barry, Barton, Cedar, Christian, Dade, Dallas, Greene, Hickory, Jasper, Laclede, Lawrence, McDonald, Newton, Polk, Stone, Taney, Webster.

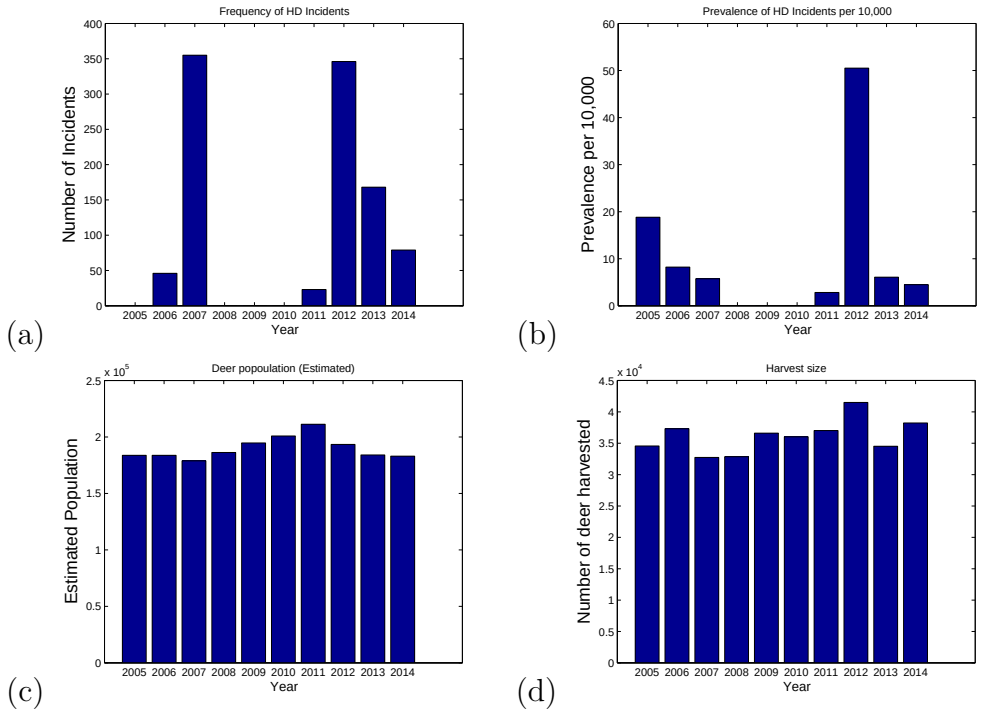


Figure 25. A summary of the data from southwest Missouri. This includes (a) number of HD incidents, (b) prevalence of HD incidents per 10,000 deer, (c) estimated population, and (d) estimated harvest for each year between 2005 and 2014.

Overall, the highest populations of deer occur in the central, northeast, and northwest part of the state, and this is consistent over all years. Estimated populations in these regions were between 2 and 2.5 million, but those dropped to about 1.5 million during and after 2012. Harvest percentages were consistent across the state as well. Most regions harvested between 18% to 20% each year, though in 2005, the harvest was as low as 16% in the south and southeast, and in 2013 and 2014 was as high as 27% in the central and northeast regions. The two largest outbreaks of HD

in the last decade have occurred in 2007 and 2012. The number of incidents in 2007 occurred mostly in the southern and eastern parts of the state. In 2012 the severity was statewide, with all counties reporting at least one incidence. Most of these occurred in the central, southern, and Kansas City regions. During the non-outbreak years, the prevalence per 10,000 is fairly consistent overall all regions, though in 2013, the northeast region had levels three to four times the rest of the state.

APPENDIX B:

MATLAB CODES

The following two sets of code were used in numerically simulating the steady state solutions of the migration model as part of Chapter 3.

The first code is the model.

```
function dy= Chain_trick_model(t,y,p)

dy = zeros(6,1);

betaD=p(1);

lambdaD=p(2);

rho=p(3);

muD=p(4);

mu2D=p(5);

dS=p(6);

dI=p(7);

gammaD=p(8);

betaM=p(9);

lambdaM=p(10);

sigma=p(11);

muM=p(12);

mu2M=p(13);

alpha=p(14);

deltaS=p(15);

deltaI=p(16);


% The following is the chain trick version of the model
```

```
dy(1) = lambdaD*(y(1)+y(2)) - (betaD*y(4)*y(1))/(y(1)+y(2))
   - (muD+rho+dS+mu2D*(y(1)+y(2)))*y(1) + y(5);
dy(2) = (betaD*y(4)*y(1))/(y(1)+y(2)) - (muD+rho+gammaD+dI+
+ mu2D*(y(1)+y(2)))*y(2) + y(6);
dy(3) = lambdaM*(y(3)+y(4)) - (betaM*y(2)*y(3))/(y(1)+y(2))
         - (muM+sigma+mu2M*(y(3)+y(4)))*y(3);
dy(4) = (betaM*y(2)*y(3))/(y(1)+y(2)) - (muM
         +sigma+mu2M*(y(3)+y(4)))*y(4);
dy(5) = -(deltaS+alpha)*y(5) + dS*alpha*y(1);
dy(6) = -(deltaI+alpha)*y(6) + dI*alpha*y(2);
```

The second code runs the numerical simulation of the model with a proposed set of parameters and graphs the results.

```
clc
clear


betaD=10.1;
lambdaD=196.315;
rho=0.1958;
muD=0.4385;
mu2D=2.608*10^(-11);
dS=0.9;
```

```
dI=0.9;

gammaD=0.2243;

gtildedeltaI=0.6;

gtildedeltaS=0.6;

betaM=460.2656;

lambdaM=.6980;

sigma=0.01;

muM=0.9363;

mu2M=2.29*10^(-6);

alpha=0.9;

deltaS=0.1;

deltaI=0.1;

% The values for betaD, dS, dI, gtildedeltaI, gtildedeltaS,

% sigma, alpha, deltaS, and deltaI were assumed.  The

% remaining values came from the parameter estimates as

% part of Chapter 3 of my dissertation


DSstar=(alpha + (deltaS+alpha)*(lambdaD-muD-rho-dS))

        /((deltaS+alpha)*mu2D);

MSstar=(lambdaM-muM-sigma)/mu2M;

% These are the disease free equilbrium values


C=betaM*MSstar/DSstar;
```

```
E=muD+rho+gammaD+mu2D*DSstar;

F=lambdaM+sigma+2*mu2M*MSstar;

G=alpha+deltaI;

% These values simplify the expression for R0


R0=sqrt(((betaD*betaM*MSstar*G/DSstar)+(dI*alpha*F))/(E*F*G))


options = odeset('RelTol',1e-4,'AbsTol',

           [1e-5 1e-5 1e-5 1e-5 1e-5 1e-5]);

p=[betaD,lambdaD,rho,muD,mu2D,dS,dI,gammaD,betaM,lambdaM,

  sigma,muM,mu2M,alpha,deltaS,deltaI];

IC=[1500 772 100.000000 2000.0000 10.0 1];

% The initial conditions for the deer are scaled values

% of the actual populations.  Values for the last four

% conditions are assumed

ts=[0 50];

[T,Y] = ode45(@Chain_Trick_model,ts,IC,options,p);

plot(T,Y(:,1),'-',T,Y(:,5),':g','LineWidth',2.5)

xlabel('time','FontWeight','bold')

ylabel('Deer population','FontWeight','bold')

leg1=legend('$${D}_{S}$$','$$\overline{D}_{S}$$')

set(leg1,'Interpreter','latex');

set(leg1,'FontSize',18);
```

```
figure
plot(T,Y(:,2),'-.r',T,Y(:,6),'--b','LineWidth',2.5)
xlabel('time','FontWeight','bold')
ylabel('Deer population','FontWeight','bold')
leg1=legend('$${D}_{I}$$','$$\overline{D}_{I}$$')
set(leg1,'Interpreter','latex');
set(leg1,'FontSize',18);
figure

plot(T,Y(:,3),'-',T,Y(:,4),'-.r','LineWidth',2.5)
xlabel('time','FontWeight','bold')
ylabel('Midge population','FontWeight','bold')
leg2=legend('$${M}_{S}$$','$${M}_{I}$$')
set(leg2,'Interpreter','latex');
set(leg2,'FontSize',18);
```

The following three sets of code were used in estimating the parameters as part of Chapter 4.

The first code is the simplified model.

```
function dy = Modelwithcurvefitting(t,y,muD,lambdaM,betaM,muM,mu2M)
timescale=10^3;
% the timescale is here to help MATLAB get through the
```

```
% computation


mu2D=2.608*10^(-11);

gammaD=.2243;

rho=.1958;

lambdaD=rho+muD+7.6*10^(-5);

% These values were derived directly from the MDC data


a11=1.021e+04;

b11=8.002;

c11=0.03872;

a12=3099;

b12=2.998;

c12=0.04318;

a13=2696;

b13=9;

c13=0.01103;


f1 = a11*exp(-((t-b11)/c11).^2) + a12*exp(-((t-b12)/c12).^2) +

     a13*exp(-((t-b13)/c13).^2);


a1 =2422;

b1 =0.0347;
```

```
c1 =2.756;

a2 =548;

b2 =0.4754;

c2 =2.758;


f2 = a1*sin(b1*t+c1) + a2*sin(b2*t+c2);


DI=f1+f2;

% These two functions were best-fit functions via MATLAB for

% outbreak years (f1) and nonoutbreak years (f2)


dy = zeros(3,1);


dy(1) = timescale*(lambdaD*y(1) -(muD+rho+(mu2D*y(1)))*y(1) - gammaD.*DI);

dy(2) = lambdaM*(y(2)+y(3)); - (betaM.*DI*y(2))/y(1) -

        (muM+mu2M*(y(2)+y(3)))*y(3);

dy(3) = (betaM.*DI*y(2))/y(1) - (muM+mu2M*(y(2)+y(3)))*y(3);
```

The second code generates a temporary set of estimated parameter values, then graphs the estimated solution curve (with the smallest SSE) against the population data values.

```
clear
clc

options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-5 1e-5]);

muD=.0171;
lambdaM=.00000052;
betaM=465;
muM=.98;
mu2M=.00092;
scale=1;
% these values were based on a previous running of this code

datatime=[1 2 3 4 5 6 7 8 9 10];
datapopulation=[1550105.95 1575757.002 1494703.004 1508661.607
    1539096.876 1548465.228 1599525.101 1406513.946 1245358.923 1177235.128];
% Actual deer population vales for 2005 - 2014

dataC=zeros(2,10);
dataC(1,:)=datatime;
dataC(2,:)=datapopulation;

iterxy=500
```

```
par=[muD;lambdaM;betaM;muM;mu2M]


LB = [0;0;0;0;0];
UB = [7.95*10^(-4);1.6;1500;2;1];
% upper bounds were based on trial and error


options2 =optimset('MaxFunEvals',8000,'MaxIter',5000)
[x,fval] =fmincon(@EstimDeer,par,[],[],[],[],LB,UB,[],options2);
SSE=fval
muD=x(1)
lambdaM=x(2)
betaM=x(3)
muM=x(4)
mu2M=x(5)



lambdaD=.1967;
mu2D=2.608*10^(-11);
gammaD=.2243;
rho=.1958;
alpha=lambdaD-muD-rho
K=alpha/mu2D
[T,Y] = ode45(@Modelwithcurvefitting,[1 10],[1550105.95 10000 100],
```

```
     options,muD,lambdaM,betaM,muM,mu2M);
% the first initial condition is the true deer population
% the second two are assumed values for the midges
plot(T,Y(:,1),'-b','linewidth',2)
hold on
plot(dataC(1,:),dataC(2,:),'ob')
set(0,'DefaultAxesFontSize', 14);
xlabel('time, t')
ylabel('D_N')
title('Time series of the model solution' ,'FontWeight','bold')
grid on
```

The third code computes the SSE of the current parameter values and passes this information back to the second code.

```
function error = EstimDeer(p)

muD=p(1)
lambdaM=p(2)
betaM=p(3)
muM=p(4)
```

```
mu2M=p(5)
scale=1;

penalty=1;

time=[1 2 3 4 5 6 7 8 9 10];
population=[1550105.95 1575757.002 1494703.004 1508661.607 1539096.876
     1548465.228 1599525.101 1406513.946 1245358.923 1177235.128];
% Actual deer population vales for 2005 - 2014

C=zeros(2,10);
C(1,:)=time;
C(2,:)=population;

options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-5 1e-5]);
[t,y] =ode45(@Modelwithcurvefitting,C(1,:),[1550105.95 10000 100],
     options,muD,lambdaM,betaM,muM,mu2M);
% the first initial condition is the true deer population
% the second two are assumed values for the midges
value = penalty*(y(:,1)-C(2,:)').^2;
error = sum(value);
```

REFERENCES

1. P. Auger, E. Benoît, *A prey-predator model in a multi-patch environment with different time scales*, J. Biol. Syst., 1: 187 - 197, 1993.
2. M. Bani-Yaghoub, R. Gautam, R. Ivanek, P. Van Den Driessche, Z. Shuai, *Reproduction numbers for infections with free-living pathogens growing in the environment*, Journal of Biological Dynamics. 6 (2): 923-940, 2012.
3. G. Baygents, M. Bani-Yaghoub, *A mathematical model to analyze spread of hemorrhagic disease in white-tailed deer population*, Journal of Applied Mathematics and Physics, 5(11): 2262 - 2282, 2017.
4. _____, *Cluster analysis of hemorrhagic disease in Missouri's white-tailed deer population: 2005 - 2013*, BMC Ecol., revision and pre-print.
5. G. Baygents, *Methods of parameter estimation in ODE models with small data sets*, pre-print.
6. J. Beringer, L. P. Hansen, D. E. Stallknecht. *An epizootic of hemorrhagic disease in white-tailed deer in Missouri*, J. Wildl. Dis. 36: 588591, 2000.
7. J. D. Caton, *The Antelope and Deer of America: A Comprehensive Scientific Treatise Upon the Natural History, Including the Characteristics, Habits, Affinities, and Capacity for Domestication of the Antilocapra and Cervidae of North America*, 2nd Edition, Forest and Stream Publishing Company, New York, 1877.
8. K. Dietz, *The estimation of the basic reproduction number for infectious diseases.* Statistical Methods in Medical Research, 2(1): 23 - 41, 1993.

9. P. Diggle, A. G. Chetwynd, R. Haggkvist, S. E. Morris 1995, *Second-order analysis of space-time clustering*, Stat. Methods Med. Res. 4: 124136, 1995.

10. W. E. Fitzgibbon, J. J. Morgan, G. B. Webb, *An Outbreak Vector-Host Epidemic Model With Spatial Structure: The 2015-2016 Zika Outbreak in Rio De Janeiro*, Theor. Biol. Med. Modell., 2017.

11. E. Flinn, J. Sumners, *Breaking Down the Hemorrhagic Disease Outbreak*, Missouri Conservationist, 74(7): 24-29. 2013a.

12. _____, *State of the States Deer Herd*, Missouri Conservationist, 74(8): 24-29. 2013b.

13. N. M. Foster, R. D. Breckon, A. J. Luedke, R. H. Jones, H. E. Metcalf, *Transmission of two strains of epizootic hemorrhagic disease virus in deer by Culicoides variipennis*, J. Wildl. Dis., 13(1): 9 - 16, 1997.

14. R. Gautam, I. Srinath, A. Clavijo, B. Szonyi, M. Bani-Yaghoub, S. Park, R. Ivanek, *Identifying areas of high risk of human exposure to Coccidioidomycosis in Texas using serology data from dogs*, Zoonoses Public Health, 60 (2):174-81, 2013.

15. J. K. Gaydos, J. M. Crum, W. R. Davidson, S. S. Cross, S. F. Owen, et al., *Epizootiology of an epizootic hemorrhagic disease outbreak in West Virginia*, J. Wildl. Dis. 40: 383393, 2004.

16. R. E. Hawkins, W. D. Klimstra, *A preliminary study of the social organization of white-tailed deer*, J. Wildl. Manage., 34 (2): 407 - 419, 1970.

17. J. A. P. Heesterbeek, K. Dietz, K. *The concept of R_0 in epidemic theory*, Statistica Neerlandica, 50: 89110, 1996.

18. J. A. P. Heesterbeek, M. G. Roberts, *The type-reproduction number T in models for infectious disease control*, Math. Biosci., 206: 310, 2007.

19. G. Hoff, D. O. Trainer, *Hemorrhagic disease in wild ruminants*, In: Infectious Diseases of Wild Mammals, Ames: Iowa State University Press, 45-53, 1981.

20. R. D. Holt, *Spatial heterogeneity, indirect interaction, and the the coexistence of prey species*, Am. Nat., 124:377 - 406, 1984.

21. A. Hurford, D. Cownden, T. Day, *Next-generation tools for evolutionary invasion analyses*, J. R. Soc. Interface, 7: 561571, 2010.

22. S. Hwang, *Extending spatial hot spot detection techniques to temporal dimensions*, In: Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS, University of Glamorgan, Wales, UK, September, pp. 58, 2005.

23. V. Jansen, L. Lloyd, *Local stability analysis of spatially homogeneous solutions of multi-patch systems*, Journal of Mathematical Biology, 41: 232, 2000.

24. D. W. Jordan, P. Smith, *Nonlinear Ordinary Differential Equations: An Introduction to Dynamical Systems*, Oxford University Press, Oxford, 1999.

25. P. Klepac, M. G. Neubert, P. Van Den Driessche, *Dispersal delays, predator-prey stability, and the paradox of enrichment*, Theoretical Population Biology, 7(4): 436-444, 2007.

26. E. Kouokam, P. Auger, H. Hbid, M. Tchuente, *Effect of the number of patches in a multi-patch SIRS model with fast migration on the basic reproduction rate*, Acta Biotheor., 56: 75, 2008.

27. Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*. Academic Press, Inc., San Diego, 1993.

28. M. Kulldorff, *A spatial scan statistic*, Communications in Statistics: Theory and Methods, 26:1481-1496, 1997.

29. _____, SaTScan User Guide for version 9.4, http://www. satscan.org/, 2006.

30. M. Kulldorff, R. Heffernan, J. Hartman, R. M. Assuno, F. Mostashari, *A space-time permutation scan statistic for the early detection of disease outbreaks*, PLoS Medicine. 2: 216-224. 2005.

31. M. Kulldorff, W. F. Athas, E. J. Feuer, B. A. Miller, C. R. Key, *Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico*, Am. J. Public Health. 88, 13771380, 1998.

32. A. Leblond, A. Sandoz, G. Lefebvre, H. Zeller, D. J. Bicout, *Remote sensing based identification of environmental risk factors associated with West Nile disease in horses in Camargue, France*, Prev. Vet. Med. 79(1): 2031, 2007.

33. Y. Lou, X-Q. Zhao, *The periodic Ross-Macdonald model with diffusion and advection*, Applicable Analysis, 89:7, 1067-1089, 2009.

34. G. Macdonald, *The Epidemiology and Control of Malaria*, Oxford University Press, London, 1957.

35. P. S. Mellor, J. Boorman, M. Baylis, *Culicoides biting midges: their role as arbovirus vectors*, Annu. Rev. Entomol., 45:307-340, 2000.

36. S. E. Mirghani, B. Y. Nour, S. M. Bushra, I. M. Elhassan, R. W. Snow, A. M. Noor. *The spatialtemporal clustering of plasmodium falciparum infection over eleven years in Gezira State, The Sudan*, Malar. J. 9: 172, 2010.

37. V. F. Nettles, W. R. Davidson, D. E. Stallknecht. *Surveillance for hemorrhagic disease in white-tailed deer and other wild ruminants, 1980-1989*, In: Proceedings

of Annual Conference of Southeastern Association of Fish and Wildlife Agencies, 138-146, 1992.

38. V. F. Nettles, D. E. Stallknecht, *History and progress in the study of hemorrhagic disease of deer*, T. N. Am. Wildl. Nat. Res. 57, 499-516, 1992.
39. M. G. Neubert, P. Klepac, P. Van Den Driessche, *Stabilizing Dispersal Delays in Predator-Prey Metapopulations Models*, Theoretical Population Biology, 61(3): 339 - 347, 2001.
40. G. Ngwa, W. Shu, *A Mathematical Model for Endemic Malaria with Variable Human and Mosquito Populations*, United Nations Educational Scientific and Cultural Organization and International Atomic Energy Agency, The Abdus Salam International Centre for Theoretical Physics, Miramare-Trieste, 1999.
41. M. Norstrm M, D. U. Pfeiffer, J. Jarp, *A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds*, Prev. Vet. Med. 47(200): 107119, 2000.
42. A. W. Park, K. Magori, B. A. White, D. E. Stallknecht, *When more transmission equals less disease: reconciling the disconnect between disease hotspots and parasite transmission*, PLoS ONE 8(4): e61501, 2013.
43. A. W. Park, C. A. Cleveland, T. A. Dallas, J. L. Corn, *Vector species richness increases haemorrhagic disease prevalence through functional diversity modulating the duration of seasonal transmission*, Parasitology, 143(7):874-9, 2016.
44. L. Perko, 2001. *Differential Equations and Dynamical Systems*, 3rd ed. Springer, New York, 2001.
45. Purdue University, Extension E-250-W, *Biting Midges: Biology and Public*

Health Risk,

https://extension.entm.purdue.edu/publichealth/insects/bitingmidge.html

46. R. Ross, *The Prevention of Malaria*, John Murray, London, 1911.
47. E. J. Routh, *A Treatise on the Stability of a Given State of Motion: Particularly Steady Motion*, Macmillan, 1877.
48. L. Sedda, H. E. Brown, B. V. Purse, L. Burgin, J. Gloster, D. J. Rogers, *A New Algorithm Quantifies the Roles of Wind and Midge Flight Activity in the Bluetongue Epizootic in Northwest Europe*, Proceedings: Biological Sciences 279 (1737). Royal Society: 235462, 2012. http://www.jstor.org.proxy.library.umkc.edu/stable/41549546
49. C. W. Severinghaus, E. L. Cheatum, Life and times of the white-tailed deer, In: The deer of North America, The Stackpole Co., Harrisburg, PA, 1956.
50. J. M. Sleeman, J. E. Howell, W. M. Knox, P. J. Stenger, *Incidence of hemorrhagic disease in white-tailed deer is associated with winter and summer climactic conditions*, EcoHealth 6, 11 - 15, 2009.
51. C. Song, M. Kulldorff, 2003. *Power evaluation of disease clustering tests*, Int. J. Health Geogr. 2: 9, 2003.
52. D. E. Stallknecht, E. W. Howerth, *Epidemiology of bluetongue and epizootic hemorrhagic disease in wildlife: surveillance methods*, Vet. Ital. 40(3): 203-207, 2004.
53. D. E. Stallknecht, M. P. Luttrell, K. E. Smith, V. F. Nettles, *Hemorrhagic disease in white-tailed deer in Texas: a case for enzootic stability*, J. Wildl. Dis. 32(4): 695-700, 1996.

54. G. Stevens, B. McCluskey, A. King, E. OHearn, G. Mayr, *Review of the 2012 epizootic hemorrhagic disease outbreak in domestic ruminants in the United States*, PLoS ONE. 10(8): e0133359, 2015.

55. D. O. Trainer, *Epizootic hemorrhagic disease of deer*, J. Wildl. Manage., 28(2): 377-381, 1964.

56. P. Van Den Driessche, J. Watmough, *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission*, Math. Biosci., 180: 2948, 2002.

57. L. J. Verme, *Movements of white-tailed deer in upper Michigan*, J. Wildl. Manage., 37(4): 545 - 552, 1973.

58. W. Wang, X-Q. Zhao, *A nonlocal and time-delayed reaction diffusion model of dengue transmission*, SIAM J. Appl. Math., 71(1): 147 - 168, 2011.

59. M. P. Ward, *Clustering of reported cases of leptospirosis among dogs in the United States and Canada*, Prev. Vet. Med. 56(2002): 215226, 2002.

60. W. E. Weisser, V. A. A. Jansen, M. P. Hassell, *The effects of a pool of dispersers on host-parasitoid systems*, J. Theor. Biol., 189:413-425, 1997.

61. L. Wieser-Schimpf W. C. Wilson, D. D. French, A. Baham, L. D. Foil, *Bluetongue virus in sheep and cattle and Culicoides variipennis and C. stellifer (Diptera: Ceratopogonidae) in Louisiana*, J. Med. Entomol 30(4): 719-724, 1993.

62. E. J. Wittmann, P. S. Mellor, M. Baylis, *Effect of temperature on the transmission of orbiviruses by the biting midge, Culicoides sonorensis*, Medical and Veterinary Entomology, 16: 147156, 2002.

63. B. Xu, M. Madden, D. E. Stallknecht, T. Hodler, K. Parker, *Spatial and spatial-*

temporal clustering analysis of hemorrhagic diseases in white-tailed deer in the southeastern USA: 1980 2003, Prev. Vet. Med. 106 (2012): 339 347, 2012.

VITA

Gerald Walker Baygents (Gerry) has been an on-and-off student of mathematics since 1989. He received his B.S. in Mathematics (magna cum laude) at Georgia Southern University in 1993 and his M.S. in Mathematics (cum laude) at the University of South Carolina. While at UMKC, he has given presentations about his research, has helped to coordinate the Math Department's Integration Bee, and has chaperoned UMKC's Missouri Collegiate Mathematics Competition team. At the time of this writing, he has one published paper, *A Mathematical Model to Analyze Spread of Hemorrhagic Disease in White-tailed Deer Population*, in the Journal of Applied Mathematics and Physics.

In addition to his mathematical studies, he has been a tutor and teacher at both the high school and college levels. Some of his other jobs have included selling and demonstrating board games, volunteering and interning at the Columbia Zoo in Columbia, SC, interning at Wolf Park in Lafayette, IN, and working at the front desk and stables at Mammoth Lodge in Yellowstone National Park. In addition, he plays French horn in community music groups. He also enjoys road trips on his motorcycle, owns too many board games, and wishes he played disc golf better.