Climate and Infection-age on West Nile Virus Transmission

by

Marina Mancuso

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2023 by the
Graduate Supervisory Committee:

Fabio Milner, Chair
Yang Kuang
Eric Kostelich
Carrie Manore
Steffen Eikenberry

ARIZONA STATE UNIVERSITY

December 2023

ABSTRACT

Climate change is one of the most pressing issues affecting the world today. An important negative impact of climate change is on the transmission of mosquito-borne diseases (MBDs), such as West Nile Virus (WNV). Climate is known to influence vector and host demography as well as MBD transmission. This dissertation addresses some key questions of how vector and host demography impact MBD dynamics, exemplified herein by WNV, and how expected and likely climate change scenarios will affect demographic and epidemiological processes of MBD (here, WNV) transmission. First, a data fusion method is developed that connects non-autonomous logistic model parameters to mosquito time series data. This method captures the inter-annual and intra-seasonal variation of mosquito populations within a geographical location. Next, a three-population WNV model for mosquito vectors, bird hosts, and human hosts, with infection-age structure for the vector and bird host populations, is introduced. Sensitivity analysis uncovers which parameter changes have the most influence on WNV outbreak variability. Finally, the WNV model is extended to include the non-autonomous population model and temperature-dependent processes. Model parameterization using historical temperature and human WNV case data from the Greater Toronto Area (GTA) is performed and the results are then used to analyze possible future WNV (as an example for many other VBDs) dynamics under two climate change scenarios. The results from these scenarios suggest that WNV risk for the GTA will substantially increase as average annual temperature increases due to climate change, even under the most conservative assumptions. This demonstrates the importance of ensuring that the warming of the planet is limited as much as possible.

DEDICATION

This dissertation is dedicated to my late grandparents, Anthony and Rosalie Mancuso.

ACKNOWLEDGMENTS

I would first like to express my gratitude to my thesis advisor, Dr. Fabio Milner, for supporting my academic and research endeavors throughout my time here at ASU. As someone who did not have as strong of a mathematical background as most of my peers in the graduate program, Dr. Milner could still recognize my capability as a researcher in applied mathematics, even when I could not see it myself. I am looking forward to continuing to work together on future projects.

A very special thank you to Drs. Carrie Manore and Kaitlyn Martinez for their mentorship during my internship at Los Alamos National Laboratory. I am incredibly grateful to have had the opportunity to be a part of the CIMMID team, and even more grateful to have received 3 semesters of academic support. I am excited to continue my work at LANL as a postdoc, and continue working under your guidance.

I would like to acknowledge the remaining members of my committee: Drs. Yang Kuang, Eric Kostelich, and Steffen Eikenberry for their guidance. Each of you have pushed me to become a better researcher. I would like to thank Mrs. Sandra Matteucci for sponsoring my education for two years as an ARCS scholar.

Finally, I would like to thank my parents, my fiancé, Brandon Chandra, the current and former SoMSS graduate student cohort, and my AZ family for their endless support and encouragement. Thanks to you all, I have had an amazing past 5 years as a graduate student.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES xii

CHAPTER

1 INTRODUCTION 1
  - 1.1 Overview of Mosquito-Borne Diseases (MBDs) 1
  - 1.2 Environmental Factors on Mosquito Biology and MBDs 1
  - 1.3 Overview of the West Nile Virus (WNV) Transmission Cycle 3
    - 1.3.1 Transmission from Vectors to Birds 3
    - 1.3.2 Transmission from Birds to Vectors 4
  - 1.4 Mathematical Modeling Paradigms for MBDs 5
  - 1.5 Climate Integrated Model for Mosquito-Borne Infectious Diseases (CIMMID) 7
    - 1.5.1 Intergovernmental Panel on Climate Change (IPCC) Scenarios 8
    - 1.5.2 The Relevance of Data Fusion Techniques 10
  - 1.6 Scope of the Dissertation 11

2 A DATA FUSION FRAMEWORK FOR MODELING MOSQUITO POPULATIONS WITH time-dependent LOGISTIC GROWTH PARAMETERS 14
  - 2.1 Motivation for Data Fusion Framework 14
    - 2.1.1 Data Sources for the Greater Toronto Area 16
    - 2.1.2 Mosquito Process-Based Model (PBM) 18
    - 2.1.3 Connecting Mosquito PBM with Human Epidemiological Models 19
  - 2.2 The Logistic Model 20

- 2.3 The Non-Autonomous Logistic Model 22
- 2.4 Model Selection Procedure 24
- 2.5 Parameter Optimization 28
  - 2.5.1 Total Mosquito Population 29
  - 2.5.2 Active Mosquito Population 30
- 2.6 Sensitivity Analysis 31
- 2.7 Parameter Prediction From Weather Variables 33
- 2.8 Results of Model Selection 34
- 2.9 Parameter Fitting Results for Total Population 35
- 2.10 Parameter Fitting Results for Active Population 38
- 2.11 Results of Sensitivity Analysis 42
- 2.12 Results of Parameter Prediction From Weather Variables 44
- 2.13 Discussion 46

3 INFECTION-AGE DEPENDENT WEST NILE VIRUS MODEL 55
- 3.1 Mathematical Model 55
- 3.2 Integral Form 59
- 3.3 Existence and Uniqueness of Solutions 62
- 3.4 Basic Reproduction Number and Local Asymptotic Stability of Disease-Free Equilibrium 63
- 3.5 Choice of Infection-Age Dependent Parameter Functions 64
  - 3.5.1 Vector-to-Bird Transmission Probability 64
  - 3.5.2 Bird-to-Vector Transmission Probability 65
  - 3.5.3 Bird Recovery Rate 66
- 3.6 Sensitivity Analysis 68

- 3.6.1 Baseline Sensitivity Indices 70
- 3.6.2 Sensitivity Results for Infection Transmission and Recovery Parameters 71
- 3.6.3 Sensitivity Results for Demographic and Biting Parameters . 73
- 3.7 Discussion 74
- 4 A NON-AUTONOMOUS WEST NILE VIRUS MODEL 78
  - 4.1 Temperature-Dependent Processes on WNV Transmission 78
  - 4.2 Time-Dependent Processes on WNV Transmission 81
  - 4.3 Full Mathematical Model 83
  - 4.4 Data Sources for Total Populations 89
    - 4.4.1 Total Mosquito Population 89
    - 4.4.2 Total Bird Population 89
    - 4.4.3 Total Human Population 92
  - 4.5 Data Fitting and Simulations 92
    - 4.5.1 Fitting to Human Case Data 92
    - 4.5.2 Numerical Method 93
    - 4.5.3 Simulations from Parameter Fitting 95
  - 4.6 Climate Change Scenario Analysis 97
    - 4.6.1 Data Sources for Climate Change Projections 97
    - 4.6.2 Projections for Total Population Sizes 99
    - 4.6.3 Simulations of Climate Change Scenarios 102
  - 4.7 Discussion 104
- 5 CONCLUSION 108

REFERENCES 111

APPENDIX

- A JOURNAL PERMISSION 120
- B LINKAGE MODEL 122
- C SIMPLIFIED WNV MODEL 130
  - C.1 Method of Characteristics for Infected Birds 131
  - C.2 Integral Form of the Model 134
  - C.3 Proof of Existence and Uniqueness of the Model 138
  - C.4 Local Asymptotic Stability of Disease-Free Equilibrium 141
  - C.5 Forward Normalized Sensitivity Indices 144
- D FULL WNV MODEL 145
  - D.1 Parameter Values and Sources 146
  - D.2 Parameter Fittings for Bird Net Growth Rate 147
  - D.3 Numerical Implementation 148

LIST OF TABLES

Table Page

2.1 Candidate Models for the Model Selection Procedure. For Each Candidate Model We Fit a Subset of the Parameters of the Non-Autonomous Logistic Model (2.2). For Models A–D We Consider a Constant Carrying Capacity (i.e., $K_b = K$ and $K_s = 0$), While For Models E–K We Consider a Time-Dependent Carrying Capacity. 27

2.2 Akaike Information Criterion (AIC) Values For the Ten Candidate Models of Greater Toronto Area's Total Mosquito Population. The Fitting Season Refers to the Start and End Dates of Model Fitting for Each Year of Time Series Data. Each Fitting Season Lasts 154 Days, or 22 Weeks. The Four-Parameter Non-Autonomous Logistic Model (Model K) Produces the Lowest AIC Values for Each Fitting Season. . . 35

2.3 Akaike Information Criterion (AIC) Values From the Ten Candidate Models of Greater Toronto Area's Active Mosquito Population. "Start" Refers to the Initial Date of Model Fitting for Each Year of the Time Series Data. Model Fitting Lasts 365 Days, or One Full Year. The Four-Parameter Non-Autonomous Logistic Model (Model K) Produces the Lowest AIC Values for Each Fitting Season. 36

3.1 State Variables and Parameters of the Model (3.3). 56

3.2 Baseline Values and Ranges of Parameters in $\mathcal{R}_0$. 67

3.3 Forward Normalized Sensitivity Indices of Parameters of $\mathcal{R}_0$ at Baseline Values as Indicated in Table 3.2. Parameters are Ranked by Sensitivity Magnitude. We See That $\mathcal{R}_0$ is Most Sensitive to the Mosquito Mortality and Biting Rates. 71

4.1 Parameters of the Full Model (4.5). 87

4.2 Components of Time-Dependent Parameters. 88
4.3 Parameters of Infection-Age-Dependent Processes. 88
4.4 Breeding Bird Survey Estimates (BBS Values) of All Birds in the Greater Toronto Area (GTA) for Years 2010-2017 [11]. Total Birds Estimate the Entire Bird Population in the GTA as the Product of the BBS Bird Density and the GTA Region Area. Initial Relative Abundance (RA) is the Proportion of WNV-Competent Birds in the GTA from eBird Data After Applying the Gaussian Process Regressor [19, 62] on the First Day of the Mosquito Season. Initial Bird Population $N_b(0)$ is Found by Multiplying the Total Birds by the Initial RA. 91
4.5 Fitted Values of the Vector Per Capita Biting Rate on Humans α_h for the Greater Toronto Area for Years 2010-2017. The Parameter α_h was Fitted to the Yearly Estimated Human WNV Cases Based on Data from Public Health Ontario [61]. Residual Values Indicate the Difference Between the Cumulative New Human Infections from the Model and the Yearly Estimated Cases. 98
4.6 Fitted Parameter Values for Projected Vector and Bird Population Sizes Under Climate Change Scenarios RCP4.5 and RCP8.5. Note That “Start Day” Refers to the First Day of the Mosquito Season. 102

Table Page

B.1 Candidate Models for the Model Selection Procedure. Models A–D Consider a Constant Carrying Capacity, While Models E–K Consider a Time-Varying Carrying Capacity. Entries with a Solid Line ("——–") Indicate that the Parameter was not Fitted for that Model, and Rather was Held Constant at the Initial Value. For Entries with Two Lines, the Upper Line Refers to the Values Used for the Total Mosquito Population, and the Lower Line Refers to the Values Used for the Active Mosquito Population. The Initial K or K_b Value for Each Model is the Maximum Value of the Mosquito Process-Based Model Time-Series used for Each Season's Fitting. "Parameters Fitted" Refers to the Parameter Components which were Optimized for Each Season. "Initial" Columns Indicate Initial Values of the Respective Parameter for Each Year. "Range" Columns Refer to the Constraint Bounds ([Min, Max]) that Directed the Search Space for the Respective Parameter Optimization. 123

B.2 Fitted Parameter Values of the Model (2.2) for Greater Toronto Area's Total Mosquito Population from Years 2005–2019. For Each Year, the Mosquito Fitting Season Lasts from the Indicated Start Date Through End Date (mm/dd/yyyy Format). 124

B.3 Fitted Parameter Values of the Model's (2.2) Mosquito Fitting Season for Greater Toronto Area's Active Mosquito Population from Years 2005–2019. For Each Year, the Mosquito Fitting Season Lasts from the Indicated Start Date through End Date (mm/dd/yyyy Format). . . . 124

D.1 Values, Equations, and Sources for Parameters of the Full Model (4.5). 146

Table Page

D.2 Values, Equations, and Sources for Parameters of Infection-Age-Dependent Processes. 147

D.3 Fitted Parameter Values of Equation (4.4.3) for the WNV-Competent Bird Relative Abundance in the Greater Toronto Area for Years 2010–2017. For Each Year, the Bird Ecological Season Lasts from the Indicated Start Date Until the Start Date of the Following Year. 147

LIST OF FIGURES

Figure Page

1.1 Schematic of the CIMMID Process. Multiple Data Streams Related to Weather, Environmental Habitat, and Human and Mosquito Demographics Must Be Fused to Inform Mathematical and Statistical Models to Produce Forecasts of Future Mosquito-Borne Disease Risk. . . 9

2.1 High-Level Overview of the CIMMID Framework. 16

2.2 (a) Daily Mosquito Trap Captures, (b) Daily Mean Temperature, and (c) Daily Precipitation of the Greater Toronto Area For 2006-2019. 17

2.3 Fitting Results for the Greater Toronto Area's Total Mosquito Population for Years 2005–2019. (a) Simulation of Model Fittings (Blue Curves) and Mosquito Process-Based Model Output (Red Dots). (b) Relative Root Mean Squared Error for Each Years' Optimal Fit. 38

2.4 Heatmaps of the Normalized Root Mean Squared Error (RMSE) Values from Total Population Fittings with Respect to Candidate Start and End Dates for Years (a) 2011 and (b) 2014. Dark Purple Regions Denote the Candidate Start- and End-Date Combinations with the Lowest RRMSE Values. 39

2.5 Fitting Results for the Greater Toronto Area's Active Mosquito Population for Years 2005–2019. (a) Simulations of Model Fittings (Blue Curves) and Mosquito Process-Based Model Output (Red Dots). (b) Relative Root Mean Squared Error for Each Year's Optimal Fit. 40

2.6 Simulations (Curves) of the Active Mosquito Population By Start Date, Mosquito Process-Based Model Output (Red Dots) for Years (a) 2013 and (c) 2019. Normalized Mean Squared Error Values for Each Start Date's Fitting (May 1st–June 1st) for Years (b) 2013 and (d) 2019. 41

2.7 Sensitivity of Peak Magnitude to Parameters for the Total (Blue Bars) and Active (Orange Bars) Mosquito Populations in the Greater Toronto Area. Interaction-Order Sensitivities Were Found by Taking the Difference Between the Total-Order and First-Order Effects. 42
2.8 Sensitivity of Peak Timing to Parameters for the Total (Blue Bars) and Active (Orange Bars) Mosquito Populations in the Greater Toronto Area. Interaction-Order Sensitivities Were Found by Taking the Difference Between the Total-Order and First-Order Effects. 43
2.9 Scatter Plots for Models of (a) the Baseline Carrying Capacity K_b of the Total Population and the Yearly Total Precipitation (Exponential), and (b) the Per Capita Baseline Net Growth Rate r_b of the Active Population and Yearly Temperature Change (Linear). Pearson Correlation and R^2 Metrics are Displayed in the Top Left Corner of Each Panel. . . . 46
2.10 Year-by-Year Comparisons of Fitted Parameter (Blue Curves) and Predictive Exponential Model (Green Curves) Simulations for the Total Population. Red Dots Show the Mosquito Process-Based Model Time Series to Which the Non-Autonomous Logistic Model was Fitted. 47
2.11 Year-by-Year Comparisons of Fitted Parameter (Blue Curves) and Predictive Linear Model (Purple Curves) Simulations for the Active Population. Red Dots show the Mosquito Process-Based Model Time Series to Which the Non-Autonomous Logistic Model was Fitted. 48

Figure Page

2.12 Scatter Plots for Regression Models of (a) the Baseline Carrying Capacity K_b of the Total Population and the Yearly Total Precipitation, and (b) the Per Capita Baseline Net Growth Rate r_b of the Active Population and Yearly Temperature Change. The Labels of Each Point on the Scatter Plot Indicate the Corresponding Year of the Fitted Parameter and Weather Variable Combination. The Colors of the Year Labels Indicate the Quality of the Predictive Simulation Using the Regression Model (Black Lines): Blue Years Were "Good" Predictions, Yellow Years Were "Acceptable" Predictions, and Red Years Were "Poor" Predictions. 49

3.1 Schematic Diagram of the Model (3.3). The Red Dashed Lines Represent the Cross Transmission Between Populations. 58

3.2 Infection-Age-Dependent Parameter Functions. Data Fitting Results for (a) the Vector-to-Bird Transmission Probability Data From [63] to a Logistic Function and (b) the Transformed Bird-to-Vector Transmission Probability Data From [39, 63]. (c) Logistic Bird Per Capita Recovery Rate Over a Range of Values of the Scale Parameter k_2. The Values of the Maximum Recovery Rate and Mean Infectious Period (Inflection Point) Were Informed from Results in [39]. 68

3.3 The Basic Reproduction Number $\mathcal{R}_0$ (Blue Lines) and Sensitivity Response Curves (Red Lines) as a Function of Transmission and Recovery Parameters of $\mathcal{R}_0$. Starred Values Indicate the Baseline $\mathcal{R}_0$ and Normalized Forward Sensitivity Index. The Top Row Shows Parameters for the Transmission Probability of Vectors to Hosts, $\beta_{vb}(\tau)$. The Middle Row Shows Parameters for the Transmission Probability of Birds to Vectors, $\beta_{bv}(\tau)$. The Bottom Row Shows Parameters for the Bird Recovery Rate, $\gamma_b(\tau)$. 73

3.4 The Basic Reproduction Number $\mathcal{R}_0$ (Blue Lines) and Sensitivity Response Curves (Red Lines) as Functions of Demographic and Biting Parameters of $\mathcal{R}_0$. Starred Values Indicate the Baseline $\mathcal{R}_0$ and Normalized Forward Sensitivity Index. 75

4.1 Heatmap of the Per-Bite Probability of Vector-to-Host Transmission with Respect to Infection-Age and Temperature. Yellow Values Indicate the Highest Probability of Infection Transmission. 80

4.2 Results of Fitting the Bird Net Growth Rate $r_b(t)$ to the GPR Time Series of WNV-Competent Bird Relative Abundance in the GTA for 2010-2017. 83

4.3 Simulations of Infected Mosquito Vectors (Top Row), Birds (Middle Row), and Symptomatic Humans (Bottom Row) for the Greater Toronto Area for Years 2010-2017. The Red Dashed Line in the Bottom Row Shows the Reported Monthly Human Cases from Public Health Ontario [61]. 97

4.4 Projected Time Series for (a) Temperature and (b) Precipitation for an Average Year Between 2080-2100 Under Climate Scenarios RCP4.5 (Yellow) and RCP8.5 (Green) Compared to an Average Year Between 2006-2020 (Blue). Precipitation Time Series Represents a 14-Day Moving Average. 100

4.5 Projections of Total Bird and Vector Population Sizes. Red Dots Represent the (a) Average Time Series of WNV-Competent Bird Relative Abundance Based on eBird Data for Years 2010-2017 (Red Dots) [19], (b) Mosquito Process-Based Model Output [74] for Average Conditions Between 2080-2100 Under Climate Change Scenario RCP4.5 and (c) RCP8.5 Based on Projected Temperature and Precipitation Data [55]. Blue Curves Show the Simulations of Equation (4.3) from Fitting Parameters of Equation (4.4) to Their Respective Time Series. 101

4.6 Simulations of Infected Vector (Top Row), Bird (Middle Row), and Symptomatic Human (Bottom Row) Population Sizes for an Average Year Between 2080-2100 in the Greater Toronto Area Under Climate Change Scenarios RCP4.5 (Left Side) and RCP8.5 (Right Side). For the Infected Symptomatic Human subfigures, the Blue Lines Represent Simulations Using the Mean Vector Per Capita Biting Rate on Humans $\alpha_h = 0.1896$ from Fitting to Historical Human Case Data, and Shaded Areas Reflect Simulations for the 95% Confidence Interval of Fitted α_h Values ([0.0136, 0.3654]). 104

Figure Page

B-1 Simulations of the Optimal Model Fits for the Total Population in the GTA (Blue Curves) and PBM Time Series (Red Dots) for Years 2005–2019. 125

B-2 Heatmaps of the Normalized Root Mean Squared Error (RMSE) Values from Total Population Fittings with Respect to Candidate Start and End Dates for Years 2005–2019. Dark Purple Regions Denote the Candidate Start and End Date Combinations with the Lowest Normalized Values. 126

B-3 Simulations of the Optimal Model fits for the Active Population in the GTA (Blue Curves) and PBM Time Series (Red Dots) for Years 2005–2019. 127

B-4 Simulations (Curves) and Normalized Root Mean Squared Error Values of the Active Population in the Greater Toronto Area for Fitting Start Dates May 1st–June 1st for Years 2005–2011. 128

B-5 Simulations (Curves) and Normalized Root Mean Squared Error Values of the Active Population in the Greater Toronto Area for Fitting Start Dates May 1st–June 1st for Years 2012–2019. 129

Chapter 1

INTRODUCTION

1.1 Overview of Mosquito-Borne Diseases (MBDs)

Mosquitoes are responsible for more human deaths per year than any other animal [14]. Mosquito-borne diseases are spread to humans and other hosts through the bites of infectious female mosquitoes. These mosquitoes are referred to as vectors, the word also used to denote the mode of disease transmission. Mosquito-borne diseases (MBDs) such as malaria, dengue fever, and West Nile Virus are responsible for 350-650 million human cases and over 630,000 deaths worldwide each year [93, 92]. Three viruses responsible for MBDs – West Nile Virus (WNV), chikungunya, and Zika – were introduced to the Americas over the past 24 years and have since become endemic to these regions [15, 91, 59]. It is of critical interest to understand the degree to which various factors affect MBD transmission and to be prepared for the next MBD outbreak.

1.2 Environmental Factors on Mosquito Biology and MBDs

Mosquitoes go through three main immature stages before reaching their adult stage. Eggs are laid in water by a female mosquito and become larvae upon hatching. From there, the larvae develop into pupae, and finally transform into adult mosquitoes. The egg, larvae, and pupae stages require an aquatic habitat. Development rates at each stage are affected by temperature and other environmental variables [74]. As a result, weather and climate directly impact mosquito biology.

Adult female mosquitoes seek out a blood meal through biting humans, birds, or other mammals in order to develop eggs – a process known as the gonotrophic cycle. It is through the biting process where vector-borne disease (VBD) transmission may occur – a susceptible mosquito may become infected from biting an infectious host, and an infectious mosquito may pass on infection when biting a susceptible host. The incubation period is the time it takes for an infection to develop after a host has been exposed to a disease-causing organism (such as bacteria, viruses, or fungi). The incubation period ends when the first signs or symptoms of the disease appear. The incubation rate denotes the rate at which the pathogen develops inside the host or vector and determines when pathogens are capable of being transmitted. It is the reciprocal of the incubation period that typically lasts several days [63]. Since the average lifespan of mosquitoes tends to be on the order of a couple weeks [86], the age at which a mosquito becomes infected and the length of the incubation period are critical factors in determining how many hosts an infectious mosquito can infect. Both vectors and hosts experience an incubation period upon acquiring infection. This is referred to as the *extrinsic* incubation period (EIP) for vectors, and *intrinsic* incubation period (IIP) for hosts. The incubation period and the survival rate of mosquitoes are both impacted by temperature in a nonlinear way [63, 16]. This is one way in which weather and climate impact the transmission of MBD in addition to the mosquito biology.

Additional biotic and abiotic factors affect the competency and transmission of the disease at both individual and community levels [39]. Competency is the ability to acquire infection and the *potential* to pass infection to others. It is a function of viremia titer in the blood and duration of viremia presence [84]. Transmission refers to the *act* of passing infection to others. Some of these factors include the feeding preference of mosquitoes [37, 75], disease-induced mortality [39], and the cross-protection by antibodies from similar infections [84, 60]. The myriad of factors influencing VBD transmission – many of which are affected by climate itself – create extensive challenges for mathematical modelers.

1.3 Overview of the West Nile Virus (WNV) Transmission Cycle

West Nile Virus (WNV) was first discovered in West Nile, Uganda in 1937 [76]. It was introduced to the United States in 1999 and has since become endemic to the continental U.S. [15]. Most human cases of WNV are asymptomatic but 20% of them result in a short, febrile illness and fully recover within a few weeks [15]. Roughly 1 in 150 infected people develop a severe illness from WNV, such as encephalitis or meningitis, and around 10% of those who contract serious illness die from central nervous complications [15]. Individuals with pre-existing medical conditions (e.g., diabetes, cancer, etc.) are at increased risk of developing serious illness from WNV. There are no vaccines or medicines available to treat WNV, but symptoms can be managed with over-the-counter drugs [15].

West Nile Virus can be transmitted to humans, birds, and other mammals through the bites of infectious female *Culex* mosquitoes. However, cross-infection from humans to mosquitoes typically does not occur because infected humans do not develop the viremia levels necessary for successful transmission. Since cross-infection only occurs from mosquitoes to humans, humans are considered “dead-end” hosts. Avian hosts are the main reservoir for WNV transmission, as cross-infection occurs in both directions between birds and mosquitoes.

1.3.1 Transmission from Vectors to Birds

The time-since-infection (infection-age) plays a role in a mosquito’s probability of passing infection to a host due to the mosquito’s short lifespan and amount of virus present in its saliva. Though mosquitoes live roughly two weeks on average [86], an infected mosquito must live long enough to survive the extrinsic incubation period – the time it takes for the pathogen to develop in the mosquito – before infecting others. The extrinsic incubation period (EIP) typically lasts several days. It was originally thought that older female mosquitoes are the most efficient vectors since the virus has more time to establish in the salivary glands [40, 65], but others argue that young infectious mosquitoes are more likely to transmit to more hosts over their lifetime due to their lower mortality rates [79]. In order for successful transmission to occur from vector to host, the mosquito must overcome the following barriers during the EIP [65]: (1) midgut infection barriers, (2) midgut escape barriers, (3) salivary gland infection barriers, and (4) salivary gland escape barriers. Moreover, successful transmission requires the host to be competent for the disease, which varies among individual birds and other species [39].

1.3.2 Transmission from Birds to Vectors

Infected birds also go through an intrinsic incubation period (IIP) before being able to transmit infection to a susceptible mosquito. It has further been shown that transmission probability from birds to mosquitoes correlates with the amount of viremia in the bird's system [63], and varies nonlinearly as a function of infection-age. Therefore, both the IIP and viremia levels of avian hosts influence the probability of WNV transmission from bird to vector.

Over 300 species of birds have acquired WNV, but WNV competence and transmission potential vary among species [15]. Corvids (e.g., crows, jays) tend to have higher WNV-competence and WNV-induced death rates [39, 60]. Many state and local agencies that monitor areas that have experienced WNV outbreaks have tracked the number of dead birds as a proxy for monitoring WNV [15]. Species richness – the number of different species in an ecological community – and relative abundance of each species contribute to the overall WNV transmission experienced within a habitat, and causes what is known as the "dilution effect" [80]. The diversity of host species coupled with issues of consistent monitoring makes mathematical modeling of WNV disease dynamics a challenging endeavor.

1.4 Mathematical Modeling Paradigms for MBDs

Mathematical modeling has been used for applications of vector-borne disease modeling since Ronald Ross explained the malaria transmission cycle in the early 20th century [68]. Ross discovered that malaria is transmitted through the bites of infectious mosquitoes and conceptualized the idea of the basic reproduction number. The basic reproduction number, $\mathcal{R}_0$, estimates the number of secondary cases that result from the introduction of one infection into a wholly susceptible population. This novel idea showed that it is not necessary to eradicate all mosquitoes to eliminate malaria, but rather that a threshold value of $\mathcal{R}_0 < 1$ is sufficient for eliminating the disease. Later, George Macdonald extended Ross's framework to include analysis of control methods used to eradicate malaria [47]. The Ross-Macdonald framework is still used today to model MBDs [82, 10, 8, 88, 44, 64, 67].

Ross's model used a single differential equation to determine $\mathcal{R}_0$. A framework for general disease epidemics was expanded into a compartmental model by Kermack & McKendrick [36]. Their simple model comprises of three coupled differential equations to represent susceptible (S), infected (I), and recovered (R) populations. Individuals move from susceptible to infectious status by coming in contact with an infectious individualThe Kermack-McKendrick model may be combined with the Ross-Macdonald framework to produce more complex models for MBD transmission.

The original Kermack-McKendrick model is formulated under the assumption of constant population size and homogeneous mixing. These assumptions can be relaxed to accommodate more realistic scenarios. It is often useful to include demographic parameters (birth, migration, natural death, and/or disease-induced death) for disease-endemic regions. Heterogeneity of the population can be incorporated by stratifying the population into classes based on particular attributes, such as sex, age, or behavior. In many compartmental ordinary differential equations (ODEs) models for MBDs researchers have explored various aspects of infection transmission and mitigation, including dependence on weather-related variables and assessment of control strategies [8, 10, 44, 88]. Laperriere, Brugger & Rubel considered temperature dependence on immature and adult mosquito demography, biting rate, and EIP [44], while the impact of precipitation on intraspecific competition and WNV transmission was studied by Wang, et al. [88]. Models that assesses strategies for control and mitigation often arrive at similar conclusions for optimal control. Both Blayneh, et al. and Bowman, et al. found that culling mosquito populations provides a more effective measure to reduce WNV transmission than controlling the bird population or enforcing use of personal protection measures by humans [8, 10].

Compartmental ODE models for disease transmission are often preferred over other types of mathematical and statistical models because of their tractability. Additionally, it is relatively straightforward to find the basic reproduction number and conduct stability analysis of the equilibria for autonomous ODE models. However, it is not always the main goal to study the asymptotic behavior of a model – particularly when systems are highly transient and complex. Extending the modeling framework of an ODE system into a partial differential equation (PDE) system may be a good option because it can provide additional insight about the spatial or age distribution of the population, and may increase accuracy when comparing simulations to data – despite being more mathematically demanding.

Recent PDE models for MBDs have been both theoretical and numerical in nature.

A theoretical PDE model by Maidana & Yang examined the spatial-temporal spread of WNV and showed that traveling wave solutions were mainly impacted by avian advection [48]. A theoretical PDE model by Richard, et al. included chronological-age and infection-age dependence for both human and *Anopheles* mosquito populations for malaria transmission [64]. Richard, et al. used integrated semigroup theory to prove the existence of unique solutions, and further analyzed the stability of equilibria. Proving the existence and uniqueness of solutions is often a more delicate endeavor for PDE models than for ODE models, but is necessary to establish minimal knowledge of the long-term behavior of solutions or to design stable numerical methods. Rock, Wood & Keeling developed a numerical PDE model for general MBD transmission that included both chronological age and biting structure of mosquitoes [67]. Simulations using this model showed that total infections were less sensitive to biting rates than chronological-age structure, suggesting that under the biting rates and age-structure assumptions therein, human treatment may actually be more effective at mitigating disease spread than mosquito control.

1.5 Climate Integrated Model for Mosquito-Borne Infectious Diseases (CIMMID)

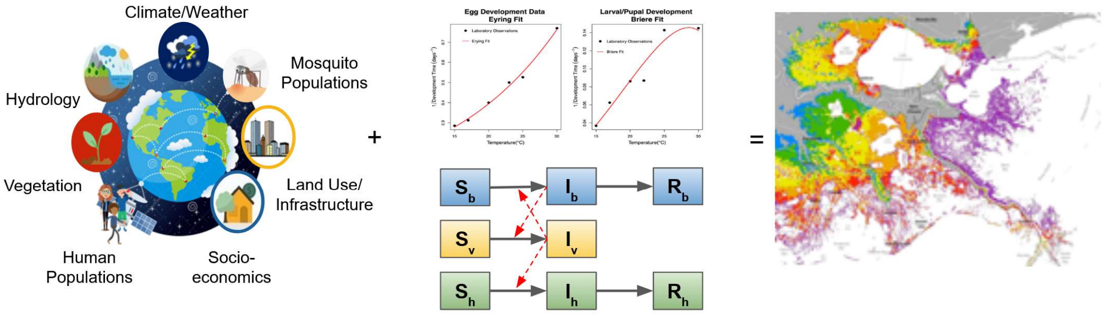
Climate change is arguably one of the most pressing issues affecting our world today, and one of the threats exacerbated by the changing climate is VBD-transmission exeplified in this work by WNV [78, 5]. Mosquito populations are directly impacted by climate in a nonlinear way – it needs to be warm enough for mosquitoes to grow and develop, but not too hot that they cannot survive [16]. Similarly, precipitation affects the availability of egg laying sites, but too much rain can flush out the eggs out of their habitats before they hatch [50, 73, 88, 89]. Temperature is also known to affect the mosquito biting rate [69] and survival [16], and the incubation period of the pathogen [63]. Thus, it is important to create a comprehensive model that considers the numerous factors affecting MBD transmission.

The Climate Integrated Model for Mosquito-borne Infectious Diseases (CIMMID) is a project lead by Carrie A. Manore, Chonggang Xu, and Jeanne Fair, and funded by the Laboratory Directed Research & Development (LDRD) at Los Alamos National Laboratory. The LDRD projects are sponsored by the National Nuclear Security Administration and promote agile responses to national security challenges, advance the frontiers of science and technology, and attract and retain the technical workforce [42].

The goal of CIMMID is to combine mechanistic modeling techniques with heterogeneous data fusion methods to assess future MBD risk under different climate change scenarios across the Americas. This continental-scale model incorporates factors from climate, human and other host species behavior, and mosquito biology. We want to not only know the role that climate has on recent surges of MBD incidence in the region, but also understand how changes in human demographics, climate, and extreme events will impact future MBD risk. To achieve this, we need an understanding of human-mosquito interactions as well as information about mosquito abundance and competence. Furthermore, we must use methods to fuse data across multiple spatial and temporal scales, calibrate parameters for high-fidelity data regions, and quantify uncertainty in parameter estimations, model structure, and forecasts. Figure 1.1 provides a schematic of the overall CIMMID process. This large-scale, interdisciplinary project involves a team of over 40 biologists, mathematicians, statisticians, and computer scientists.

1.5.1 Intergovernmental Panel on Climate Change (IPCC) Scenarios

The Intergovernmental Panel on Climate Change (IPCC) establishes scenarios for future climate conditions based on likely socioeconomic and concentration path-



Real-time, voluminous, noisy data

Mathematical, statistical, and computational models

Probabilistic forecasts with quantified uncertainty

Figure 1.1: Schematic of the CIMMID Process. Multiple Data Streams Related to Weather, Environmental Habitat, and Human and Mosquito Demographics Must Be Fused to Inform Mathematical and Statistical Models to Produce Forecasts of Future Mosquito-Borne Disease Risk.

ways [46]. Shared socioeconomic pathways (SSPs) refer to the policy conditions that influence economic behaviors that may help or harm the changing climate. Representative concentration pathways (RCPs) refer to the CO_2 emissions that lead to increased global surface air temperature (GSAT).

The two most commonly studied climate change scenarios are the SSP2-4.5 and SSP5-8.5 SSP-RCP combinations. The SSP2-4.5 is known as the "middle of the road" scenario. Under this scenario, socioeconomic factors are expected to follow historical trends without significant changes. The SSP5-8.5 is referred to as the "worst case scenario." This scenario assumes CO_2 emissions will double by 2050 due to increased fossil-fuel-driven development. Relative to the conditions from 1850-1900, the estimated GSAT is expected to rise by 2.7°C under SSP2-4.5 and as much as 4.4°C under SSP5-8.5 [46]. To put this in perspective, there has been a 1.09°C increase in GSAT from 1850-1900 to 2011-2022 [46].

In addition to increases of GSAT, global land precipitation is also expected to increase. Relative to the recent past (1995-2014), global land precipitation is expected to increase by 1.5-8.3% under SSP2-4.5 and 0.9-12.9% under SSP5-8.5 [46].

Changes in both air temperature and land precipitation will have consequences for MBD transmission, as both mosquito lifecycle and epidemiology traits are influenced by these climate factors.

1.5.2 The Relevance of Data Fusion Techniques

Mathematical and statistical models have traditionally been two separate classes of modeling – mathematical models aim to understand the underlying process that explain relationships between variables, while statistical models use data to quantify the significance of detected trends in empirical observations. Both types of modeling also have their shortcomings. For example, deterministic mathematical models cannot capture the uncertainty that occurs in nature, and statistical models cannot assess the impact of unobserved variables. But as capabilities for data storage and collection grow alongside our knowledge of underlying processes, scientists and researchers can use the best of both modeling approaches to enhance our understanding of physical and biological phenomena [71]. An outstanding challenge in integrating these two types of modeling is data fusion – the process of transforming collected data into a usable format [5]. Data fusion is needed to reconcile disparate data streams of varying spatial and temporal resolution.

The concept of determining dynamic equations from data is far from new, with notable developments ranging from symbolic regression and chaotic data analysis to adaptive inference [12, 9, 20, 22]. Crutchfield & McNamara developed a method based on chaotic data analysis to reconstruct the deterministic portion of equations of motion from data [20]. Symbolic regression was used by Bongard & Lipson to generate equations for nonlinear coupled dynamical systems from time-series data [9], and inspired the sparse regression method used by Brunton et al. based on the assumption that only a few terms are needed to find the governing dynamics of a system [12]. Other methods for determining dynamic equations from data use adaptive techniques to include an appropriate amount of complexity for the given data resolution [22].

These techniques are extremely useful when the underlying physical laws are relatively unknown. However, these techniques can be computationally expensive and elusive for researchers with limited computer science or mathematics backgrounds. Large-scale modeling frameworks like CIMMID are often comprised of multiple small-scale submodels, and the teams that develop and use the these models tend to be interdisciplinary in nature. The outputs of the individual submodels must both be in a usable format for subsequent models, as well as comprehensible to those who need to use them [52]. Accessibility of data and methods is key to pushing the boundaries of scientific knowledge – it is essential for the platforms developed by interdisciplinary teams to be understood by research scientists of diverse backgrounds.

1.6 Scope of the Dissertation

In this dissertation the goal is to help answer two questions related to improved understanding of the weather, climate, and demographic influences on WNV transmission. The first question is, "What is the impact of vector and host demography on the dynamics of WNV?" The second question is, "what are some possible impacts of expected and likely climate change scenarios on demographic and epidemiological parameters for WNV transmission?"

The second and third chapters of the dissertation are focused on answering the first question. In the second chapter, a non-autonomous logistic model is introduced to infer the dynamic behavior of vector populations from time-series data. The time-varying parameters of the logistic model introduced include periodic behavior of the vector net growth rate and carrying capacity, to describe the seasonality of the mosquito population. We develop a data fusion framework that provides a means to find the optimal model parameters to fit the start and duration of the mosquito season based on time series data derived from the CIMMID mosquito process-based model. The results from fitting the parameters highlight the inter-annual and intra-seasonal variation of mosquito seasons for two different mosquito populations within a single geographic region. We explore the sensitivity of the peak timing and magnitude to the various logistic model parameters, and we investigate the connection between model parameters and weather variables. These non-autonomous parameters are later used in the fourth chapter to estimate the mosquito vector population of a time-continuous epidemiological model for WNV.

The third chapter analyzes the effect of infection-age heterogeneity in both vector and bird host populations on WNV dynamics. We develop a three population partial differential equations (PDE) model for theoretically understanding how infection-age-dependent processes impact the basic reproduction number. To our knowledge, this is the first PDE model for WNV to incorporate infection-age dynamics on both vector and host populations. The inclusion of infection-age heterogeneity in both populations was inspired by experimental data showing the nonlinear relationship between infection-age and WNV transmission from vectors to hosts and from hosts to vectors. We proved the existence and uniqueness of a continuous solution under some model assumptions, and investigated the sensitivity of the basic reproduction number on the infection-age-dependent parameters. This chapter gives insights into how infection-age-dependent processes of vectors and birds affect WNV outbreaks.

The second question posed will be addressed in the fourth chapter of the dissertation. There, we extend the PDE model introduced in the third chapter to account for temperature dependence of the vector mortality rate and transmission from vectors to hosts. The extended model therein is fitted to historical human case data from the Greater Toronto Area for 2010-2017. The extended model incorporating said temperature-dependent processes and the parameters resulting from fittings to real-life WNV-incidence data are then used to project possible future dynamics under the RCP4.5 and RCP8.5 climate change scenarios. Both scenarios lead to significantly increased WNV risk and variability compared to current incidence data for the region considered. This highlights the need for very effective vector control strategies to mitigate future WNV transmission.

Chapter 2

A DATA FUSION FRAMEWORK FOR MODELING MOSQUITO POPULATIONS WITH TIME-DEPENDENT LOGISTIC GROWTH PARAMETERS

2.1 Motivation for Data Fusion Framework

A comprehensive model that incorporates data streams for climate, land cover, and mosquito and human populations is vital to assess future mosquito-borne disease (MBD) risk under possible climate change scenarios. In order to achieve the long-term goal of understanding how MBD transmission will be affected by climate change, it is first necessary to incorporate the seasonal dynamics of mosquito populations into the chosen model for MBD disease transmission. Currently, most epidemiological models for mosquito-borne diseases assume a constant or classical logistic recruitment rate of mosquitoes into a population [82, 10, 8]. However, it can be critical to understand how mosquito abundance in a region changes both within a season as well as over longer time spans of several years or decades. Two possible approaches for connecting seasonal dynamics to epidemiological models include discrete models and fully-coupled models. For discrete models, the dynamic population must be updated and redistributed amongst the model classes at each time for which data is available. This option is challenging because it is not known how the distribution across model classes changes in time. A fully-coupled model would combine the dynamics of the mosquito and human epidemiological models into a single, fully parameterized model. However, this option would limit the amount of complexity that can be incorporated into the model due to the increased number of parameters needed for fitting the model, computational limitations, or differing temporal resolution between data and the model. One way to reconcile the shortcomings of each modeling approach is to concatenate two submodels, first using a highly-detailed mosquito population model that considers many environmental variables and models each mosquito life-cycle stage, followed by a simpler, time-dependent model that incorporates the seasonal behavior of the highly-detailed vector population model into the subsequent epidemiological model.

The overall comprehensive framework of CIMMID involves the combined effort of earth systems models, mosquito population models, and human and other hosts epidemiological models, to produce continental-scale infection risk quantification for mosquito and human populations. Figure 2.1 shows a high-level overview of the CIMMID framework. Climate, weather, land cover, and hydrology data are generated from Earth Systems models. These data streams are then fed as inputs into the mosquito Process-Based Model (PBM) [74], which generates two daily time-series of mosquito populations: (1) the Total Population, consisting of all female mosquitoes in the considered region, and (2) the Active Population, consisting of the average number of female mosquito captures per trap per day. Mosquito population estimates from the PBM are then used in the human epidemiological models for assessing MBD risk. Each of the submodels must be integrated with the others in order for their outputs to be suitable for the next tool in the workflow. Our data fusion framework takes the discrete mosquito time series generated by the PBM and fits yearly periodic per capita baseline growth rate and carrying capacity logistic parameters aimed at reproducing the seasonal behavior observed in the PBM time series. These parameters are then used as continuous-time demographic parameters of the vector population in the human epidemiological models. The human epidemiological models generate disease risk quantification for human and mosquito populations.

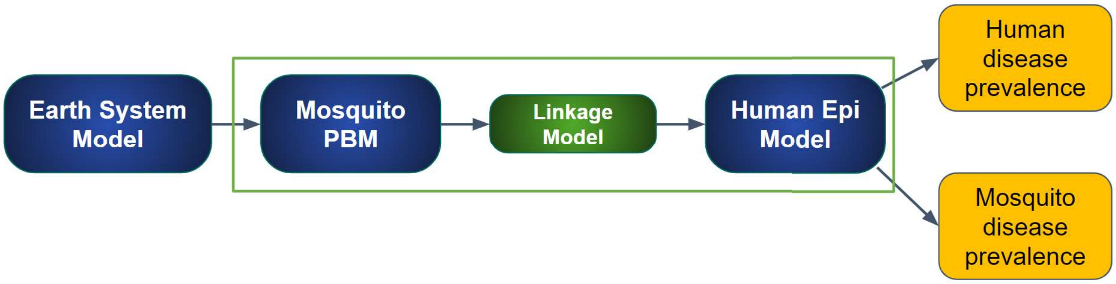


Figure 2.1: High-Level Overview of the CIMMID Framework.

This chapter is focused on the CIMMID submodels highlighted within the green box – the connection between the highly detailed mosquito PBM and the human epidemiological model.

To provide further context that motivates the implementation of our data fusion framework, we first introduce the data sources for the region of interest of this chapter – the Greater Toronto Area, and provide additional details about the mosquito PBM.

2.1.1 Data Sources for the Greater Toronto Area

The Greater Toronto Area (GTA) consists of the City of Toronto and the four surrounding Ontario Public Health Units. Figure 2.2 displays the daily time series of mosquito trap captures, mean temperature, and precipitation for years 2006-2019. Mosquito trap data was obtained from Public Health Ontario's West Nile Virus (WNV) database [61], and temperature and precipitation data were obtained from the ERA5 reanalysis product [32].

Adult mosquitoes are trapped weekly during the mosquito season, which occurs approximately from May through October. Between 2006-2019, over 115,000 mosquito observations from 2,722 trap sites were recorded. Since nearly 85% of the traps used were Light Traps that target female mosquitoes actively seeking a blood meal, we assume that all trapped mosquitoes are female. The majority of identi-

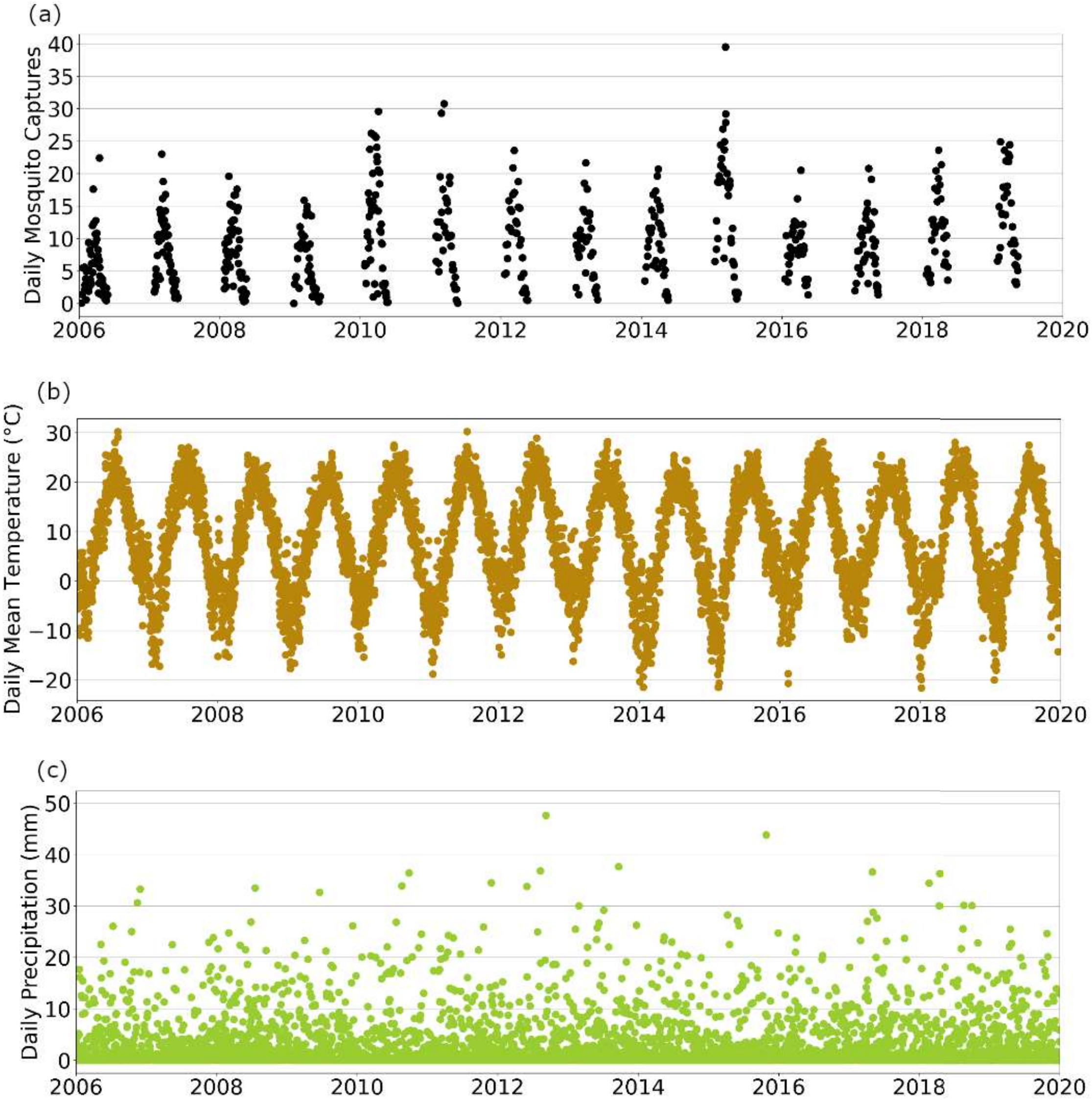


Figure 2.2: (a) Daily Mosquito Trap Captures, (b) Daily Mean Temperature, and (c) Daily Precipitation of the Greater Toronto Area For 2006-2019.

fied mosquitoes were *Culex pipiens* and *Culex restuans*, which are known to transmit WNV [74].

2.1.2 Mosquito Process-Based Model (PBM)

Our data fusion framework is capable of fitting any discrete time series for mosquitoes. However, in order to mitigate the impacts of data gaps or outliers from individual traps, the trap data was first fed as an input for a mosquito process based model (PBM) developed by Shutt, et al. [74]. Two significant challenges of applying trap data to mechanistic mosquito population models are:

1. The sparsity of the data
2. Estimating the true underlying mosquito population size

It is common for mosquito traps not to capture any mosquitoes for days or weeks at a time, making the time series of observations very sparse [74]. Further, since mosquito trap data only accounts for a fraction of the mosquitoes within an area, it is difficult to approximate the actual underlying mosquito population size from existing trap data [70]. To overcome these challenges, the non-autonomous logistic model was fitted to the daily time series mosquito data streams generated by the CIMMID [74].

The PBM mechanistically models the life stages of mosquito egg, larvae/pupae, adult gonotrophic cycles, and diapause to estimate mosquito populations. Data streams for daily temperature, daylight hours, and aquatic habitat availability inform the changes in the population size for each life stage. By incorporating mechanistic dynamics of the mosquito life cycle and heterogeneous fusion of various data streams, the CIMMID mosquito PBM provides an improved estimation of the daily true mosquito population size compared to fitting parameters to the temporally-coarse data alone.

In addition to calibrating daily time series models to mosquito trap data, the PBM can also produce synthetic time series counts of mosquito populations for locations where trap data is unavailable. The PBM was fitted to the GTA's *Culex* trap data for years 2005-2016, and then used to estimate the mosquito population sizes for years 2017-2019 based on temperature and precipitation gauge data stream records. The PBM provides two time series outputs for adult female mosquito populations:

1. Total Population: estimates the total number of female mosquitoes in the GTA each day.
2. Active Population: estimates the average number of mosquito trap captures per trap per day.

The Active Population is used to address the sparsity issue in the mosquito trap data. Trapped mosquitoes are typically in the bloodmeal-seeking stage of their gonotrophic cycle. Therefore, the Active Population can help inform the forecasting and prediction of MBD risk when connected to epidemiological models. On the other hand, the Total Population estimates the total number of female mosquitoes in the GTA each day and, therefore, can be used to inform mosquito mitigation and control measures.

2.1.3 Connecting Mosquito PBM with Human Epidemiological Models

To address the issue of connecting the discrete mosquito PBM time series to the continuous-time human epidemiological model, we fit a non-autonomous logistic model with periodic net growth rate and carrying capacity parameters to each of 15 consecutive years of the mosquito PBM time series. We refer to the non-autonomous logistic model that connects the two CIMMID submodels as the "linkage model" or the "logistic linkage model". We first introduce the classic logistic model, followed by the non-autonomous version with periodic parameters. We then considered ten variations of the non-autonomous logistic growth model, allowing for different combinations of fixed and periodic parameters to ensure a balance between parsimony and enough flexibility to capture observed patterns throughout the considered time period. A model selection procedure based on the Akaike Information Criterion (AIC) shows that a four-parameter non-autonomous logistic model best captures the dynamics of both mosquito time series generated by the PBM, (a) the Total Population including all adult female stages, and (b) the Active Population – blood-seeking females assumed to be in number directly proportional to the number of captures per trap per day. We also quantified the error in each model's projection with respect to the PBM data and determined, not only optimal parameters for each year, but also the best fitting for the start and duration of each mosquito season as these vary from year-to-year. We also explore the sensitivity of the peak timing and magnitude to the non-autonomous logistic parameters for each of the two mosquito population types, and use regression models to model parameter dependence on weather-related variables based on the strongest correlations between them and fitted parameter values. This method addresses the ongoing challenges of data and model fusion by serving as a link between discrete, noisy population data and differential equations for mosquito-borne epidemiology. The non-autonomous demographic parameters for the mosquito population of the GTA are used to estimate the seasonally-varying vector population for a partial differential equations model of WNV (or other VBDs) in the fourth chapter of the dissertation. Thus, this data fusion framework helps to answer the first question of the dissertation about the influence of seasonal demographic parameters on WNV transmission.

2.2 The Logistic Model

A population of size $P = P(t)$ with classical logistic growth dynamics is described by the following mathematical model:

$$\frac{dP(t)}{dt} = rP(t)\Big(1 - \frac{P(t)}{K}\Big). \tag{2.1}$$

The classical logistic growth model was first conceptualized by François Verhulst and has two parameters – the per capita baseline growth rate, r, and the carrying capacity, K [85]. Here we assume that r is the *net* per capita baseline growth rate, that is, the difference between the per capita recruitment and baseline mortality rates of the population. The carrying capacity is the largest population size that can be sustained by the environment. Under logistic growth, if a population is smaller than the carrying capacity, its per capita rate of change decreases linearly as its size approaches the carrying capacity; if a population is larger than the carrying capacity, its per capita rate of change – that is negative in this case – increases linearly as its size approaches the carrying capacity.

Density-dependent biological phenomena such as tumor growth, fishery management, and mosquito populations, can all be modeled quite accurately assuming logistic growth dynamics [41, 56, 69]. The classical logistic model and most of its applications assume the biological parameters remain constant with respect to time. This simplifying assumption not only allows for simple tractability and analysis – as the solution of the governing ODE is then explicit – but also serves as an approximation when detailed population data is unavailable. However, the assumption of constant parameters may fail to capture realistic behavior observed in nature, such as seasonal variations that are much better approximated by periodic functions than by constants. For the case of mosquitoes, whose populations are inherently time-dependent, their growth rate is known to be influenced by environmental temperature [74], and their carrying capacity depends on the availability of egg-laying sites and competition between larvae [50, 73, 88, 89].

These logistic growth models can then be used to assess population control strategies, or incorporated into a larger, mechanistic vector-borne disease modeling framework to study transmission dynamics [50, 90, 82]. Thus, incorporating time-dependent parameters helps to address the uncertainty associated with population and epidemiological models.

2.3 The Non-Autonomous Logistic Model

The non-autonomous logistic model consists of a single ordinary differential equation to represent the rate of change in the size of a population $P = P(t)$:

$$\frac{dP(t)}{dt} = r(t)P(t)\Big(1 - \frac{P(t)}{K(t)}\Big), \tag{2.2.1}$$

where $r(t)$ is the time-dependent per capita baseline net growth rate and $K(t)$ is the time-dependent carrying capacity.

Previous studies of the non-autonomous logistic model incorporate time-dependent parameters and have contributed several theoretical and numerical results. Vance & Coddington showed sufficient conditions for existence and uniqueness of solutions to the non-autonomous logistic model for a population that persists [83]. Coleman established the canonical solution to a non-autonomous logistic model where the carrying capacity either varies slowly in time or remains near a constant value [18]. Others have simulated scenarios modeled by the deterministic non-autonomous logistic model with periodic parameters [28] and time delay [58]. Also, stable periodic solutions have been shown to exist for a stochastic non-autonomous logistic model [35]. Banks [4] provides several applications of non-autonomous logistic models ranging from agricultural populations to railroad mileage.

While the aforementioned results provide a useful basis for understanding the underlying dynamics of various non-autonomous logistic models, current applications tend to lack validation from real-life data and also assume that the infimum of the time-dependent per capita baseline net growth rate is positive. In our application, $P(t)$ is the population size of adult female mosquitoes and $r(t)$ is the difference between the per capita rates of adult emergence and natural adult mortality. Therefore, it is possible for the per capita baseline net growth rate to be negative. Examples of a negative growth rate can occur during extreme weather events that kill a high proportion of juvenile or adult mosquitoes, or may simply be a consequence of typical environmental patterns that cause some species of mosquitoes to enter diapause during the winter season [38]. Here we use periodic functions to represent the yearly seasonal fluctuations of $r(t)$ and $K(t)$:

$$r(t) = r_b - r_s \cos\left(\frac{2\pi t}{365}\right), \tag{2.2.2}$$

$$K(t) = K_b - K_s \cos\left(\frac{2\pi t}{365}\right). \tag{2.2.3}$$

Coefficients with the b subscript denote the baseline (mean) values, and the coefficients with s subscripts represent the amplitude scaling factor of the cosine wave. We choose to use a general time dependence on the parameters instead of including parameter dependence on weather variables (e.g., temperature and precipitation) for two reasons. First, although precipitation and water availability directly relates to the carrying capacity, daily precipitation for the GTA is quite randomly distributed and there is no discernible pattern for the time interval of interest (see Panel (c) of Figure 2.2). Second, the highly-detailed nonlinear interactions between mosquito populations and weather variables are already incorporated in the mosquito PBM, and the purpose of the non-autonomous logistic model is to provide a simple way to link the mosquito PBM output to seasonal demographic parameters for the human epidemiological model. Keeping general time dependence on the parameters makes the non-autonomous logistic model easier to use for mosquito populations from other geographic locations and species whose population fluctuations may have a vastly different response to environmental variables.

2.4 Model Selection Procedure

Model selection involves finding a balance between having a model detailed enough to capture important behavior of interest, but not too detailed that it overfits the given dataset. In order to find the most suitable non-autonomous logistic model with periodic r and K that works well across all years of data, we conducted a model selection procedure based on the Akaike Information Criterion (AIC) [51]. The AIC is based on information theory and discourages overfitting by penalizing models with more parameters. Models with a smaller AIC score can reflect better goodness-of-fit than models with larger AIC score.

Ten candidate non-autonomous logistic models with periodic coefficients were included in the model selection procedure. For each candidate model a subset of the parameters of model (2.2) is fitted to PBM mosquito data while keeping the remainder of the parameters fixed. For the first four models (A–D), a constant carrying capacity is assumed (i.e., $K_b = K$ and $K_s = 0$) and, for the remaining six models (E–K) – a time-dependent carrying capacity. We show in Table 2.1 which parameters are fitted for each candidate model. For candidate models in which we do not fit K or K_b (i.e., candidate models A, B, E, F, H, and J), the carrying capacity is predetermined as the maximum value of the mosquito PBM time-series of each year. Additional aspects of each candidate model are noted below:

- Model A: we use a per capita baseline net growth rate $r_b = 0.01$ and fit the scaling factor r_s. We use the predetermined carrying capacity for K.
- Model B: we fit both coefficients of the per capita net growth rate (baseline r_b and scaling factor r_s), and use the predetermined carrying capacity for K.
- Model C: we use a net growth rate with a baseline value of $r_b = 0.01$ and fit

the scaling factor r_s as well as the carrying capacity K.

- Model D: we fit both coefficients of the per capita net growth rate, along with the carrying capacity K.
- Model E: we use a per capita net growth rate with a baseline value of $r_b = 0.01$ and fit the scaling factor r_s. We also use a carrying capacity with the baseline value K_b equal to the predetermined carrying capacity and the scaling factors $K_s = 100$ for the Total Population and $K_s = 1$ for the Active Population.
- Model F: we fit both the baseline and scaling coefficients of the per capita net growth rate, and use a carrying capacity with baseline coefficient K_b equal to the predetermined carrying capacity and scaling factor $K_s = 100$ for the Total Population and $K_s = 1$ for the Active Population.
- Model G: we use a per capita net growth rate with a baseline value of $r_b = 0.01$ and a scaling factor of $r_s = -0.07$, and fit both coefficients of the carrying capacity (baseline K_b and scaling factor K_s).
- Model H: we use a per capita net growth rate with baseline valua $r_b = 0.01$ and fit the scaling factor r_s. We also use a carrying capacity with baseline coefficient K_b equal to the predetermined carrying capacity and fit the scaling factor K_s.
- Model J: we fit both coefficients of the per capita net growth rate, use a carrying capacity with baseline K_b equal to the predetermined carrying capacity and fit the scaling factor K_s.
- Model K: we fit both coefficients of the per capita net growth rate and both coefficients of the carrying capacity.

Table B.1 provides additional details about the chosen initial values and search space of parameter fitting for each candidate model. Initial values and bounds constraints are based on estimated ranges for the per capita net growth rate [50, 90, 6] and reasonable assumptions for the carrying capacity [1]. We note that this is not an exhaustive list of all possible candidate models of hierarchical nature, but select the ten models that would likely be the most representative of ecological processes.

Parameter fitting was carried out in Python using the `least_squares` function from the SciPy Optimize library [72]. This function minimizes the sum of squared deviations between the PBM output and corresponding model simulation. We use a Trust Region Reflective algorithm for minimization, which allows us to incorporate bounds on the parameter search space as described in Table B.1. Further details about the selected search space constraints are included in Section 2.5. Simulations using the candidate models use a fourth-order Runge-Kutta method to sove the model ODEs.

The parameters for each candidate model were fitted for three different time intervals for each mosquito population (Total and Active). We refer to these tests as "Fitting Seasons," and they roughly represent the duration and timing of mosquito presence in the GTA. The three Fitting Seasons for the Total Population occur from (i) May 1–October 1, (ii) May 15–October 15, and (iii) June 1–November 1, each year. Each of them lasts 154 days, or 22 weeks. The Total Population time series poses additional challenges for fitting the non-autonomous logistic model from November to May since this is the period during which most or all adult female mosquitoes are in diapause. However, it is believed that the majority of the mosquitoes remain in diapause during this time frame and, therefore, are less relevant to model for vector control purposes. On the other hand, the Active Population can be modeled throughout the year (365 days) without issue. The three Fitting Seasons for the

Table 2.1: Candidate Models for the Model Selection Procedure. For Each Candidate Model We Fit a Subset of the Parameters of the Non-Autonomous Logistic Model (2.2). For Models A–D We Consider a Constant Carrying Capacity (i.e., $K_b = K$ and $K_s = 0$), While For Models E–K We Consider a Time-Dependent Carrying Capacity.

Model	Parameters Fitted	Parameters Not Fitted
A	r_s	r_b, K
B	r_b, r_s	K
C	r_s, K	r_b
D	r_b, r_s, K	—
E	r_s	r_b, K_b, K_s
F	r_b, r_s	K_b, K_s
G	K_b, K_s	r_b, r_s
H	r_s, K_s	r_b, K_b
J	r_b, r_s, K_s	K_b
K	r_b, r_s, K_b, K_s	—

Active Population begin (i) May 1, (ii) May 15, and (iii) June 1, each year.

For each population, the AIC for candidate model j = A, B, ... , K, was calculated from the sum of squared errors (SSE) over all $M = 15$ years of time series data [51]:

$$\text{AIC}_j = n\Big[\ln\left(\frac{\text{SSE}_j}{n}\right)\Big] + 2(k+1), \tag{2.3}$$

where,

$$\text{SSE}_j = \sum_{i=1}^{M} \text{SSE}_i^j,$$

is the sum of the SSEs for years $i = 1, 2, ...M$ for candidate model j, n is the number of data points (154 fitting days $\times M$ years for Total Population, 365 fitting days $\times M$ years for Active Population), and k is the number of parameters fitted.

The four-parameter model (Model K) has the lowest AIC for each of the three Fitting Seasons for both Total and Active Populations and, consequently, was selected as the non-autonomous logistic model to fit the GTA mosquito populations. For the remainder of this chapter, “the model” or “the linkage model” refers to the four-parameter, non-autonomous, logistic model (2.2). Further details and discussion of the AIC results are found in Section 2.8.

2.5 Parameter Optimization

The four-parameter model (Model K) was selected as it produced the lowest AIC score for each of the three Fitting Seasons for both Total and Active Populations. We then estimated the seasonal (yearly) net growth rate and carrying capacity parameter functions for the Total and Active Populations of the GTA from 2005–2019 using Model K. As explained in the previous section, a Trust Region Reflective algorithm from the SciPy Optimize library [72] was used to find the optimal seasonal parameters for the time-dependent per capita net growth rate and carrying capacity. Specifically, the coefficients r_b, r_s, K_b, and K_s of (2.2.2)-(2.2.3) were found for each of the 15 seasons for each mosquito population type (Total and Active). The Trust Region Reflective algorithm allows us to incorporate bounds constraints for parameters based on estimated ranges for the per capita net growth rate [50, 90, 6] and reasonable assumptions for the carrying capacity [1]. This algorithm is a least squares method that minimizes the mean squared error between the PBM time series and the mosquito population $P(t)$. Other methods to infer parameters from data include gradient descent [45], Bayesian inference [2, 81], and ensamble-adjusted Kalman filter [23]. Each of these methods are suitable for fitting a small number of model parameters, but we used the Trust Region algorithm so as to also incorporate biologically suitable parameter bounds [72]. This framework is related to the concept of determining dynamics by providing a useful way to connect discrete time series data to numerous time-dependent parameters in a continuous-time modeling approach [12, 22, 9, 20].

2.5.1 Total Mosquito Population

The Total Mosquito Population includes all adult female mosquitoes in the GTA – i.e., the Total Population captures the number of adult female mosquitoes through all stages of the gonotrophic cycle – bloodmeal seeking, digestion and egg maturation, and oviposition – as well as those in diapause. The optimal start and duration of the mosquito fitting season for the Total Population changes from year-to-year and reflects the inter-annual variation of mosquito populations. The majority of mosquitoes remain dormant in diapause during the cold winter and emerge once temperatures become suitable for growth and survival [74]. Mosquitoes in diapause remain non-biting, and thus do not pose a risk for contributing to infection propagation [38]. We explored May 1–June 1 as "candidate start days" to begin fitting the Total Population each year. The range of "candidate start days" refers to the time when mosquitoes begin to emerge from their overwintering state in the GTA. Similarly, "candidate end days" refer to the range of dates when most mosquitoes are likely to be in diapause. We selected October 1–November 1 as the date range for the "candidate end days." The range of candidate start and end days provides an exhaustive grid-search across the likely emergence and disappearance of mosquitoes for each season.

For each year, the initial condition, $P(0)$, was selected as the value of the mosquito PBM time series on the candidate start day. The initialized value of K_b was selected as the maximum of the mosquito PBM time series between the candidate start day and candidate end day. The bounds constraints for the parameters were selected as $-0.2 \leq r_b \leq 0.2$, $-0.35 \leq r_s \leq 0$, $1 \leq K_b \leq 100{,}000$, and $0 \leq K_s \leq 100{,}000$. Bounds constraints for the per capita growth rate parameters were estimated from literature values [50, 90] and were included to help ensure that we obtain biologically relevant values. Although less is known about the bounds constraints for the carrying capacity parameters, we assumed the carrying capacity to be no more than one order of magnitude larger than the largest value in the mosquito PBM time series. Parameters r_b and r_s were initialized as $r_b = 0$ and $r_s = -0.07$ to provide a biologically relevant starting place for optimization. Parameter K_s was arbitrarily initialized as 0.1% of the maximum baseline carrying capacity (i.e., $K_s = 100$) to reflect a small variation in carrying capacity. The selected parameter initialization and constraints direct the optimization algorithm to an appropriate solution – one that is both biologically valid and numerically stable.

Each of the 14,415 fits (31 candidate start-days × 31 candidate end-days × 15 years) returns the fitted parameters along with the root mean squared error (RMSE). The combination of the candidate start-day and candidate end-day with the lowest RMSE for each year is selected as that year's mosquito fitting season. The resulting r_b, r_s, K_b, and K_s values obtained from the mosquito fitting season determine the time-dependent per capita net growth rate and carrying capacity parameters for that year.

2.5.2 Active Mosquito Population

The Active Mosquito Population represents the number of adult female mosquitoes in the GTA currently seeking a blood meal in their gonotrophic cycle, so that they are the most likely to be trapped. In other words, this is the population of mosquitoes that are actively biting humans and other animals, and it provides a reasonable source for understanding the magnitude of mosquito-borne disease risk on a given day. The fitting season of the Active Population lasts 365 days since they go down to zero during the winter, which the linkage model can easily capture (as opposed to the Total Population time series that includes diapausing mosquito dynamics during winter, which we avoided fitting). The range of "candidate start-days" was chosen from May 1–June 1 each year and fitted until the same calendar day the following year. That is, new values for r_b, r_s, K_b, and K_s are fitted every 365 days.

To ensure that the mosquito population remains non-negative, the initial condition, $P(0)$, was selected as maximum between 0.01 and the mosquito PBM value on the candidate start-day. The initial value of K_b was chosen as the maximum value of the mosquito PBM time series during the fitting season. The carrying capacity for the Active Population is much lower than that of the Total Population because it is scaled to the number of adult female mosquitoes captured per trap per day, rather than to the entire population. Therefore, we reasonably assumed the bounds constraints for K_b and K_s to be $1 \leq K_b \leq 1,000$ and $0 \leq K_s < 1,000$, respectively. Initialization and bounds constraints for r_b and r_s are as before, and K_s was initialized as 0.1% of the maximum K_b value (i.e., $K_s = 1$).

The root mean squared error (RMSE) was found for each of the 465 fits (31 candidate start days × 15 years), and the candidate start-day with the lowest RMSE value for each year was selected as that season's start day. For each year, simulations last from the current season's start day until the following season's start day. That is, the 365-day period used to fit the linkage model is extended/truncated to align with the next season's start day. The last season of the time series is simulated for the full 365 day period from that season's start day. This produces a piecewise-continuous function to approximate the Active Population size across the 15 years of time series data.

2.6 Sensitivity Analysis

A sensitivity analysis was conducted to better understand the influence of non-autonomous logistic model parameters on two quantities of interest. The quantities of interest we explored were the peak magnitude and peak timing. The peak magnitude is defined as the greatest daily mosquito population within a season ($\max P(t)$), and the peak timing is the time t at which the peak magnitude occurs. Sobol sensitivity indices [77] of each parameter were computed for the peak magnitude and timing of the Total and Active Populations. Sobol sensitivity analysis is a variance-based sensitivity measure that computes the percentage of variance that can be attributed to each input parameter and their interactions [77, 94].

For each population type (Total and Active), we generated 1,024 (or 2^{10}) samples of parameter combinations for the Sobol sensitivty analysis using the `sobol.sample` function of the SALib sample package [34, 31]. To avoid selecting samples that would be biologically invalid, we first re-scaled the carrying capacity equation (2.2.3) to ensure that the linkage model would always produce a non-negative carrying capacity:

$$K(t) = K_b \left(1 - K_a \cos\left(\frac{2\pi t}{365}\right)\right), \tag{2.4}$$

where $0 < K_a < 1$. Sensitivity indices were computed for five model parameters: the per capita baseline and scaling net growth rate parameters (r_b and r_s), the baseline and re-scaled carrying capacity parameters (K_b and K_a), and the initial condition $P(0)$. Parameter ranges for r_b, r_s, K_b, and $P(0)$ were selected from the ranges obtained from the parameter fitting of each population type (see Sections 2.9 and 2.10, along with Tables B.2 and B.3). We decided to use these ranges for the parameter sampling bounds instead of the original ranges selected for the parameter fitting (see Sections 2.5.1 and 2.5.2) because some combinations of parameters could produce numerically unstable output. In particular, numerical simulations tended to be unstable when large carrying capacity fluctuations were selected with large per capita net growth rate values. Because of this, we limited the range of K_a values to be between 0 and 0.9.

Simulations of the non-autonomous logistic model were run for 184 days for the

Total Population and 300 days for the Active Population for each parameter sample. The time frame for the Total Population aligns with the time frame used for fitting parameters (May 1 – Nov 1), and the time frame for the Active Population was selected to ensure that only one peak would be generated during the period of simulation. The peak magnitude and timing was computed from each simulation for both populations, and the first and total order Sobol sensitivity indices were found using the `sobol.analyze` function of the SALib sample package [34, 31]. First-order indices measure the contribution to output variance by a single model input parameter alone, while the total-order indices measure the contribution to the output variance caused by the model input parameter and all higher-order interactions with other model input parameterss [77, 31].

2.7 Parameter Prediction From Weather Variables

It is well known that climate and weather affect mosquito populations [16, 25, 54]. In order to decide if weather variables can be used to predict parameters of the logistic linkage model from a simple linear relationship, we calculated Pearson correlations between the four fitted coefficients in the parameter functions of the Total and Active Populations (r_b, r_s, K_b, K_s) and five yearly weather variables: (1) mean temperature – the average daily temperature across the calendar year, (2) maximum temperature – the highest daily temperature observed in the calendar year, (3) minimum temperature – the lowest daily temperature observed in the calendar year, (4) temperature range – the difference between the maximum and minimum daily temperatures of the calendar year, and (5) total precipitation – the cumulative precipitation across the calendar year. Temperature and precipitation data was collected from the ERA5 data product [32].

For each mosquito population type, linear regression was applied to the combination of parameter and weather variable producing the strongest correlation. Linear regression was carried out in Python using `Scikit Learn`'s `LinearRegression` function. Generally, weak correlations were found between the 14 years of parameter values and weather variables for both mosquito populations, with a few cases of strong correlations. The strongest correlation for the Total Population was between K_b and total precipitation (correlation of 0.714) and the strongest correlation for the Active population was between r_b and the temperature range (correlation of 0.774). Simulations were performed using predicted values of K_b (for the Total Population) and r_b (for the Active Population) from the linear regression models and compared to the original simulations with mosquito-PBM-time-series-fitted parameters. For the Total Population's simulations using predicted values of K_b, the values of r_b, r_s, initial condition, and start and end days remained at their original fitted values for each year. To avoid the possibility of a negative carrying capacity, values of K_s were selected as the minimum between the original fitted K_s value and 1 less than the regression-predicted K_b. For the Active Population simulations, we used regression-predicted values of r_b, and all other parameters and initial conditions remained at their original mosquito-PBM-time-series-fitted values. Prediction performance was determined from visual parsimony and was divided into three categories: good, acceptable, and poor.

2.8 Results of Model Selection

Akaike Information Criterion (AIC) values from the three tests for the Total and Active Populations in the GTA are shown in Tables 2.2 and 2.3, respectively. For both populations, the four-parameter Model K has the lowest AIC value out of the ten candidate models for each test. The three-parameter model with constant carrying capacity, Model D, also performs relatively well for both populations. Model

Table 2.2: Akaike Information Criterion (AIC) Values For the Ten Candidate Models of Greater Toronto Area's Total Mosquito Population. The Fitting Season Refers to the Start and End Dates of Model Fitting for Each Year of Time Series Data. Each Fitting Season Lasts 154 Days, or 22 Weeks. The Four-Parameter Non-Autonomous Logistic Model (Model K) Produces the Lowest AIC Values for Each Fitting Season.

Fitting Season	**Model A**	**Model B**	**Model C**	**Model D**	**Model E**
May 1–Oct 1	18472	18247	18409	17929	18471
May 15–Oct 15	18735	17923	18580	17716	18730
Jun 1–Nov 1	19428	18125	18966	17909	19426

Fitting Season	**Model F**	**Model G**	**Model H**	**Model J**	**Model K**
May 1–Oct 1	18247	17953	18178	18014	17485
May 15–Oct 15	17924	18572	18164	17876	17644
Jun 1–Nov 1	18124	19141	18798	18048	17853

G, the two-parameter model with time-dependent carrying capacity, performs fairly well for the Total Population when the fitting season occurs from May 1–October 1, but has poor performance when fitting from June 1–November 1. Since the fitting season varies from year-to-year, Model K was selected to ensure better fitting quality throughout the range of potential fitting seasons. The remainder of this chapter uses "the model" or "the linkage model" to refer to the four-parameter non-autonomous model (2.2) (Model K).

2.9 Parameter Fitting Results for Total Population

Parameter fitting results for the GTA's Total Population are shown in Figure 2.3. Panel (a) of Figure 2.3 shows the seasonal fittings of the linkage model (2.2) along with the mosquito PBM output. Panel (b) of Figure 2.3 shows the relative root mean squared error (RRMSE) value for each year's optimal fitting. Although the mean squared error metric was used as the cost function for parameter optimization,

Table 2.3: Akaike Information Criterion (AIC) Values From the Ten Candidate Models of Greater Toronto Area's Active Mosquito Population. "Start" Refers to the Initial Date of Model Fitting for Each Year of the Time Series Data. Model Fitting Lasts 365 Days, or One Full Year. The Four-Parameter Non-Autonomous Logistic Model (Model K) Produces the Lowest AIC Values for Each Fitting Season.

Start	**Model A**	**Model B**	**Model C**	**Model D**	**Model E**
May 1	12223	9137	11125	8400	11821
May 15	13337	8819	12545	8060	12991
Jun 1	14836	8707	13269	7883	14292

Start	**Model F**	**Model G**	**Model H**	**Model J**	**Model K**
May 1	9134	13501	8661	8197	7107
May 15	8826	12912	9930	8263	7384
Jun 1	8725	13363	12381	8658	7671

RRMSE values are presented in Panel (b) of Figure 2.3 to provide comparison of the fitting performance across years. The parameters were fitted separately for each year due to the deterministic nature of the linkage model and yearly variation in the data. Nipa, Yang, & Allen used a stochastic differential equation model with seasonality incorporated for the mosquito population in a dengue transmission model [57]. As an alternative, we could select similar seasonal constants through averaging many realizations of this stochastic model, but that method would be much more computationally intensive.

For the Total Population, it can be observed that the PBM time series exhibits a "tail" towards the end of each mosquito season. This occurs during the period from November through April, where the majority of *Culex* mosquitoes in the GTA remain dormant in a diapause state. During diapause, female mosquitoes are neither biting nor seeking a bloodmeal, and are therefore not a hazard for spreading infection [38]. The non-autonomous logistic model is fitted to the PBM time series data during the period when mosquitoes have emerged from diapause, which occurs roughly between May and October in the GTA.

Figure B-1 provides a closer look at each year's best fitting. Overall, the time-dependent parameters of the non-autonomous logistic model lead to a good resolution of the intra-seasonal and inter-annual variability of the Total Mosquito Population. However, the linkage model fitting struggles to capture the peak magnitude for some seasons – particularly for years 2007 and 2014, both of which had multiple peaks during the mosquito season. Year 2013 had the lowest fitting performance and and it happens to be the year with the greatest peak value and a distinctly bi-modal mosquito season. The substantial increase of peak magnitude for 2013 is likely attributable to unusually high flooding in the GTA for that year [24]. The linkage model (2.2) was not fitted to the Total Population time series from late fall through early spring, as it was assumed that the majority of adult female mosquitoes during that period were in diapause. The mosquito fitting season and fitted parameter values for each year are found in Table B.2 in the Appendix. We note that the minimum carrying capacity value throughout the season is always nonnegative. Interestingly, nearly half of the years show a constant carrying capacity throughout the mosquito fitting season (for 7 of 15 years, $K_s \approx 0$), while other years show larger variability in the carrying capacity where K_s ranges from 31-99% of the baseline value. This explains why the three-parameter model with constant carrying capacity was the second-best performing model after the four-parameter model with time-dependent carrying capacity.

The difference in optimal candidate start- and end-date combinations between years 2011 and 2014 highlight the intra-seasonal variation in mosquito populations within a geographical location, as is apparent in the heatmaps of Figure 2.4. The heatmaps present the normalized RMSE to best compare the fitting performance

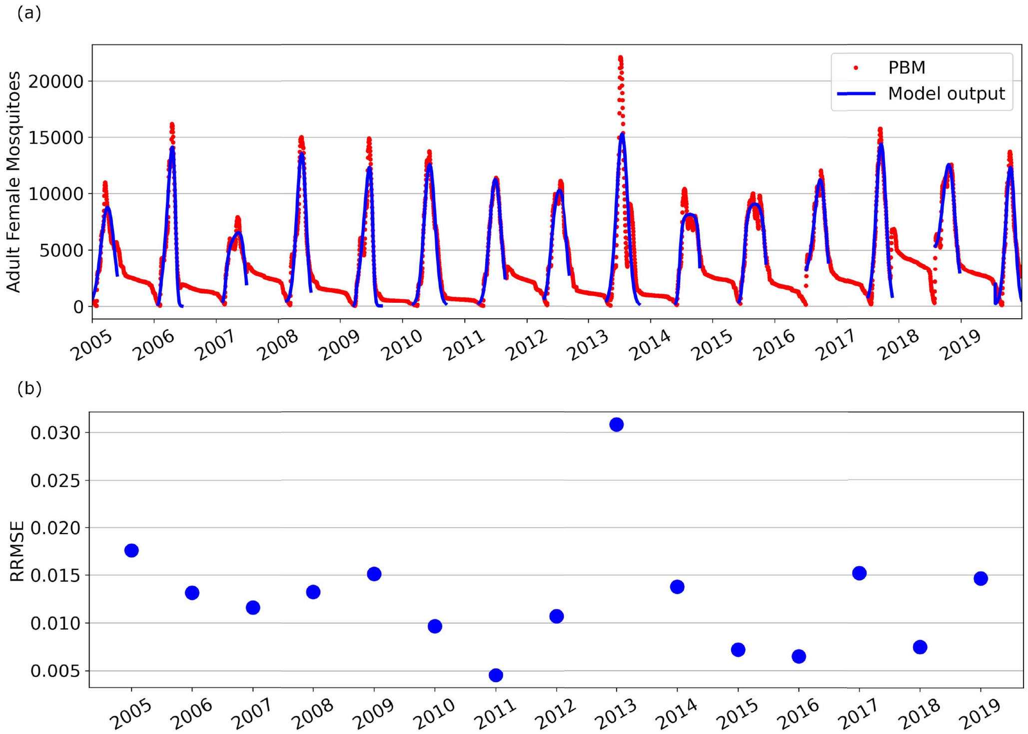


Figure 2.3: Fitting Results for the Greater Toronto Area's Total Mosquito Population for Years 2005–2019. (a) Simulation of Model Fittings (Blue Curves) and Mosquito Process-Based Model Output (Red Dots). (b) Relative Root Mean Squared Error for Each Years' Optimal Fit.

across start and end date combinations within a single year. For 2011, fitting is optimal when the Mosquito Fitting Season begins the first week of May and ends the first week of October. In 2014, a Mosquito Fitting Season from late May to mid October provides the best fitting. Heatmaps of normalized RMSE values for GTA's Total Population for years 2005–2019 are shown in Figure B-2.

2.10 Parameter Fitting Results for Active Population

Parameter fitting results for the GTA's Active Population are shown in Figure 2.5. Panel (a) shows the seasonal fittings of the linkage model (2.2) along with the mosquito PBM output. Panel (b) shows the RRMSE values for each year's optimal fitting. A more detailed look at each year's best fittings is provided in Figure B-3. The

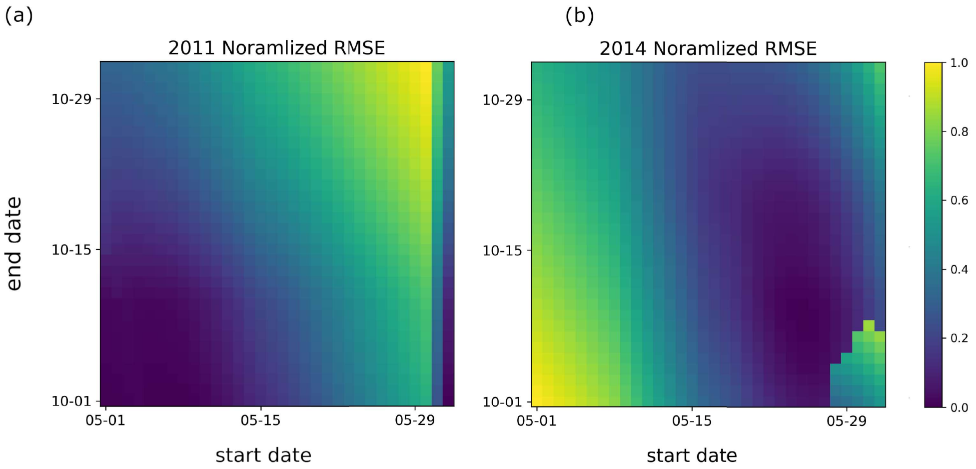


Figure 2.4: Heatmaps of the Normalized Root Mean Squared Error (RMSE) Values from Total Population Fittings with Respect to Candidate Start and End Dates for Years (a) 2011 and (b) 2014. Dark Purple Regions Denote the Candidate Start- and End-Date Combinations with the Lowest RRMSE Values.

linkage model can accurately capture the peak magnitude for most mosquito seasons but may struggle to capture peaks with greater magnitudes or stronger nonlinearities reflected in steeper derivatives in the differential equation, and possibly causing instability in the numerical solver. Additionally, a clear discontinuity is observed between the end of the 2011 season and the beginning of the 2012 season, showing that the linkage model may not always provide a smooth, continuous fitting from year-to-year. Nonetheless, the linkage model simulation generally avoids the noise in the time series signal present at the beginning of each season.

Table B.3 in the Appendix provides the fitted parameters for each season. Similarly to what we observed for the Total Population, the minimum carrying capacity for each season is also nonnegative, and 6 of 15 years show a constant carrying capacity. Interestingly, five of the six years with constant carrying capacity were the same for both the Total and Active Populations.

Intra-seasonal variation of the linkage model fitting performance by start date is shown in Figure 2.6 for the 2013 and 2019 seasons. Panels (a) and (c) show the

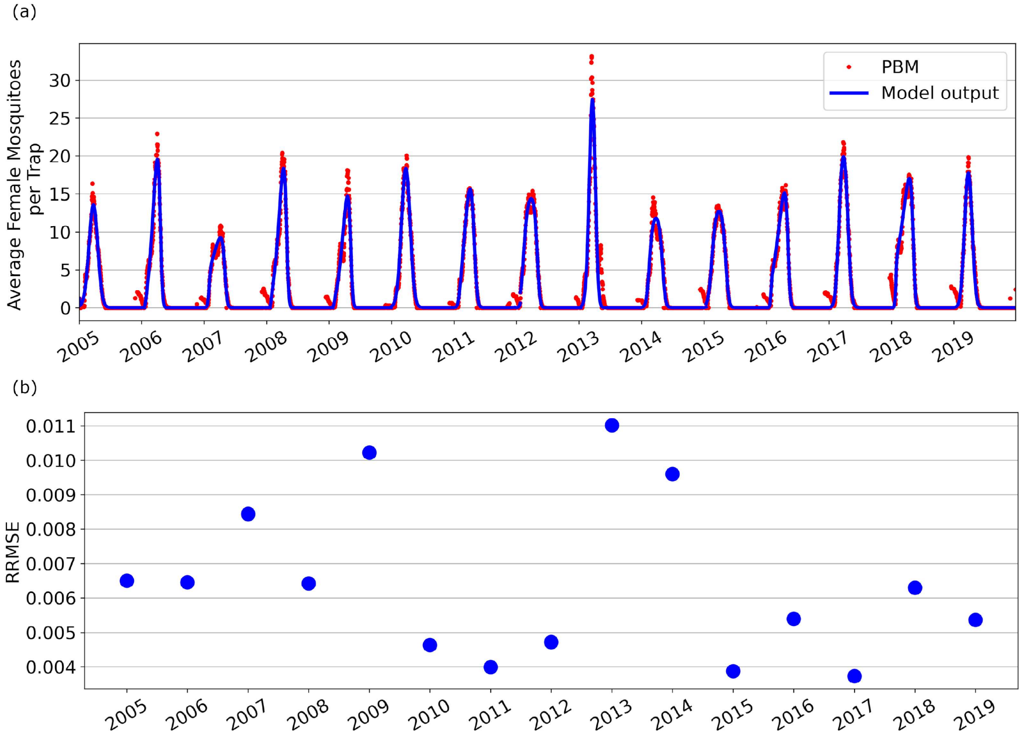


Figure 2.5: Fitting Results for the Greater Toronto Area's Active Mosquito Population for Years 2005–2019. (a) Simulations of Model Fittings (Blue Curves) and Mosquito Process-Based Model Output (Red Dots). (b) Relative Root Mean Squared Error for Each Year's Optimal Fit.

fitting for each start date of the respective seasons, and Panels (b) and (d) show the normalized RMSE for each start date. Similarly to what was observed for the Total Population, the season start date of the Active Population varies year-to-year. The start date of the fitting season directly relates to the peak magnitude achieved in the simulation. May 20th is the optimal start date for the 2013 season (Panel (b)), and May 1st is the optimal start date for the 2019 season (Panel (d)). Intra-seasonal variation for the Active Mosquito Population for all years 2005–2019 is included in Figures B-4 and B-5.

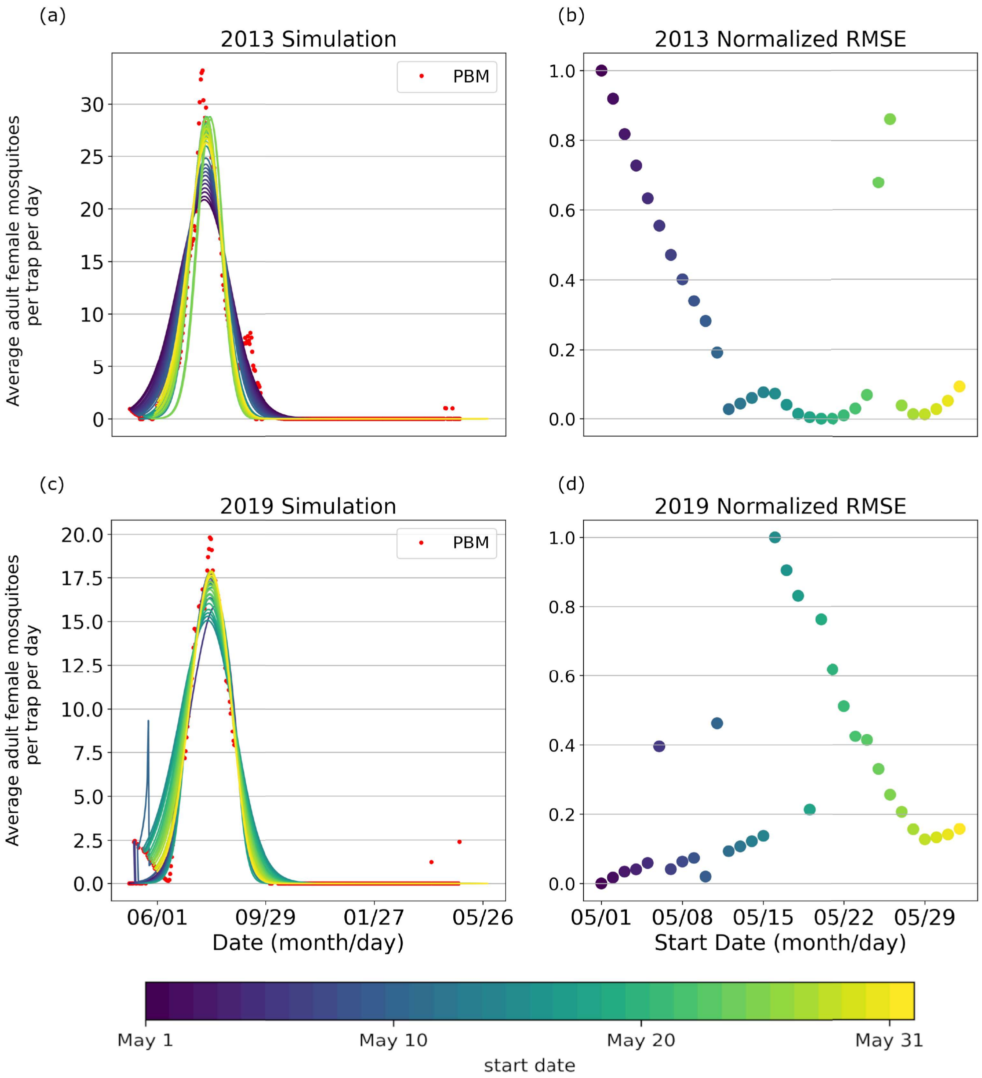


Figure 2.6: Simulations (Curves) of the Active Mosquito Population By Start Date, Mosquito Process-Based Model Output (Red Dots) for Years (a) 2013 and (c) 2019. Normalized Mean Squared Error Values for Each Start Date's Fitting (May 1st–June 1st) for Years (b) 2013 and (d) 2019.

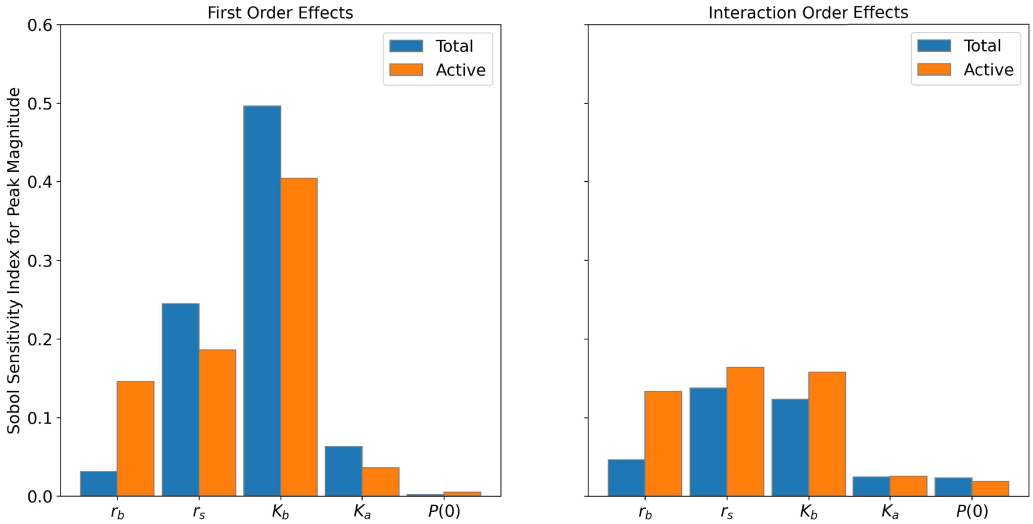


Figure 2.7: Sensitivity of Peak Magnitude to Parameters for the Total (Blue Bars) and Active (Orange Bars) Mosquito Populations in the Greater Toronto Area. Interaction-Order Sensitivities Were Found by Taking the Difference Between the Total-Order and First-Order Effects.

2.11 Results of Sensitivity Analysis

Sensitivity results to model parameters are displayed for the peak magnitude (Figure 2.7) and peak timing (Figure 2.8) of the Total and Active Mosquito Populations in the GTA. The first-order effects measure the single model-parameter's contribution to the output variance, and the interaction-order effects measure the higher-order interactions between model parameter input combinations [77] responsible for output variance. The interaction-order effects were found by taking the difference between the total-order and first-order effects.

The peak magnitude for both mosquito populations is most sensitive to the baseline carrying capacity (K_b) (Figure 2.7), followed by the variation in per capita net growth rate (r_s) – between 40-50% of the output variance is explained by K_b alone, and 18-25% of the variance is explained by r_s alone. We notice that the peak magnitude of the Active Population has greater first- and interaction-order sensitivity to the per capita baseline net growth rate (r_b) than the Total Population. The interaction-order

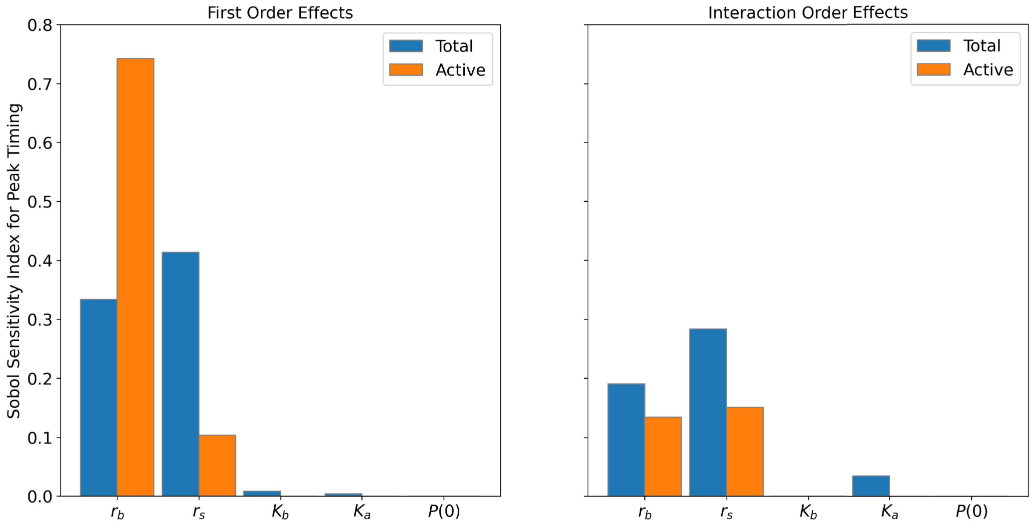


Figure 2.8: Sensitivity of Peak Timing to Parameters for the Total (Blue Bars) and Active (Orange Bars) Mosquito Populations in the Greater Toronto Area. Interaction-Order Sensitivities Were Found by Taking the Difference Between the Total-Order and First-Order Effects.

sensitivities of the peak magnitudes for both populations are similar with respect to other parameters.

For peak timing (Figure 2.8), both populations show the greatest sensitivity to the per capita net growth rate parameters (r_b and r_s). However, the Active Population has substantially greater first-order effects for r_b and less sensitivity to r_s compared to the Total Population. Nearly 88% of the output variance can be attributed to r_b alone for the Active Population, compared to only 53% for the Total Population. On the other hand, roughly 70% of output variance is attributed to r_s for the Total Population, whereas only 25% for the Active Population.

Discrepancies in sensitivity indices between the two populations are likely the result of the slightly different parameter ranges used in the parameter sampling. These differing parameter ranges were selected based on the differing characteristics of the population size's time series. When compared to the Active Population, the fitted baseline carrying capacities and initial conditions for the Total Population were up to two and four orders of magnitude greater, respectively (see Tables S2 and S3).

Further investigation of parameter identifiability and sensitivity may yield additional explanation about these differences [21, 94]. Nonetheless, both population types clearly show that the peak magnitude is most sensitive to K_b and r_s, while the peak timing is most sensitive to r_b and r_s. This suggests that the fluctuations in per capita net growth rate drive both peak magnitude and timing of the mosquito season. Furthermore, the baseline carrying capacity is most influential for the peak size, and the per capita baseline net growth rate is most influential for the timing of the peak.

2.12 Results of Parameter Prediction From Weather Variables

Pearson correlations showed that the strongest linear correlations between parameter functions' coefficients and weather variables occurred between the log of K_b and yearly total precipitation for the Total Population (Panel (a) of Figure 2.9) and between r_b and yearly temperature range for the Active Population (Panel (b) of Figure 2.9). Biologically, this indicates that the baseline carrying capacity of the Total Population increases approximately exponentially as the yearly total precipitation increases, and that the per capita baseline net growth rate of the Active Population increases approximately linearly as the yearly temperature change increases.

Linear regression was used to define linear models to predict K_b and r_b values for the Total and Active Populations, respectively. The exponential model $K_b = 2.217e^{1.0054P_tot}$ can be used to predict the K_b value of the Total Population based on the yearly total precipitation P_{tot}. The linear model $r_b = 0.0129T_{range} - 0.612$ can be used to predict the r_b value of the Active Population based on yearly temperature range T_{range}.

Year-by-year comparisons of fitted parameters and predictive exponential model simulations for the Total Population are displayed in Figure 2.10. Prediction performance was assessed by visual parsimony. Six years were deemed to have "good" predictions by the exponential model (2007, 2011-2013, 2017, 2018), two years with "acceptable" predictions (2010 and 2016), and six years with "poor" predictions (2006, 2008, 2009, 2014, 2015, 2019). Predictions were considered "good" when their corresponding simulations aligned closely to the PBM-time-series-fitted-parameter simulations in both shape and magnitude. Predictions were considered "acceptable" if they were relatively close to the fitted simulations, but may have slight variation in shape or magnitude. Predictions were deemed "poor" when they greatly over or underestimated the magnitude of the fitted simulation peak. Predictions tended to perform more poorly when K_b was overestimated – of the seven years that had a predicted K_b value greater than the fitted value, only two of those years were deemed "good" fits, while the remaining five years were "poor" fits. On the other hand, of the seven years that had a predicted K_b value less than the fitted value, only one was deemed a "poor" fit, while the other six years were considered "good" or "acceptable" fits.

Year-by-year comparisons of PBM-time-series-fitted-parameter and predictive linear model simulations for the Active Population are displayed in Figure 2.11. Nine years were deemed to have "good" predictions by the linear model (2006-2009, 2011, 2012, 2014, 2016, 2018), two years had "acceptable" predictions (2015 and 2019), and three years had "poor" predictions (2010, 2013, 2017). Poor predictions occurred when the predicted r_b value was much greater or much smaller than the fitted value. An interesting observation to make is that both "acceptable" predictions occurred when the the fitted value of $r_b < 0$, but the linear model predicted $r_b > 0$.

Figure 2.12 labels each data point's year from the scatter plots and the regression models from Figure 2.9. The years are labeled by color according to quality of the linear model simulation: blue years represent "good" predictions, yellow years show "acceptable" predictions, and red years indicate "bad" predictions. Overall, the linear model for the Active Population performed better than the exponential one for the

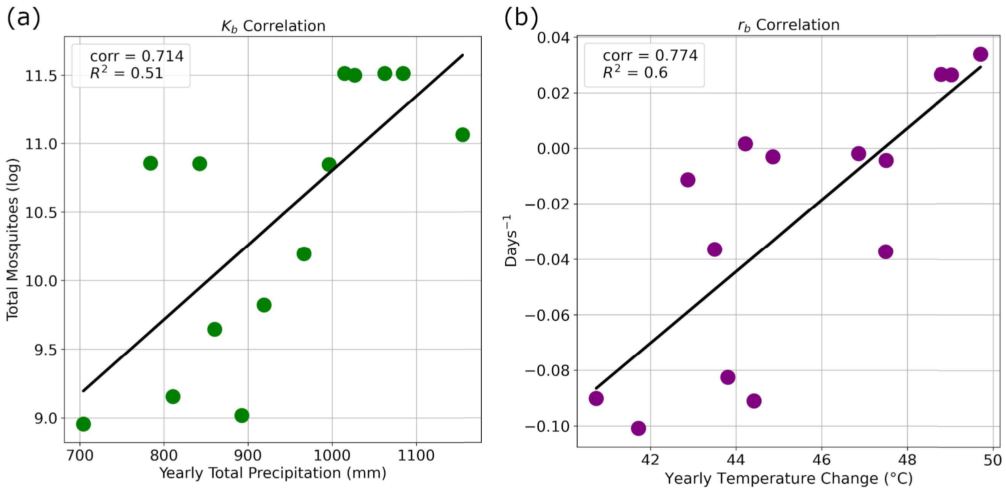


Figure 2.9: Scatter Plots for Models of (a) the Baseline Carrying Capacity K_b of the Total Population and the Yearly Total Precipitation (Exponential), and (b) the Per Capita Baseline Net Growth Rate r_b of the Active Population and Yearly Temperature Change (Linear). Pearson Correlation and R^2 Metrics are Displayed in the Top Left Corner of Each Panel.

Total Population. However, the number of "bad" predictions from either model is too high to consider the predictions reliable. The purpose of exploring fitted-parameter correlations to weather variables and associated regression models was to see if simple relationships with weather variables could be used to predict the parameter functions' coefficients of the non-autonomous logistic model. These results emphasize the nonlinear relationships between mosquito population parameters and weather variables, and how these relationships differ by type of population.

2.13 Discussion

As data availability and storage continue to expand, data-driven mechanistic and mathematical models provide useful insights into the dynamics of various biological processes, such as population growth [70, 12]. Mosquito populations have seasonal fluctuations dependent on nonlinear, exogenous variables related, for example, to climate, land cover, and human behavior [73, 5, 89], and modeling those fluctuations

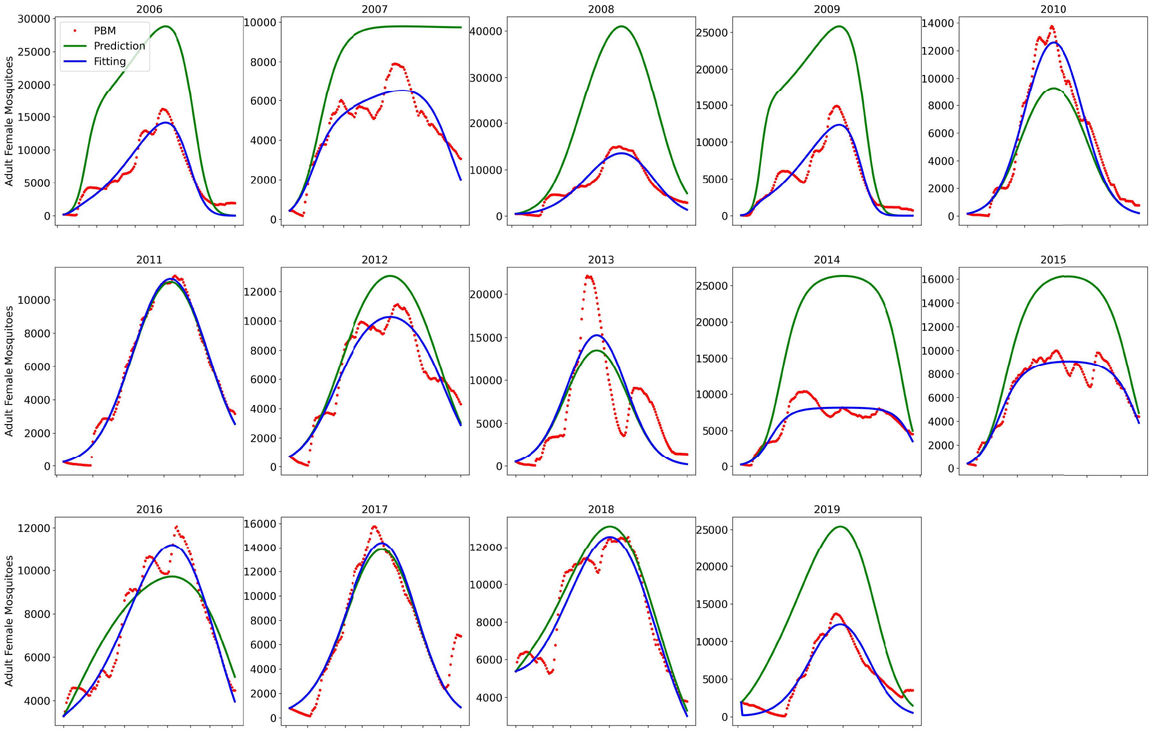


Figure 2.10: Year-by-Year Comparisons of Fitted Parameter (Blue Curves) and Predictive Exponential Model (Green Curves) Simulations for the Total Population. Red Dots Show the Mosquito Process-Based Model Time Series to Which the Non-Autonomous Logistic Model was Fitted.

is an ongoing challenge. The ability to accurately quantify mosquito populations is important – not only to understand their distribution across various spatial and temporal scales, but also to inform vector-borne disease control and mitigation efforts, particularly under the continued threat of climate change [2, 81, 54].

In this chapter, we fitted a non-autonomous logistic model with piecewise periodic per capita net growth rate and carrying capacity parameter functions, to adult female *Culex* mosquitoes in the Greater Toronto Area (GTA) for years 2005–2019. Although most applications using non-autonomous logistic growth typically have only one of the parameters vary through time – either the per capita growth rate or the carrying capacity [4] – there is evidence that seasonal, time-dependent processes affect both parameters, making it preferable to allow possible time-dependence for both of them

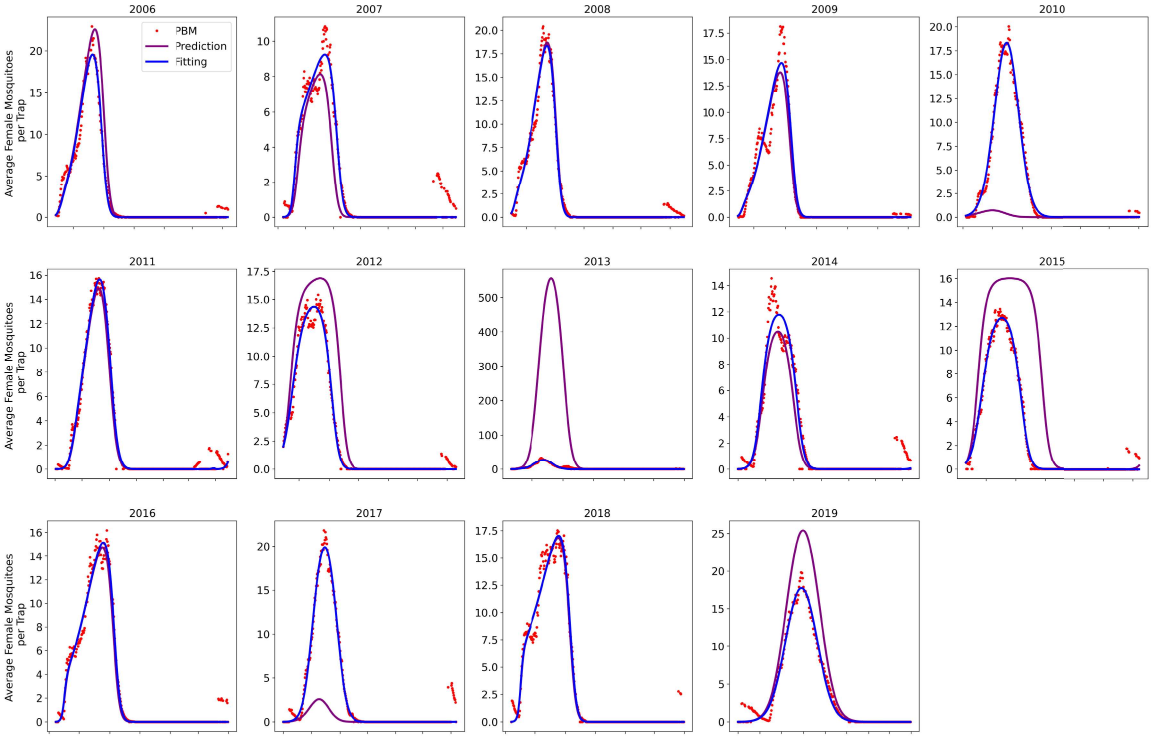


Figure 2.11: Year-by-Year Comparisons of Fitted Parameter (Blue Curves) and Predictive Linear Model (Purple Curves) Simulations for the Active Population. Red Dots show the Mosquito Process-Based Model Time Series to Which the Non-Autonomous Logistic Model was Fitted.

in our application [73, 88]. This assumption was supported through the model selection procedure, which showed that fitting a four-parameter non-autonomous logistic model provided the most consistent fit across the time series data for two types of mosquito populations, Total and Active.

Common challenges of fitting temporally sparse data of an unknown underlying distribution were averted by fitting the per capita net growth rate and carrying capacity parameters to a mosquito process-based model (PBM) output that considers the biological mechanisms of each life stage of the mosquito population [74]. The joint implementation of the PBM and the non-autonomous logistic model allows us to connect discrete time series values to parameters for continuous-time mechanistic models [5]. The submodels are informed by real data, and challenges related to

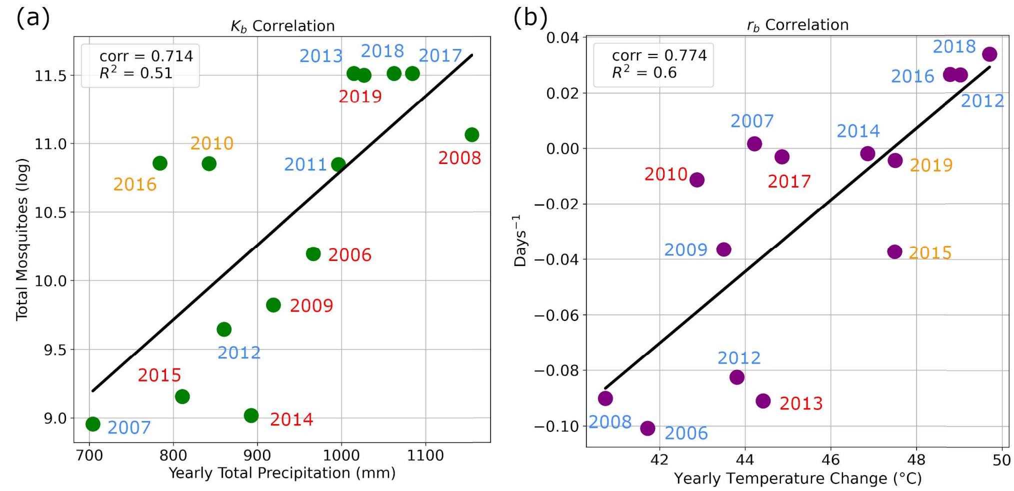


Figure 2.12: Scatter Plots for Regression Models of (a) the Baseline Carrying Capacity K_b of the Total Population and the Yearly Total Precipitation, and (b) the Per Capita Baseline Net Growth Rate r_b of the Active Population and Yearly Temperature Change. The Labels of Each Point on the Scatter Plot Indicate the Corresponding Year of the Fitted Parameter and Weather Variable Combination. The Colors of the Year Labels Indicate the Quality of the Predictive Simulation Using the Regression Model (Black Lines): Blue Years Were "Good" Predictions, Yellow Years Were "Acceptable" Predictions, and Red Years Were "Poor" Predictions.

sample bias are reconciled through the mechanistic framework. Moreover, historical weather data was used as a proxy for estimating the PBM's time series output in the absence of mosquito data (years 2017–2019) [74].

The non-autonomous logistic model captured the inter-annual variability observed in the PBM time series for two mosquito populations in the GTA – the Total Mosquito Population, which considers all adult female mosquitoes, and the Active Mosquito Population, which estimates the average number of captured adult female mosquitoes per trap per day. The quality and performance of model fitting are sensitive to the start date of the fitting, with the optimal day to begin the fitting varying from year-to-year. This highlights the need to incorporate temporal population variation in subsequent models for mosquitoes. An outstanding question of interest is how to select the parameter search space – the final parameter values obtained from the optimization framework were notably sensitive to the initialization values and bounds constraints. Parameter initialization and bounds constraints were selected based on available knowledge of the per capita net growth rate [50, 90] and reasonable assumptions for the carrying capacity [1], which cannot be directly observed in nature. An interesting result of the linkage model fitting was that it showed nearly half the years considered best modeled using a constant carrying capacity ($K_s = 0$) for both populations despite the model having the capability of modeling a time-dependent one. This is consistent with Model D, the three-parameter model with constant carrying capacity, being the second-best performing candidate model. We observe that constant carrying capacities tend to be returned when the PBM time series has a more symmetric peak or when the best fitting curve is symmetric. However, the curves of the PBM time series are greatly non-symmetric for several years of each population, and thus the flexibility of having a time-dependent carrying capacity is desirable to produce a parsimonious output. As such, the interpretation of a time-dependent carrying capacity may provide more of a phenomenological than a biological insight. While our model does capture intra-annual variation, it can struggle to accurately reproduce seasons with multiple, distinct mosquito peaks. This suggests an opportunity for modifying the linkage model for multi-modal data within a season. This is particularly worthwhile for capturing the Total Population during the "off season" (November though April for the GTA) when most mosquitoes are diapausing and not at risk of spreading disease. Incorporating a semi-discrete framework in the linkage model structure can allow us to quantify diapausing mosquitoes that can influence initial infected populations as they emerge from overwintering [38]. Nonetheless, the parameters fitted in the linkage model align with biologically sensible ranges to produce parsimonious fits that balance model accuracy with complexity.

Investigating the idenitifiability of model parameters and the sensitivity of the outputs on them is a critical aspect of understanding data-driven models [94, 21].

We performed a variance-based sensitivity analysis of the mosquito season's peak magnitude and timing to the logistic model parameters using Sobol sensitivity indices [77]. Our results show that both the peak magnitude and timing are sensitive to the amplitude variation of the per capita net growth rate (r_s). Additionally, the peak magnitude is notably sensitive to the baseline carrying capacity (K_b), and the peak timing is highly sensitive to the per capita baseline net growth rate (r_b). Sensitivity indices of the quantities of interest with respect to the model parameters can vary based on the specific characteristics of the mosquito population – we noticed that peak timing was much more sensitive to r_b and less sensitive to r_s for the Active Population compared to the Total Population. Further identifiability studies are required to explain the causes that drive these differing sensitivities [21]. We suspect that the non-autonomous logistic model is practically unidentifiable, as numerous combinations of parameter values can produce equally good fits. The fitted parameters obtained in this study cannot be assumed as "ground truth" and should be interpreted with caution.

Attempts to connect the logistic coefficient functions' parameters to weather variables were less successful than we had hoped. We fitted exponential and linear models of a single weather variable and parameter function coefficient to predict mosquito dynamics using the non-autonomous logistic model. The yearly total precipitation was selected to predict the baseline carrying capacity of the Total Population, and the yearly temperature range was selected to predict the per capita baseline net growth rate of the Active Population based on these relationships having the highest Pearson correlations out of any parameter function coefficient and weather variable combination explored. Despite the relatively high correlations, only 58%, or 8 out of 14 years of the Total Population simulations produced "good" or "acceptable" regression-based-predictions when compared to their PBM-time-series-parameter-fitted simulations.

The Active Population was fitted quite satisfactorily, showing high correlation and R^2 metric with the yearly temperature range, and led to an excellent 78%, or 11 out of 14 years of “good” or “acceptable” simulations. We suspect that these results stem from the nonlinear effect weather has on population dynamics. Immature development time, egg viability, and adult lifespan are all nonlinear functions of temperature [54]. The interactions of these nonlinear processes make it difficult to reliably identify the relation between growth rate and temperature. Though it is possible to extend the non-autonomous parameter functions to allow for more complex nonlinearities and multiple weather variables, we aimed to create a simple model to match its output to the discrete mosquito PBM time series [74]. The PBM is a highly detailed, high-fidelity model that considers not only temperature and precipitation, but other water level data such as runoff and soil moisture, to construct a daily time series from observed trap data. Thus, many of the nonlinear interactions between weather variables and mosquito populations are already captured in the PBM, and our logistic linkage method provides a way to seamlessly connect discrete time series data to continuous-time epidemiological models. The PBM essentially fills in the gaps associated with sparse data sets and is used to infer the total number of mosquitoes within each annual life cycle and gonotrophic stage. Trap data captures only a fraction of the mosquitoes that are currently biting. Some epidemiological modeling studies informed by trap data estimate the true mosquito population by multiplying the trap data by a constant factor [75], but this assumption is neither tied to ecological processes nor considers the impact of environmental variables. Consequently, parameters in our model were fitted to the PBM time series instead of the raw trap data. These considerations highlight the intricacies and complexities of modeling mosquito population dynamics as it relates to weather and climate variables.

Although the results shown here are specific to the *Culex pipiens* and *Culex restu-*

ans populations in the GTA, the non-autonomous logistic model framework can be adapted to other mosquito species and geographic locations.

Parameters obtained for the Total Population can be used to analyze and inform mitigation and control measures of mosquito populations, while parameters obtained for the Active Population can aid the forecasting and prediction of mosquito-borne disease risk in deterministic epidemiology models. The purpose of our linkage model is to reconstruct an observed signal to be used as a subsequent model input for a mosquito-borne epidemiological model. The non-autonomous logistic mosquito parameters will be used to estimate the seasonal vector population for the WNV partial differential equations model in chapter 4.

The concept of determining dynamic equations from data is far from new, with notable developments ranging from symbolic regression and chaotic data analysis to adaptive inference [12, 9, 20, 22]. These techniques are useful when the underlying physical laws are relatively unknown. However, these techniques can be computationally expensive and elusive for researchers with limited computer science or mathematics backgrounds [52]. Our framework can be very useful for applications where the functional form of time-dependent parameters or candidate models are assumed to be known, but optimization of numerous time-dependent parameter functions is desired. Other applications that could benefit from this method include reactions with Michaelis-Menten dynamics, the initiation of action potentials in neuroscience, gene regulation, and circuit signals [22, 12]. The presented method provides a useful way to connect discrete time-series data to a continuous-time modeling framework.

While cloud computing has increased accessibility for researchers to analyze large, heterogeneous datasets [52], lack of publicly available data can still inhibit the ability to obtain and analyze high-resolution data. There is currently no national open-access mosquito data repository or standardized protocol for mosquito data collection for the United States [70]. This creates challenges for both data acquisition, data fusion, and population modeling [5].

The future is promising for data-informed mechanistic models [5]. Statistical models alone are insufficient to estimate parameters in some coupled systems or systems with unobserved variables [71], yet data is also necessary to validate mathematical models and produce reliable modeling forecasts. Mechanistic models are useful to assess dynamics in locations where data is lacking or underlying drivers of the dynamics are changing, for example, estimating the effect of the carrying capacity that cannot be directly measured or assessing the impacts of mitigation strategies. Comprehensive frameworks, such as CIMMID, that combine statistical with mathematical modeling approaches, can provide robust output to inform decision makers and resource allocation [5].

Chapter 3

INFECTION-AGE DEPENDENT WEST NILE VIRUS MODEL

3.1 Mathematical Model

A deterministic system of partial differential equations models the West Nile Virus (WNV) dynamics among three populations: mosquito vectors, bird hosts, and human dead-end hosts. The total population sizes of the vectors, birds, and humans are represented by N_v, N_b, and N_h, respectively. Each species' population is divided into compartments based on infection status:

$$N_v(t) = S_v(t) + I_v(t), \tag{3.1.1}$$

$$N_b(t) = S_b(t) + I_b(t) + R_b(t), \tag{3.1.2}$$

$$N_h(t) = S_h(t) + I_h(t) + R_h(t), \tag{3.1.3}$$

where S-compartments denote susceptible populations, I-compartments denote infectious, and R-components denote recovered populations. Due to the relatively short lifespan of mosquito vectors, the model assumes infected mosquitoes remain infectious until death.

The infectious compartments for vectors and birds are further structured by infection-age τ, defined as the time since acquired infection. Therefore,

$$I_v(t) = \int_0^\infty i_v(\tau, t) d\tau, \tag{3.2.1}$$

$$I_b(t) = \int_0^\infty i_b(\tau, t) d\tau, \tag{3.2.2}$$

where $i_v(\tau, t)$ and $i_b(\tau, t)$ are the population densities of infected vectors and birds with infection-age τ, respectively.

Table 3.1: State Variables and Parameters of the Model (3.3).

State Variable	**Description**
$S_v(t)$	Number of susceptible vectors at time t
$I_v(t)$	Number of infected vectors at time t
$i_v(\tau, t)$	Infection-age density of vectors with infection-age τ at time t
$S_b(t)$	Number of susceptible birds at time t
$I_b(t)$	Number of infected birds at time t
$i_b(\tau, t)$	Infection-age density of birds with infection-age τ at time t
$R_b(t)$	Number of recovered birds at time t
$S_h(t)$	Number of susceptible humans at t
$I_h(t)$	Number of infected humans at time t
$R_h(t)$	Number of recovered humans at time t

Parameter	**Description**
Λ_v	Vector recruitment rate
μ_v	Natural vector per captia death rate
α	Per capita vector biting rate
p	Probability that vector bite is on a bird
$\beta_{vb}(\tau)$	Probability of disease transmission from infected vectors with infection-age τ to susceptible birds
$\beta_{vh}(\tau)$	Probability of disease transmission from infected vectors with infection-age τ to susceptible humans
$\beta_{bv}(\tau)$	Probability of disease transmission from infected birds with infection-age τ to susceptible vectors
$\gamma_b(\tau)$	Per capita recovery rate of infected birds with infection-age τ
γ_h	Per capita recovery rate of infected humans

The model consists of the following partial differential equation (PDE) system:

$$\frac{dS_v}{dt} = \Lambda_v - \lambda_{bv}(t)S_v(t) - \mu_v S_v(t), \tag{3.3.1}$$

$$\frac{\partial i_v}{\partial t} + \frac{\partial i_v}{\partial \tau} = -\mu_v i_v(\tau, t), \tag{3.3.2}$$

$$\frac{dS_b}{dt} = -\lambda_{vb}(t)S_b(t), \tag{3.3.3}$$

$$\frac{\partial i_b}{\partial t} + \frac{\partial i_b}{\partial \tau} = -\gamma_b(\tau)i_b(\tau, t), \tag{3.3.4}$$

$$\frac{dR_b}{dt} = \int_0^\infty \gamma_b(\tau)i_b(\tau, t)d\tau, \tag{3.3.5}$$

$$\frac{dS_h}{dt} = -\lambda_{vh}(t)S_h(t), \tag{3.3.6}$$

$$\frac{dI_h}{dt} = \lambda_{vh}(t)S_h(t) - \gamma_h I_h(t), \tag{3.3.7}$$

$$\frac{dR_h}{dt} = \gamma_h I_h(t). \tag{3.3.8}$$

Adult vectors are recruited into their susceptible population at constant per capita rate Λ_v and die naturally at constant rate μ_v. Vectors become infected at per capita rate $\lambda_{bv}(t)$ – the force of infection from infected birds to susceptible vectors:

$$\lambda_{bv}(t) = \frac{\alpha p}{N_b + N_h} \int_0^\infty \beta_{bv}(\tau)i_b(\tau, t)d\tau. \tag{3.4.1}$$

The per capita vector biting rate is α, p is the percentage of vector bites that are on birds, and $\beta_{bv}(\tau)$ denotes the per-bite probability of disease transmission from birds with infection-age τ to susceptible vectors. Similarly, $\lambda_{vh}(t)$ and $\lambda_{vb}(t)$ represent the forces of infection from infectious vectors to susceptible humans and from infectious vectors to susceptible birds, respectively.

$$\lambda_{vh}(t) = \frac{\alpha(1-p)}{N_b + N_h} \int_0^\infty \beta_{vh}(\tau)i_v(\tau, t)d\tau, \tag{3.4.2}$$

$$\lambda_{vb}(t) = \frac{\alpha p}{N_b + N_h} \int_0^\infty \beta_{vb}(\tau)i_v(\tau, t)d\tau. \tag{3.4.3}$$

Infected birds with infection-age τ recover from infection at per capita rate $\gamma_b(\tau)$. Infected humans recover at per-capita rate γ_h. Table 3.1 includes the state variables and parameters of the model (3.3). The model (3.3) uses the following initial conditions,

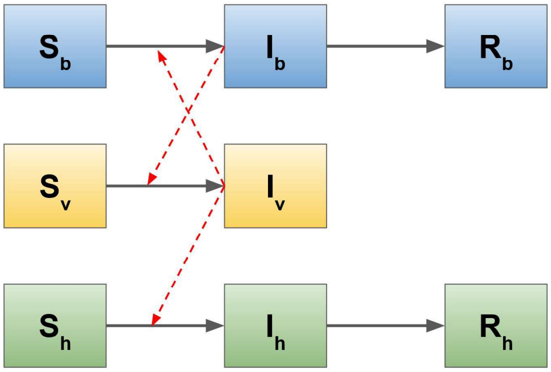


Figure 3.1: Schematic Diagram of the Model (3.3). The Red Dashed Lines Represent the Cross Transmission Between Populations.

$$\begin{aligned}(S_v(0), i_v(\tau,0), &S_b(0), i_b(\tau,0), R_b(0), S_h(0), I_h(0), R_h(0)) \\ &= \left(S_v^0, i_v^0(\tau), S_b^0, i_b^0(\tau), R_b^0, S_h^0, I_h^0, R_h^0\right),\end{aligned} \tag{3.5.1}$$

and boundary conditions – corresponding to new infections, with infection-age 0,

$$i_v(0,t) = \lambda_{bv}(t)S_v(t), \quad i_b(0,t) = \lambda_{vb}(t)S_b(t). \tag{3.5.2}$$

A schematic diagram of the model (3.3) is shown in Figure 3.1. The following assumptions pertain to the model (3.3):

- The total population sizes of vectors, birds, and humans remain constant. This assumption facilitates the proof of the existence and uniqueness of solutions.
- Although WNV-induced death affects certain bird species [39] and a small fraction of infected humans [15], we neglect disease-induced death to make the analysis of the model simpler.
- The model is intended to capture the *seasonal* dynamics of WNV, which is typically on the order of several months. Therefore, demographic parameters

(i.e., birth and death rates) for birds and humans are neglected. Demographic parameters for vectors are included since their lifespan is generally on the order of days to weeks [86].

- For simplicity, the model only considers vector bites on humans and birds. Vector bites on other mammals are neglected.
- Vectors transmit infections to both birds and humans, but only birds can transmit infection back to vectors due to the low amounts of viremia found in infected human blood [15]. The probability of infection transmission depends on the infection-age of the vector or bird.
- Due to the short lifespan of the vectors, we assume that infected vectors remain infectious until death [86].
- Birds and humans may recover from infection. Birds recover at an infection-age-dependent rate, and humans recover at an infection-age-independent rate.
- Due to the lack of conclusive evidence surrounding WNV re-infection, it is assumed that humans and birds who have recovered from infection maintain lifelong immunity against future WNV infection.

3.2 Integral Form

It is more convenient to work with the integral form of the model (3.3). First, Equations (3.3.2) and (3.3.4) are integrated along the characteristics to yield the infection-age-density of each population:

$$i_v(\tau,t) = \begin{cases} i_v^0(\tau-t)\frac{\Pi_v(\tau)}{\Pi_v(\tau-t)} & t<\tau, \\ \lambda_{bv}(t-\tau)S_v(t-\tau)\Pi_v(\tau) & \tau\le t, \end{cases} \tag{3.6.1}$$

$$i_b(\tau,t) = \begin{cases} i_b^0(\tau - t)\frac{\Pi_b(\tau)}{\Pi_b(\tau - t)} & t < \tau, \\ \lambda_{vb}(t-\tau)S_b(t-\tau)\Pi_b(\tau) & \tau \leq t, \end{cases} \tag{3.6.2}$$

where,

$$\Pi_v(\tau) = e^{-\mu_v \tau}, \tag{3.7.1}$$

$$\Pi_b(\tau) = e^{-\int_0^\tau \gamma_b(s)ds}, \tag{3.7.2}$$

denote, respectively, the probability of vector survival to age τ and the probability of infected birds recovering at exactly age τ. Detailed calculations determining the characteristic form of (3.3.2) can be found in Appendix B. The method to determine the characteristic form of (3.3.4) follows similarly. Equation (3.7.1) is the probability that an infected vector survives to infection-age τ, and Equation (3.7.2) is the probability that an infectious bird remains infectious at infection-age τ.

We would like to reformulate the model to be a function of t only. We solve the formally linear ODEs (3.3.1), (3.3.3), (3.3.5)-(3.3.8) to find expressions for state variables $S_v(t)$, $S_b(t)$, $R_b(t)$, $S_h(t)$, $I_h(t)$, and $R_h(t)$, and Equation (3.6) with respect to τ to obtain $I_v(t)$ and $I_b(t)$. The integral form of the model (3.3) follows:

$$S_v(t) = S_v^0 \Pi_v(t) e^{-\int_0^t \lambda_{bv}(s)ds} + \Lambda_v \int_0^t e^{-\int_0^\eta \lambda_{bv}(a)da} \Pi_v(\eta)\eta, \tag{3.8.1}$$

$$\begin{aligned} I_v(t) = &S_v^0 \Pi_v(t) \int_0^t \lambda_{bv}(\eta) e^{-\int_0^\eta \lambda_{bv}(a)da} d\eta \\ &+ \Lambda_v \int_0^t \Pi_v(t-s) \int_s^t \lambda_{bv}(\eta) e^{-\int_0^{\eta - s} \lambda_{bv}(a)da} d\eta ds + F_1(t), \end{aligned} \tag{3.8.2}$$

$$S_b(t) = S_b^0 e^{-\int_0^t \lambda_{vb}(s)ds}, \tag{3.8.3}$$

$$I_b(t) = S_b^0 \int_0^t \lambda_{vb}(\eta) \Pi_b(t-\eta) e^{-\int_0^t \lambda_{vb}(s)ds} d\eta + F_2(t), \tag{3.8.4}$$

$$R_b(t) = S_b^0 \int_0^t \int_0^s \gamma_b(\tau) \lambda_{vb}(s-\tau) \Pi_b(\tau) e^{-\int_0^{s-\tau} \lambda_{vb}(a)da} d\tau ds + F_3(t), \tag{3.8.5}$$

$$S_h(t) = S_h^0 e^{-\int_0^t \lambda_{vh}(s)ds}, \tag{3.8.6}$$

$$I_h(t) = I_h^0\Pi_h(t) + S_h^0\int_0^t \lambda_{vh}(s)\Pi_h(t-s)e^{-\int_0^s \lambda_{vh}(a)da}ds, \tag{3.8.7}$$

$$R_h(t) = \gamma_h S_h^0\int_0^t\int_0^\eta \lambda_{vh}(s)\Pi_h(\eta-s)e^{-\int_0^\eta \lambda_{vh}(a)da}dsd\eta - \frac{\gamma_h I_h^0\Pi_h(t)}{\mu_v}, \tag{3.8.8}$$

where,

$$\begin{aligned}\lambda_{vb}(t) = &\frac{\alpha p}{N_b(t)+N_h(t)}\Bigg(S_v^0\Pi_v(t)\int_0^t \beta_{vb}(t-\tau)\lambda_{bv}(\tau)e^{-\int_0^\tau \lambda_{bv}(\eta)d\eta}d\tau \\ &+ \Lambda_v\int_0^t \Pi_v(t-\tau)\int_\tau^t \beta_{vb}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^{\eta-\tau}\lambda_{bv}(\xi)d\xi}d\eta d\tau + G_1(t)\Bigg),\end{aligned} \tag{3.9.1}$$

$$\begin{aligned}\lambda_{vh}(t) = &\frac{\alpha(1-p)}{N_b(t)+N_h(t)}\Bigg(S_v^0\Pi_v(t)\int_0^t \beta_{vh}(t-\tau)\lambda_{bv}(\tau)e^{-\int_0^\tau \lambda_{bv}(\eta)d\eta}d\tau \\ &+ \Lambda_v\int_0^t \Pi_v(t-\tau)\int_\tau^t \beta_{vh}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^{\eta-\tau}\lambda_{bv}(\xi)d\xi}d\eta d\tau + G_2(t)\Bigg),\end{aligned} \tag{3.9.2}$$

$$\lambda_{bv}(t) = \frac{\alpha p}{N_b(t)+N_h(t)}\Big(S_b^0\int_0^t \beta_{bv}(t-\tau)\Pi_v(t-\tau)\lambda_{vb}(\tau)e^{-\int_0^\tau \lambda_{vb}(\eta)d\eta}d\tau + G_3(t)\Big), \tag{3.9.3}$$

$$F_1(t) = \int_t^\infty i_v^0(\tau-t)\frac{\Pi_v(\tau)}{\Pi_v(\tau-t)}d\tau, \tag{3.9.4}$$

$$F_2(t) = \int_t^\infty i_b^0(\tau-t)\frac{\Pi_b(\tau)}{\Pi_b(\tau-t)}d\tau, \tag{3.9.5}$$

$$F_3(t) = \int_0^t\int_s^\infty \gamma_b(\tau)i_b^0(\tau-s)\frac{\Pi_b(\tau)}{\Pi_b(\tau-s)}d\tau ds, \tag{3.9.6}$$

$$\Pi_h(t) = e^{-\gamma_h t}, \tag{3.9.7}$$

and,

$$G_1(t) = \int_0^\infty \beta_{vb}(t+\eta)i_v^0(\eta)\frac{\Pi_v(t+\eta)}{\Pi_v(\eta)}d\eta, \tag{3.9.8}$$

$$G_2(t) = \int_0^\infty \beta_{vh}(t+\eta)i_v^0(\eta)\frac{\Pi_v(t+\eta)}{\Pi_v(\eta)}d\eta, \tag{3.9.9}$$

$$G_3(t) = \int_0^\infty \beta_{bv}(t+\eta)i_b^0(\eta)\frac{\Pi_b(t+\eta)}{\Pi_b(\eta)}d\eta. \tag{3.9.10}$$

Detailed calculations to find the integral form of the model (3.3) can be found in Appendix B. Equations (3.9.4)-(3.9.10) can be parameterized directly from available data. Additionally, (3.9.1)–(3.9.3) can be shown as a self-contained system – notice that, with the exception of initial data, λ_{vb} and λ_{vh} depend only on λ_{bv}, and λ_{bv} depends only on λ_{vb} in their expressions therein.

3.3 Existence and Uniqueness of Solutions

The existence of a solution for (3.4.1)-(3.4.3) implies the existence of a solution for the integral system (3.8). Let $T^* > 0$ be arbitrary and $t \in [0, T^*]$. Assume there exists some maximal infection-age A^*, and let $A_T = A^* + T^*$. We employ the following assumptions to guarantee the existence of a unique continuous solution for (3.4.3)-(3.4.1):

(H1) The per capita biting rate α and mosquito recruitment rate Λ_v are positive constants.

(H2) Disease transmission probabilities $\beta_{vb}(\tau)$, $\beta_{vh}(\tau)$, and $\beta_{bv}(\tau)$ are nonnegative continuous functions that vanish beyond A_T. That is,

$$\beta_{vb}(\tau) = \beta_{vh}(\tau) = \beta_{bv}(\tau) = 0 \text{ for } \tau > A_T.$$

(H3) The human per capita recovery rate γ_h is a positive constant and the bird per capita recovery rate $\gamma_b(\tau)$ is a nonnegative continuous function.

(H4) The natural per capita death rate for mosquitoes μ_v is a positive constant.

(H5) Initial densities $i_b^0(\tau)$, $i_v^0(\tau)$ are nonnegative, continuous, and vanish beyond $\tau > A^*$.

(H6) Initial and boundary conditions are compatible at $t = 0$. That is,

$$i_b^0(0) = \lambda_{vb}(0)S_b^0, \quad i_v^0(0) = \lambda_{bv}(0)S_v^0.$$

Theorem 1. *The system of integral equations* (3.4.3)-(3.4.1) *has a unique continuous solution.*

The proof of Theorem 1 is in Appendix B.

3.4 Basic Reproduction Number and Local Asymptotic Stability of Disease-Free Equilibrium

The basic reproduction number, $\mathcal{R}_0$, quantifies the average number of secondary infectious hosts produced by one infected host during its infectious period in a wholly susceptible host population. The basic reproduction number for our system is computed as:

$$\mathcal{R}_0 = \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2} \bigg(\int_0^\infty \beta_{vb}(\tau)\Pi_v(\tau)d\tau \bigg) \bigg(\int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau)d\tau \bigg). \tag{3.10}$$

The factor $\frac{\alpha \Lambda_v}{\mu_v (S_b^0 + S_h^0)} \int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau)d\tau$ represents the average number of infectious vectors produced by one infectious bird during its infectious period in a wholly susceptible population. Similarly, the factor $\frac{\alpha S_b^0}{S_b^0 + S_h^0} \int_0^\infty \beta_{vb}(\tau)\Pi_v(\tau)d\tau$ represents the average number of infectious birds produced by one infectious vector during its infectious period in a wholly susceptible population.

Theorem 2. *The disease-free equilibrium is locally asymptotically stable if* $\mathcal{R}_0 < 1$ *and unstable if* $\mathcal{R}_0 > 1$.

The proof of Theorem 2 is in Appendix B. A $\mathcal{R}_0$ less than unity implies that a small introduction of WNV will not produce a sustained outbreak. We note that the epidemiological parameters for human infection do not appear in the $\mathcal{R}_0$ expression because WNV only transmits in one direction – from vectors to humans. Nonetheless, reducing the value of $\mathcal{R}_0$ will in turn reduce the amount of human WNV infections since less WNV-infected mosquitoes will be generated.

3.5 Choice of Infection-Age Dependent Parameter Functions

The model assumes that the vector-to-host transmission probabilities (with "host" meaning birds or humans), bird-to-vector transmission probability from birds to vectors, and bird recovery rate are functions of infection-age τ.

3.5.1 Vector-to-Bird Transmission Probability

Data from Reisen et al. [63] was fitted to a logistic curve to determine the infection-age dependent per-bite probability of vector-to-bird transmission $\beta_{vb}(\tau)$. This data shows the fraction of WNV transmitting *Culex tarsalis* mosquitoes at 22°C as a function of infection-age [63]. The per-bite vector-to-bird transmission probability takes the form:

$$\beta_{vb}(\tau) = \frac{M_1}{1 + e^{-k_1(\tau - \tau_{0,1})}}, \tag{3.11.1}$$

where M_1 is the maximum per-bite transmission probability, $\tau_{0,1}$ is the mean extrinsic incubation period in days, and k_1 is the scale parameter with units of days^{-1} that affects the steepness of the graph of β_{vb}. We assume a baseline value of $M_1 = 1$ and fit $\tau_{0,1}$ and k_1 to the data in [63]. Parameter fitting was carried out using a Levenberg-Marquardt algorithm from [72]. This algorithm uses a nonlinear least squares approach to minimize the mean squared error between the data and model output. Panel (a) of Figure 3.2 shows the results of the data fitting. The fitted values of $\tau_{0,1}$ and k_1 provide baseline estimates and are listed in Table 3.2.

3.5.2 Bird-to-Vector Transmission Probability

Data from Komar et al. [39] and Reisen et al. [63] are used to inform the infection-age-dependent bird-to-vector transmission probability $\beta_{bv}(\tau)$. A unimodal relationship between bird host viremia and infection-age is presented in [39] for ten passerine species infected with WNV. Passerine species (e.g., crows, sparrows, robins, etc.) are known to be highly competent WNV reservoirs and are present throughout North America [39]. A linear relationship between bird-to-vector infection probability and bird host viremia is shown in [63]. To construct a continuous function for $\beta_{bv}(\tau)$, we first use the linear relationship in [63] to transform the host viremia data from [39]. We then fitted a Gaussian function to this transformed data set. The bird-to-vector infection transmission probability takes the form:

$$\beta_{bv}(\tau) = ae^{-\frac{(\tau-b)^2}{2c^2}}, \tag{3.11.2}$$

where a is the maximum transmission probability, b is the time of occurrence of maximum transmission probability in days, and c is the standard deviation in days. We used a trust region reflective algorithm [72] to fit parameters a, b, c to the transformed data from [39, 63]. Similar to the Levenberg-Marquardt algorithm used to fit parameters of (3.11.1), the trust region reflective algorithm is a nonlinear least squares method for optimization. Unlike the Levenberg-Marquardt algorithm, the trust region algorithm accepts bound constraints to restrict the parameter search space. We constrained the parameter search space to $0.85 \leq a \leq 1$, $3 \leq b \leq 5$, and $1 \leq c \leq 2$ to ensure fitted parameters would have biologically realistic values [39]. Panel (b) of Figure 3.2 shows the results from fitting the transformed data. Fitted values of a, b, c provide baseline estimates and are listed in Table 3.2.

3.5.3 Bird Recovery Rate

The bird per capita recovery rate is typically measured through the level of WNV antibody response [84] or the reciprocal of the infectious period [39]. Antibody response is measured as a function of infection-age and has been observed to follow a logistic pattern [84]. A bird is considered recovered from WNV infection when the antibody response exceeds a certain threshold [84]. On the other hand, the bird infectious period measures the length of time viremia is present [39]. It is unclear whether or not antibody response and viremia levels are correlated [84, 39]. However, it may be of interest to understand how the behavior of an infection-age-dependent recovery rate impacts the dynamics of WNV transmission. Here, the bird per capita recovery rate is assumed to have a logistic relationship with infection-age, as seen with the antibody response data in VanDalen et al. [84]:

$$\gamma_b(\tau) = \frac{M_2}{1 + e^{-k_2(\tau - \tau_{0,2})}}. \tag{3.11.3}$$

Data from Komar et al. informs the baseline maximum per capita recovery rate M_2 and the inflection point of the logistic curve $\tau_{0,2}$ [39]. The mean infectious period in days was selected as the baseline value of $\tau_{0,2}$. What is unknown is the scale parameter k_2 that affects how steep the curve is at $\tau_{0,2}$. In panel (c) of Figure 3.2, we show the logistic reovery rate over a range of values of the scale parameter k_2 to display possible variation.

Table 3.2 displays the baseline and ranges of each parameter in $\mathcal{R}_0$ when assuming the functional forms of $\beta_{vb}(\tau)$, $\beta_{bv}(\tau)$, $\gamma_b(\tau)$ in (3.11). Baseline values and ranges for demographic and biting parameters were obtained from experimental and other modeling studies found in extant when possible. Parameters belonging to infection-age-dependent functions had their baseline values fitted from experimental data and

Table 3.2: Baseline Values and Ranges of Parameters in $\mathcal{R}_0$.

Parameter	**Baseline**	**Range**	**Source**
Λ_v	3,776	1,302 – 5,183	[61]; [74]
α	0.5	0.30 – 1.00	[90]; [1]; [87]; [88]; [69]
μ_v	0.066	0.05 – 0.33	[6]; [88]; [23]
M_1	1.00	0.85 – 1.00	Fitted from [63]
$\tau_{0,1}$	15.24	12 – 16	Fitted from [63]
k_1	0.495	0.20 – 0.60	Fitted from [63]
a	1.0	0.85 – 1.0	Fitted from [63]; [39]
b	3.06	2.0 – 5.0	Fitted from [63]; [39]
c	1.41	1.0 – 2.0	Fitted from [63]; [39]
M_2	0.50	0.30 – 0.60	Assumed from [39]
$\tau_{0,2}$	4.50	3.0 – 5.0	Assumed from [39]
k_2	0.50	0.30 – 1.0	Assumed
S_b^0	5,000	1,000 – 10,000	Assumed
S_h^0	50,000	10,000 – 100,000	Assumed

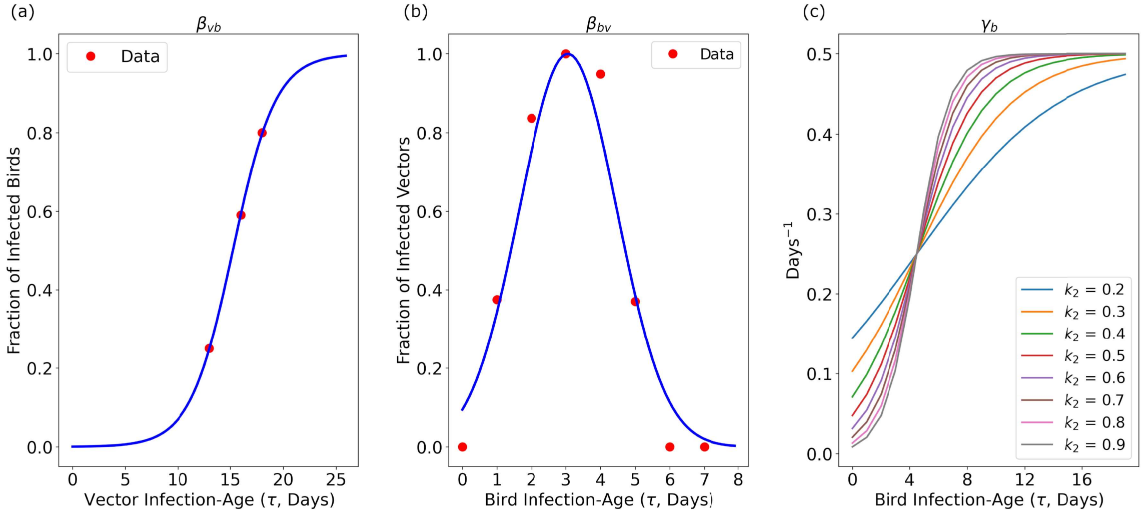


Figure 3.2: Infection-Age-Dependent Parameter Functions. Data Fitting Results for (a) the Vector-to-Bird Transmission Probability Data From [63] to a Logistic Function and (b) the Transformed Bird-to-Vector Transmission Probability Data From [39, 63]. (c) Logistic Bird Per Capita Recovery Rate Over a Range of Values of the Scale Parameter k_2. The Values of the Maximum Recovery Rate and Mean Infectious Period (Inflection Point) Were Informed from Results in [39].

assumed a biologically realistic range. The parameter baseline and range values were used to inform a sensitivity analysis on the basic reproduction number $\mathcal{R}_0$. The sensitivity analysis can provide insight on which demographic and epidemiological parameters have the most influence on the dynamics of WNV outbreaks and inform mitigation or control efforts.

3.6 Sensitivity Analysis

A sensitivity analysis was conducted to understand the impact of demographic and epidemiological parameters on the basic reproduction number $\mathcal{R}_0$. We used the forward normalized sensitivity index on 14 model parameters using the approach in [30].

The normalized forward sensitivity index is the ratio of relative rate of change in a quantity of interest q with respect to parameter of interest p:

$$\gamma_p^q = \frac{\partial q}{\partial p} \times \frac{p}{q}.$$

Our analysis uses $q = \mathcal{R}_0$ and p belongs to the set:

$$\mathbf{p} := \{\Lambda_v, \mu_v, \alpha, M_1, k_1, \tau_{0,1}, M_2, k_2, \tau_{0,2}, a, b, c, S_h^0, S_b^0\}.$$

We detail the process for finding the normalized forward sensitivity index of the basic reproduction number $\gamma_{\mu_v}^{\mathcal{R}_0}$ on the mosquito per capita mortality rate.

To find the partial derivative of $\mathcal{R}_0$ with respect to parameter p, let $A = \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}$, $B = \int_0^\infty \beta_{vb}(\tau)\Pi_v(\tau)d\tau$, and $C = \int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau)d\tau$. Then, the basic reproduction number may be re-written as $\mathcal{R}_0 = ABC$. By the product rule,

$$\frac{\partial \mathcal{R}_0}{\partial p} = A\left(B\frac{\partial C}{\partial p} + C\frac{\partial B}{\partial C}\right) + BC\frac{\partial A}{\partial p}, \tag{3.12}$$

where,

$$\frac{\partial B}{\partial p} = \int_0^\infty \left(\frac{\partial \beta_{vb}(\tau)}{\partial p}\Pi_v(\tau) + \beta_{vb}(\tau)\frac{\partial \Pi_v(\tau)}{\partial p}\right)d\tau, \tag{3.13.1}$$

$$\frac{\partial C}{\partial p} = \int_0^\infty \left(\frac{\partial \beta_{bv}(\tau)}{\partial p}\Pi_b(\tau) + \beta_{bv}(\tau)\frac{\partial \Pi_b(\tau)}{\partial p}\right)d\tau. \tag{3.13.2}$$

For $p = \mu_v$, notice that:

$$\frac{\partial A}{\partial \mu_v} = \frac{-\alpha^2 \Lambda_v S_b^0}{\mu_v^2 (S_b^0 + S_h^0)^2} = \frac{-A}{\mu_v}, \tag{3.14.1}$$

$$\frac{\partial B}{\partial \mu_v} = \int_0^\infty \beta_{vb}(\tau)\frac{\partial \Pi_v(\tau)}{\partial \mu_v}d\tau = -\int_0^\infty \beta_{vb}(\tau)\tau\Pi_v(\tau)d\tau, \tag{3.14.2}$$

$$\frac{\partial C}{\partial \mu_v} = 0. \tag{3.14.3}$$

By substituting (3.14) into (3.12), we obtain:

$$\frac{\partial \mathcal{R}_0}{\partial \mu_v} = -AC\int_0^\infty \beta_{vb}(\tau)\tau\Pi_v(\tau)d\tau - \frac{\mathcal{R}_0}{\mu_v}. \tag{3.15}$$

It follows that the normalized forward sensitivity index of μ_v on $\mathcal{R}_0$ is:

$$\gamma_{\mu_v}^{\mathcal{R}_0} = \frac{\partial \mathcal{R}_0}{\partial \mu_v} \times \frac{\mu_v}{\mathcal{R}_0} = -\frac{\mu_v}{B} \int_0^\infty \beta_{vb}(\tau)\tau\Pi_v(\tau)d\tau - 1. \tag{3.16}$$

Appendix C includes the expressions for the sensitivity indices of $\mathcal{R}_0$ with respect to the remaining parameters of $\mathbf{p}$. The normalized forward sensitivity indices of $\mathcal{R}_0$ with respect to the various parameters can provide valuable information on the impact of control strategies that target WNV. We first compute the baseline value of $\mathcal{R}_0$ using the baseline parameter values found in Table 3.2. We then compute $\gamma_p^{\mathcal{R}_0}$ at the baseline value for each parameter $p \in \mathbf{p}$. Additionally, for each $p \in \mathbf{p}$, we compute the $\mathcal{R}_0$ and $\gamma_p^{\mathcal{R}_0}$ for 20 uniformly sampled points from the range described in Table 3.2. This process allows us to generate response curves of $\gamma_p^{\mathcal{R}_0}$ that show how the relative sensitivities of $\mathcal{R}_0$ with respect to each parameter p change as their values deviate from the baseline. By sampling points from the ranges presented in Table 3.2, we can find the relative sensitivities of $\mathcal{R}_0$ for parameter values that are within a biologically realistic context. Computation of $\mathcal{R}_0$ and $\gamma_p^{\mathcal{R}_0}$ for each parameter sample and baseline value assumed a maximum infection-age of 30 days for both vectors and birds, and used a stepsize of 0.01 days.

3.6.1 Baseline Sensitivity Indices

Table 3.3 shows the baseline sensitivity index for $\mathcal{R}_0$ with respect to each parameter. The most sensitive parameters of $\mathcal{R}_0$ at the baseline value include the mosquito mortality rate (μ_v) and mosquito biting rate (α). At the baseline value, a 1% increase in μ_v results in a 2.36% decrease in $\mathcal{R}_0$. A 1% increase in α results in a 2% increase in $\mathcal{R}_0$ at the baseline value. These results are consistent with what was found in [6, 82]. The top three parameters for the highest sensitivity of $\mathcal{R}_0$ at their baseline values are related to the demographic and biting processes. This highlights the importance of

Table 3.3: Forward Normalized Sensitivity Indices of Parameters of $\mathcal{R}_0$ at Baseline Values as Indicated in Table 3.2. Parameters are Ranked by Sensitivity Magnitude. We See That $\mathcal{R}_0$ is Most Sensitive to the Mosquito Mortality and Biting Rates.

Parameter	$\gamma_p^{\mathcal{R}_0}$		**Parameter**	$\gamma_p^{\mathcal{R}_0}$
μ_v	-2.36		c	0.83
α	2.00		S_b^0	0.82
S_h^0	1.81		$\tau_{0,2}$	0.47
$\tau_{0,1}$	-1.60		b	-0.35
a	1.00		M_2	-0.29
Λ_v	1.00		k_2	0.26
M_1	1.00		k_1	-0.09

quantifying uncertainty of demographic and biting parameters when modeling MBDs like WNV. This further demonstrates the impact that mosquito population control measures may have on epidemic quantities like $\mathcal{R}_0$. The infection-age-dependent function with the greatest sensitivity for $\mathcal{R}_0$ is the mean extrinsic incubation period ($\tau_{0,1}$) that appears in the transmission probability – a 1% increase in $\tau_{0,1}$ results in a 1.6% decrease in $\mathcal{R}_0$.

3.6.2 Sensitivity Results for Infection Transmission and Recovery Parameters

Figure 3.3 shows the basic reproduction number $\mathcal{R}_0$ and the sensitivity response curve $\gamma_p^{\mathcal{R}_0}$ with respect to parameters of the infection-age-dependent functions (3.11). $\mathcal{R}_0$ is proportional to the maximum transmission probabilities M_1 and a. Therefore, sensitivity indices with respect to these parameters always have a value of one – a 1% increase in M_1 or a results in a 1% increase of $\mathcal{R}_0$.

The sensitivity of $\mathcal{R}_0$ to the maximum bird recovery rate M_2, mean intrinsic incubation period b, and mean extrinsic incubation period $\tau_{0,1}$, increases as these parameter values increase. Furthermore, $\mathcal{R}_0$ is a decreasing function of each of these parameters. Small increases at greater values of M_2, b, or $\tau_{0,1}$ result in greater reduction of $\mathcal{R}_0$ compared to the same relative increases at their baseline values. For example, if the mean intrinsic incubation period increases from three days to five days, then a 1% increase in b leads to 1.2% reduction of $\mathcal{R}_0$ compared to only 0.4% reduction at the baseline value. Since increasing the incubation period equates to decreasing the incubation rate, lowering incubation rates can have a substantial impact on lowering $\mathcal{R}_0$.

On the other hand, the sensitivity of $\mathcal{R}_0$ with respect to the scale parameter for vector-to-bird transmission probability k_1 and the standard deviation of bird-to-vector transmission probability c increases as these parameter values decrease. Smaller values of k_1 correspond to larger values of $\mathcal{R}_0$, but increasing k_1 at these smaller values can have a greater impact on reducing $\mathcal{R}_0$ compared to increases at the baseline value. On the contrary, smaller values of c correspond to smaller values of $\mathcal{R}_0$, but slight increases in c at these smaller values can have a greater impact on increasing $\mathcal{R}_0$.

The most interesting sensitivity responses occur for the mean infection period for birds $\tau_{0,2}$ and the scale parameter for the bird recovery rate k_2. The non-monotonic sensitivity behavior results from the negative exponential applied to the bird recovery rate function (3.7.2). Interestingly, sensitivity to $\mathcal{R}_0$ peaks near the baseline value for both $\tau_{0,2}$ and k_2. As values of $\tau_{0,2}$ and k_2 deviate from the baseline, changes in these parameters have less of an impact on $\mathcal{R}_0$. Although parameters of the bird recovery rate function show the most interesting sensitivity responses, they are generally the least impactful on $\mathcal{R}_0$.

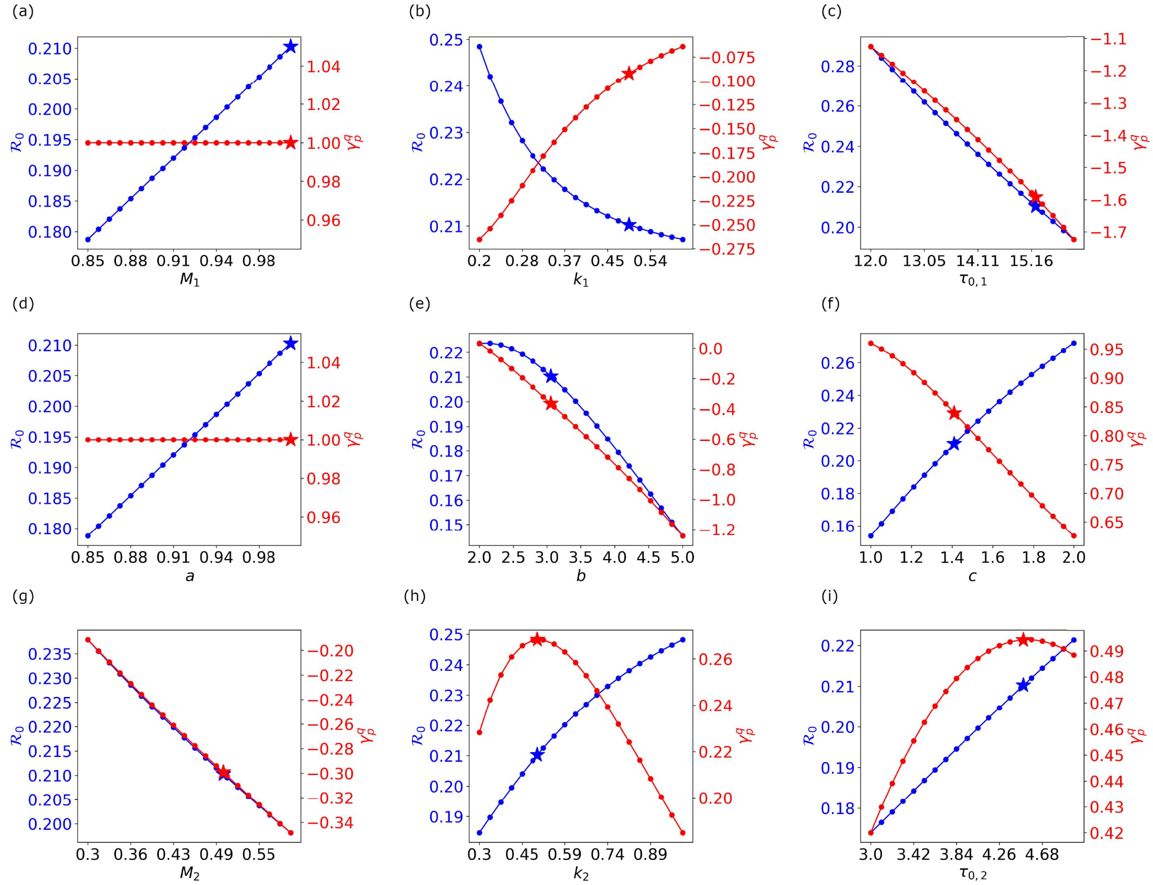


Figure 3.3: The Basic Reproduction Number $\mathcal{R}_0$ (Blue Lines) and Sensitivity Response Curves (Red Lines) as a Function of Transmission and Recovery Parameters of $\mathcal{R}_0$. Starred Values Indicate the Baseline $\mathcal{R}_0$ and Normalized Forward Sensitivity Index. The Top Row Shows Parameters for the Transmission Probability of Vectors to Hosts, $\beta_{vb}(\tau)$. The Middle Row Shows Parameters for the Transmission Probability of Birds to Vectors, $\beta_{bv}(\tau)$. The Bottom Row Shows Parameters for the Bird Recovery Rate, $\gamma_b(\tau)$.

3.6.3 Sensitivity Results for Demographic and Biting Parameters

Figure 3.4 shows the basic reproduction number $\mathcal{R}_0$ and the sensitivity response curve $\gamma_p^{\mathcal{R}_0}$ with respect to the demographic and biting parameters of (3.3). The sensitivity of $\mathcal{R}_0$ to the mosquito recruitment rate Λ_v and mosquito biting rate α are constant as a result of $\mathcal{R}_0$ depending linearly on Λ_v and quadratically on α.

Changes in the initial susceptible human population S_h^0 at greater values result in greater changes to $\mathcal{R}_0$ compared to the same relative changes at the baseline value. The model assumes that mosquitoes are equally likely to seek out a bloodmeal from humans as they are from birds. Increasing the human population decreases the fraction of birds in the host population. Since humans do not transmit WNV back to mosquitoes and their infection parameters do not appear in the $\mathcal{R}_0$ expression, increasing S_h^0 will decrease $\mathcal{R}_0$ values. Alternatively, increasing the initial susceptible bird population S_b^0 will increase $\mathcal{R}_0$, but with decreasing impact from changes at greater values of S_b^0.

The mosquito mortality rate μ_v is by far the most impactful parameter on $\mathcal{R}_0$ across all ranges and parameters assessed. This is likely due to μ_v being involved with both the steady state initial vector population (recall, $S_v^0 = \frac{\Lambda_v}{\mu_v}$) and the mosquito survival probability $\Pi_v(t)$. All other parameters explored belong to a single processes – i.e., they appear in exactly one of A, B, or C of (3.12), whereas μ_v appears in both A and B expressions. A 1% increase of μ_v when $\mu_v = 0.33$ results in a reduction of $\mathcal{R}_0$ by over 5.22%. The mosquito mortality rate is the reciprocal of the mosquito lifespan. Therefore, decreasing the mosquito lifespan equates to increasing its mortality rate. Reducing the mosquito lifespan will have the greatest impact on reducing $\mathcal{R}_0$.

3.7 Discussion

This chapter introduces a PDE model for WNV transmission between mosquito vectors, bird hosts, and human dead-end hosts. The model includes infection-age-dependent processes for both mosquito vectors and bird hosts, and we demonstrate the impact of demographic and epidemiological parameters on WNV transmission through sensitivity analysis. As far as we know, this is the first PDE model to include infection-age dependence on both vector and host populations. The model assumes the vector-to-host transmission probability $\beta_{vb}(\tau)$, bird-to-vector transmission probability $\beta_{bv}(\tau)$, and bird recovery rate $\gamma_b(\tau)$ are functions of infection-age

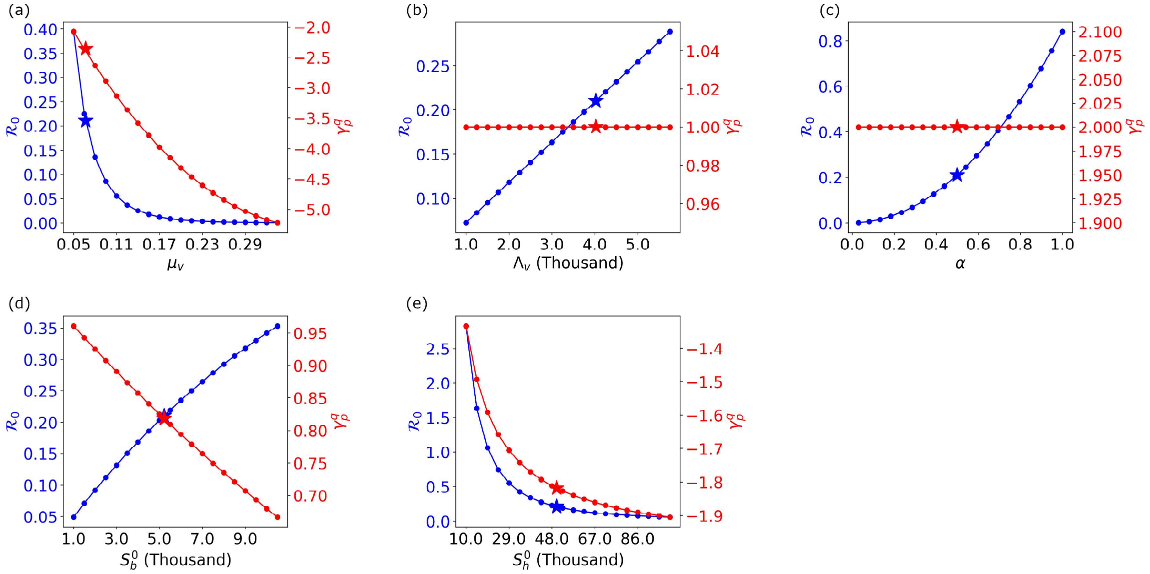


Figure 3.4: The Basic Reproduction Number $\mathcal{R}_0$ (Blue Lines) and Sensitivity Response Curves (Red Lines) as Functions of Demographic and Biting Parameters of $\mathcal{R}_0$. Starred Values Indicate the Baseline $\mathcal{R}_0$ and Normalized Forward Sensitivity Index.

τ. We proved that the model has a unique continuous solution under biologically realistic assumptions for the infection-age-dependent processes. Moreover, we found an analytic expression for the basic reproduction number $\mathcal{R}_0$ and showed that the disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 < 1$. We assumed logistic functions for $\beta_{vb}(\tau)$ and $\gamma_b(\tau)$ and a Gaussian function for $\beta_{bv}(\tau)$. Parameters for the infection-age-dependent processes were fitted to experimental data when available or assumed biologically reasonable ranges. We determined the forward normalized sensitivity indices for parameters of the basic reproduction number $\mathcal{R}_0$ for our sensitivity analysis. The sensitivity analysis uses the aforementioned assumptions on functions $\beta_{vb}(\tau)$, $\beta_{bv}(\tau)$, and $\gamma_b(\tau)$. Under these assumptions, the mosquito mortality and biting rates were found to have the greatest impact on $\mathcal{R}_0$. Other families of functions for the infection-age-dependent processes can be explored and may provide different sensitivity results [30].

Our sensitivity analysis results suggest that control strategies for mitigating WNV transmission should focus on increasing the mosquito mortality rate and decreasing the mosquito biting rate. Increasing the mosquito mortality rate may be achieved through the use of adulticides [13]. Humans can decrease their risk of being bitten by a mosquito by using mosquito repellent or insecticide nets. However, all of these strategies may bring about unintended negative consequences, as use of these control measures have shown to increase insecticide resistance in mosquitoes [66]. We also note that the model assumes that mosquitoes are equally likely to bite birds as they are humans. Decreasing the biting rate of mosquitoes on humans may not decrease the biting rate overall, since mosquitoes will then resort to biting birds or other mammals when a human bloodmeal is unavailable. Some modeling studies have explored the seasonal shift in mosquito's biting preference between birds and humans [23]. Overall, identifying the best control strategy for minimizing WNV and other MBD risk is a nontrivial challenge with no straightforward solution. This issue is compounded further due to the effect that weather and climate have on MBD dynamics [5] In view of the explanation provided above for the mosquito mortality rate, it is likely that this parameter will be the most important in terms of curtailing infection prevalence and incidence.

The next step is to incorporate explicit weather dependence on WNV transmission. Studies have evaluated temperature dependence on the mosquito mortality rate [63], which happens to be the most influential parameter on $\mathcal{R}_0$ in our model. Others have shown temperature dependence on the extrinsic incubation period [16]. Interestingly, the WNV infection and recovery parameters that had the most impact on $\mathcal{R}_0$ for our model were the length of the incubation and infection periods. The following chapter will introduce our extended WNV model that includes temperature-dependent parameter functions, more realistic demographic processes, and connect the model to true WNV human case data. We also attempt to explore possible future

WNV incidence under different climate change scenarios.

Chapter 4

A NON-AUTONOMOUS WEST NILE VIRUS MODEL

4.1 Temperature-Dependent Processes on WNV Transmission

Mosquitoes are incapable of adequately regulating their body temperature due to their ectothermic nature. Consequently, environmental temperature impacts both mosquito lifespan [16] and mosquito-borne disease (MBD) transmission [63]. It is critical to include temperature dependence in mosquito-borne disease models like West Nile Virus (WNV), particularly for assessing the impact of climate change on MBD transmission. Recent approaches for modeling MBD transmission under climate change have included both mechanistic [82, 54] and statistical models [2]. Trejo et al. included a temperature-dependent extrinsic incubation period (EIP) and mosquito lifespan in a dengue transmission model and found that increasing temperatures under climate change alters both dengue risk and efficacy of control measures [82]. Moser et al. explored the effect of life history traits of *Culex* mosquitoes on a temperature-dependent basic reproduction number for WNV [54]. The authors found that climate change may cause WNV transmission by *Cx. pipiens* to increase along the northern part of the US and decrease for the southern part of the US, while WNV transmission by *Cx. quinquefasciatus* is expected to increase in most of the US [54]. Albrecht & Kaufeld found that environmental temperature is a strong predictor of human WNV cases in Ontario, Canada using a zero-inflated Poisson model with lagged temporal covariates [2]. All studies highlighted the nonlinear effects of temperature on MBD transmission and noted how these effects may vary based on characteristics of particular geographical regions.

Here we extend our infection-age-dependent WNV model to include temperature dependent processes for mosquito mortality and transmission probability. Recall that the mosquito mortality rate μ_v was the single most influential parameter on the basic reproduction number $\mathcal{R}_0$ of the simplified WNV PDE model (3.3), and the extrinsic incubation period (EIP) $\tau_{0,1}$ was the most influential parameter on $\mathcal{R}_0$ related to infection transmission. $\mathcal{R}_0$ was also moderately sensitivite to the maximum vector-to-host transmission probability per bite M_1. The EIP and maximum transmission probability appear in the expression for the per-bite probability of vector-to-host transmission β_{vb}. Both the average adult lifespan of mosquitoes and EIP decrease as a function of temperature [16, 63], while the maximum transmission probability is a unimodal function of temperature [43].

The temperature-dependent mosquito mortality rate is modeled using results from Ciota et al., which presents the longevity of various *Culex* species as a linear function of temperature [16]. The reciprocal of the mosquito lifespan gives the per capita mortality rate as a function of temperature T:

$$\mu_v(T) = \begin{cases} \frac{1}{-4.57T+157.6}, & T < 34.48°\text{C}, \\ 37.8, & \text{otherwise.} \end{cases} \tag{4.1.1}$$

The expression for the temperature-dependent EIP is found by taking the multiplicative inverse of the extrinsic incubation rate presented in Reisen et al. [63]:

$$\tau_{0,1}(T) = \begin{cases} \frac{1}{0.0093T-0.1352}, & T > 15°\text{C}, \\ 232, & \text{otherwise.} \end{cases} \tag{4.1.2}$$

The temperature-dependent maximum vector-to-host transmission probability is modeled from the Briére relationship presented in Lambrechts et al. [43]:

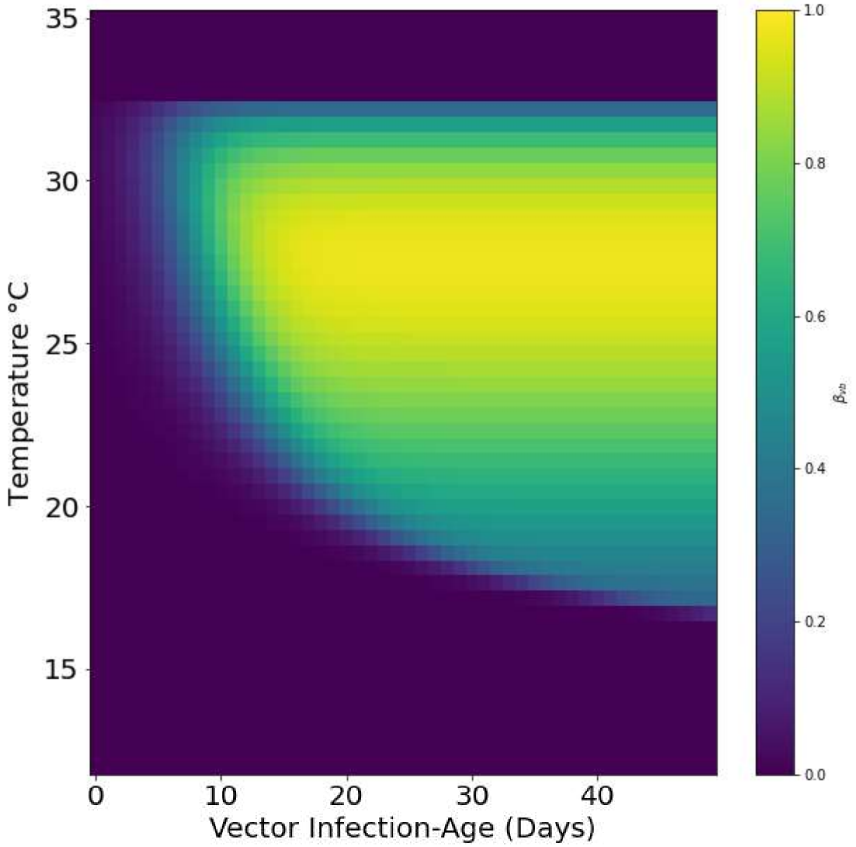


Figure 4.1: Heatmap of the Per-Bite Probability of Vector-to-Host Transmission with Respect to Infection-Age and Temperature. Yellow Values Indicate the Highest Probability of Infection Transmission.

$$M_1(T) = \begin{cases} 0.001044T(T - 12.2586)\sqrt{32.461 - T}, & 12.2586 \leq T \leq 32.461, \\ 0, & \text{otherwise}. \end{cases} \tag{4.1.3}$$

The temperature-dependent EIP and maximum transmission probability are used in the per-bite probability of vector-to-host transmission $\beta_{vb}(\tau, T)$, which is now both infection-age and temperature-dependent,

$$\beta_{vb}(\tau, T) = \frac{M_1(T)}{1 + e^{-k_1(\tau - \tau_{0,1}(T))}} \tag{4.2}$$

Figure 4.1 depicts the transmission probability with respect to temperature and vector infection-age. The highest probability of transmission occurs beyond vector infection-age of 15 days and between 25-28°C.

The above expressions for $\mu_v(T)$ and $\tau_{0,1}(T)$ indicate that suitable temperatures for WNV transmission occur within the range 15.8–33.2°C where the EIP is less than the adult mosquito lifespan. Under global increases in average temperature, certain areas are expected to become too hot for mosquito survival while others become more suitable for longer WNV seasonal transmission [54]. Incorporating temperature dependence on μ_v, $\tau_{0,1}$, and M_1 allows us to assess the impact of climate change scenarios on WNV transmission.

4.2 Time-Dependent Processes on WNV Transmission

Before we assess the impact of climate change scenarios on WNV transmission, we must first understand the role that demographic processes have on current WNV transmission. Both mosquito and bird populations are inherently time-varying with inter-annual and inter-seasonal variation [49]. Many current mechanistic MBD models neglect to include time-varying populations as a simplifying model assumption or due to lacking reliable data for parameterizing demographic processes [5, 52]. Some mechanistic models for WNV that do incorporate time-varying populations have modeled juvenile mosquito processes [44, 88] and fitted bird populations to observed relative frequency of WNV-competent birds [44, 69]. In addition to adding temperature-varying processes for mosquito mortality and disease transmission, we also extend the simplified WNV model (3.3) by incorporating time-varying demographic processes for vector and bird populations that are parameterized by data from the Greater Toronto Area (GTA). The total population of mosquitoes $N_v(t)$ and birds $N_b(t)$ follow the non-autonomous differential equations:

$$\frac{dN_v(t)}{dt} = r_v(t)N_v(t)\left(1 - \frac{N_v(t)}{K_v(t)}\right), \tag{4.3.1}$$

$$\frac{dN_b(t)}{dt} = r_b(t)N_b(t), \tag{4.3.2}$$

where, in order to model seasonality, we let

$$r_v(t) = r - r_s \cos\left(\frac{2\pi t}{365}\right), \tag{4.4.1}$$

$$K_v(t) = K - K_s \cos\left(\frac{2\pi t}{365}\right), \tag{4.4.2}$$

$$r_b(t) = \phi - \phi_s \cos(\omega(t + \Delta t)). \tag{4.4.3}$$

The time-varying mosquito population $N_v(t)$ follows a non-autonomous logistic model with periodically-varying net growth rate $r_v(t)$ and carrying capacity $K_v(t)$. The components of $r_v(t)$ and $K_v(t)$ are determined from fitting the logistic linkage model to the total mosquito process-based model (PBM) time series for the Greater Toronto Area [49]. Seasonal (yearly) values of r, r_s, K, and K_s are fitted to the PBM to capture the inter-annual variation in mosquito population. Further details about the fitting process are found in [49].

The time-varying bird population $N_b(t)$ assumes a periodically-varying net growth rate $r_b(t)$ to capture the population recruitment via hatching and migration. Seasonal values of the baseline net growth rate ϕ, amplitude scaling factor ϕ_s, and frequency ω are fitted to relative abundance data for WNV-competent birds in the GTA. Relative abundance is defined as the proportion of birds that are WNV-competent. The WNV-competent birds in the GTA and include American Crows, American Robins, Blue Jays, Common Grackles, European Starlings, House Finches, House Sparrows, Northern Cardinals, Red-wing Blackbirds, Ring-billed Gulls, and Song Sparrows [39]. Relative abundance data was obtained from eBird [19] and a Gaussian Process Regressor (GPR) [62] was applied on a weekly timescale to approximate daily time

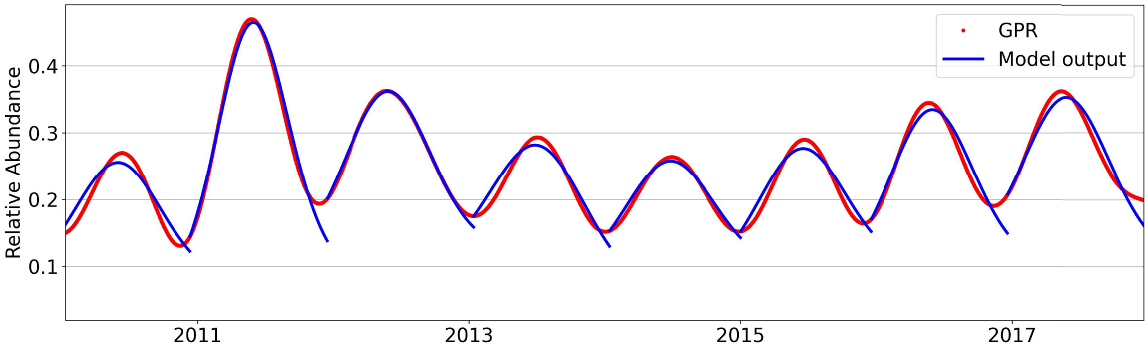


Figure 4.2: Results of Fitting the Bird Net Growth Rate $r_b(t)$ to the GPR Time Series of WNV-Competent Bird Relative Abundance in the GTA for 2010-2017.

series values. Fitting of parameter components of $r_b(t)$ follow a similar process to the fitting of mosquito parameters $r_v(t)$ and $K_v(t)$. The parameter Δt denotes the time in days between the beginning of the bird and mosquito ecological seasons. For the GTA, the ecological season for WNV competent birds begins in December [19], whereas the mosquito ecological season starts around May [49]. Figure 4.2 shows the result of fitting the bird net growth rate $r_b(t)$ to the GPR time-series. Table D.3 includes the yearly fitted parameter values of ϕ, ϕ_s, and ω for years 2010-2017. More details about the data sources used for parameterizing the bird population are found in Section 4.4.2.

4.3 Full Mathematical Model

The population compartments are as previously defined (3.1)–(3.2). The full model consists of the following partial-ordinary differential equation system:

$$\frac{dS_v}{dt} = r_v(t)N_v(t)\Big(1 - \frac{N_v(t)}{K_v(t)}\Big) - \lambda_{bv}(t)S_v(t), \tag{4.5.1}$$

$$\frac{\partial i_v}{\partial t} + \frac{\partial i_v}{\partial \tau} = -\mu_v(T)i_v(\tau, t), \tag{4.5.2}$$

$$\frac{dS_b}{dt} = r_b(t)N_b(t) - \lambda_{vb}(t)S_b(t), \tag{4.5.3}$$

$$\frac{\partial i_b}{\partial t} + \frac{\partial i_b}{\partial \tau} = -(\gamma_b(\tau) + \delta_b + \mu_b)i_b(\tau, t), \tag{4.5.4}$$

$$\frac{dR_b}{dt} = \int_0^\infty \gamma_b(\tau) i_b(\tau, t) d\tau - \mu_b R_b(t), \tag{4.5.5}$$

$$\frac{dS_h}{dt} = \Lambda_h S_h(t) - \lambda_{vh}(t) S_h(t), -\mu_h S_h(t) \tag{4.5.6}$$

$$\frac{dI_h}{dt} = \lambda_{vh}(t) S_h(t) - \gamma_h I_h(t) - \delta_h I_h(t) - \mu_h I_h(t), \tag{4.5.7}$$

$$\frac{dR_h}{dt} = \gamma_h I_h(t) - \mu_h R_h(t). \tag{4.5.8}$$

Adult vectors are recruited into their susceptible population at a periodically-varying logistic rate and die naturally at temperature-dependent per capita rate $\mu_v(T)$. Vectors become infected at per capita rate $\lambda_{bv}(t)$. Birds are recruited into their susceptible population at a periodically-varying per capita net growth rate, $r_b(t)$, die naturally at constant per capita rate μ_b, and become infected at per capita rate $\lambda_{vb}(t)$. Infected birds with infection-age τ either recover from infection at per capita infection-age-dependent rate $\gamma_b(\tau)$, die from disease-induced mortality at constant per capita rate δ_b, or die naturally at constant per capita rate μ_b.

Humans are recruited into their susceptible population at constant rate Λ_h and die naturally at constant per capita rate μ_h. Humans become infected per capita at rate $\lambda_{vh}(t)$. Infected humans can recover from infection at constant per capita rate γ_h, die from disease-induced mortality at constant per capita rate δ_h, or die naturally at constant per capita rate μ_h.

The expressions for the forces of infection are as follows:

$$\lambda_{vb}(t) = \frac{\alpha_b \eta}{N_b(t) + N_h(t)} \int_0^\infty \beta_{vb}(\tau, T) i_v(\tau, t) d\tau, \tag{4.6.1}$$

$$\lambda_{vh}(t) = \frac{\alpha_h \eta}{N_b(t) + N_h(t)} \int_0^\infty \beta_{vh}(\tau, T) i_v(\tau, t) d\tau, \tag{4.6.2}$$

$$\lambda_{bv}(t) = \frac{\alpha_b \eta}{N_b(t) + N_h(t)} \int_0^\infty \beta_{bv}(\tau) i_b(\tau, t) d\tau, \tag{4.6.3}$$

where $\lambda_{vb}(t)$ represents the force of infection from vectors to birds, $\lambda_{vh}(t)$ is the force of infection from vectors to humans, and $\lambda_{bv}(t)$ is the force of infection from birds to vectors. Moreover, $\beta_{vb}(\tau, T)$ denotes the per-bite probability of disease transmission from vectors with infection-age τ to birds, at temperature T. Similarly, $\beta_{vh}(\tau, T)$ denotes the probability of disease transmission from vectors with infection-age τ to humans at, temperature T, and $\beta_{bv}(\tau, t)$ denotes the probability of disease transmission from birds with infection-age τ to vectors at time t. Recall in Equation (3.4) of the simplified WNV model that α represented the per capita biting rate on any hosts (birds and humans). For the full model, we assume different vector biting rates between birds and humans, which is a more realistic assumption. We let $\alpha = \alpha_b + \alpha_h$, where α_b is the per capita biting rate on birds and α_h is the per capita biting rate on humans. Let $p = \alpha_b/\alpha$ be the the percentage of vector bites on birds. Then $\alpha_b = p\alpha$ and $\alpha_h = (1-p)\alpha$. Further, we include the modification parameter η to tune the vector-to-host ratio. The total vector population obtained from the mosquito PBM [74] is a structure estimate from trap data and vector lifecycle knowledge. The demographic parameters for the total vector population were obtained from the logistic linkage model in Chapter 2 of this dissertation and in [49] and construct a continuous signal from a discrete time series. However, there still remains uncertainty of the exact translation from the observed trap data to the true vector population. This uncertainty is compounded from the lack of reliable data about the true total bird population. The parameter η tunes the vector-to-host ratio to allow numerical simulations to align with observed human case data.

Table 4.1 includes the definitions of the parameters of the full model. Table 4.2 contains the definitions of the time-varying parameter components. Table 4.3 shows the definitions of the parameters involved with infection-age-dependent processes. Values and sources for the parameters of the full model and infection-age varying processes are listed in Tables D.1 and 4.3, respectively. The initial and boundary conditions of the full model are as before (3.5.1)–(3.5.2). The following assumptions pertain to the full model (4.5):

- For simplicity, the model only considers vector bites on humans and birds. Vector bites on other mammals are neglected.
- Vectors transmit infections to both birds and humans, but only birds can transmit infection back to vectors due to the low amounts of viremia found in infected human blood [15]. The probability of infection transmission depends on both the infection-age of the vector or bird and environmental temperature.
- Due to the short lifespan of the vectors, we assume that infected vectors remain infectious until death.
- Infected birds and humans may either recover from infection, die from disease-induced mortality, or die from natural causes. The per capita recovery rate is infection-age-dependent for birds and constant for humans. Natural and disease-induced per capita death rates are constant for both birds and humans.
- Despite evidence showing that vertical transmission can occur in laboratory settings [3], we neglect vertical transmission in this model. All recruited populations are assumed to be susceptible to infection.
- Due to the lack of conclusive evidence surrounding WNV re-infection, it is assumed that humans and birds who have recovered from infection maintain lifelong immunity against future WNV infection.

There is currently no method to assess the qualitative dynamics or stability of the full model (4.5) due to the time dependence of various parameters. However, we expect the qualitative results to be the same as those in the simplified model (3.3), but the addition of temperature and infection-age dependence will change the

Table 4.1: Parameters of the Full Model (4.5).

Parameter	**Description**	**Dimension**
$r_v(t)$	Vector net per capita growth rate	days^{-1}
$K_v(t)$	Vector carrying capacity	vectors
$r_b(t)$	Bird net per capita growth rate	days^{-1}
Λ_h	Recruitment rate of humans	days^{-1}
α_b	Per capita vector biting rate on birds	bites/(vector×day)
α_h	Per capita vector biting rate on humans	bites/(vector×day)
η	Modification parameter for scaling vector-host ratio	—
$\beta_{vb}(\tau, T)$	Per-bite probability of disease transmission from infected vectors with infection-age τ to susceptible birds, at temperature T	bites^{-1}
$\beta_{vh}(\tau, T)$	Per-bite probability of disease transmission from infected vectors with infection-age τ to susceptible humans, at temperature T	bites^{-1}
$\beta_{bv}(\tau, t)$	Per-bite probability of disease transmission from infected birds with infection-age τ to susceptible vectors at time t	bites^{-1}
$\gamma_b(\tau)$	Per capita recovery rate of infected birds with infection-age τ	days^{-1}
γ_h	Per capita recovery rate of infected humans	days^{-1}
$\mu_v(T)$	Natural per capita death rate of vectors, at temperature T	days^{-1}
μ_b	Natural per capita death rate of birds	days^{-1}
μ_h	Natural per capita death rate of humans	days^{-1}
δ_b	Disease-induced per capita death rate of infected birds	days^{-1}
δ_h	Disease-induced per capita death rate of humans	days^{-1}

quantitative outcome of the simulations. By fitting time, temperature, and infection-age-dependent parameters to existing data, we hope to increase the accuracy of future outbreak predictions compared with a model without temperature or infection-age structure. The following section discusses the data sources used for quantifying the total vector, bird, and human populations.

Table 4.2: Components of Time-Dependent Parameters.

Parameter	**Description**	**Dimension**
r	Baseline vector per capita net growth rate	days^{-1}
r_s	Amplitude scaling factor for vector per capita net growth rate	days^{-1}
K K_s	Baseline vector carrying capacity Amplitude scaling factor for vector carrying capacity	vectors days^{-1}
ϕ	Baseline bird per capita net growth rate	days^{-1}
ϕ_s	Amplitude scaling factor for bird per capita net growth rate	days^{-1}
ω	Frequency of bird per capita recruitment rate	days^{-1}
Δt	Time lag between vector and bird ecological seasons	days

Table 4.3: Parameters of Infection-Age-Dependent Processes.

Parameter	**Description**	**Dimension**
$M_1(T)$	Maximum vector-to-host transmission probability	—
k_1	Scaling parameter for vector-to-host transmission probability	days^{-1}
$\tau_{0,1}(T)$	Extrinsic incubation period	days^{-1}
a	Maximum bird-to-vector transmission probability	—
b	Mean intrinsic incubation period	days
c	Standard deviation of intrinsic incubation period	days
M_2	Maximum per capita bird recovery rate	days^{-1}
k_2	Scaling parameter for bird per capita recovery rate	days^{-1}
$\tau_{0,2}$	Mean duration of bird viremia	days

4.4 Data Sources for Total Populations

We use data sources from the Greater Toronto Area (GTA) for years 2010-2017 to parameterize the total vector, bird, and human populations. The GTA consists of the five Ontario public health units: Durham, Halton, Toronto, Peel, and York.

4.4.1 *Total Mosquito Population*

The mosquito process-based model (PBM) by Shutt et al. uses daily temperature, water data, daylight hours, and mosquito lifecycle knowledge to generate a daily time series from the sparse mosquito trap data from Public Health Ontario [74, 61]. Yearly parameters of the mosquito net per capita growth rate and carrying capacity were fitted to the PBM time series. That is, r, r_s, K, and K_s values of Equations (4.4.1) and (4.4.2) were fitted for each year of the PBM time series. The time frame of the mosquito season is also determined through the parameter fitting optimization process, and the initial condition $N_v(0)$ becomes the value of the PBM time series on the first day of the mosquito season. More details about the fitting process are found in Chapter 2 of the dissertation and in [49].

4.4.2 *Total Bird Population*

To estimate the total number of WNV-competent birds in the GTA for each year, we multiply the relative abundance of the WNV-competent birds in the GTA by the total number of birds in the GTA.

Relative abundance is defined as the proportion of birds that are WNV competent. Relative abundance data was obtained from eBird [19]. eBird is a citizen science database managed by the Cornell Lab of Ornithology. Anyone can contribute to eBird data by creating an online profile and submitting information about their bird sightings. The eBird database contains documentation of bird distribution and abundance throughout time. Over 100 million bird sightings are submitted annually by eBirders across the globe. Due to the unstructured and sparse nature of the eBird data, a Gaussian process regressor (GPR) [62] was applied to the eBird dataset at a monthly timescale to produce a daily time series of relative abundance estimates for 2010-2017. This process is similar to how the mosquito PBM was used to transform the mosquito trap data to a daily time series, but mechanistic information is not taken into account for the GPR. Yearly values for the bird net growth rate were found from fitting ϕ, ϕ_s, and ω of Equation (4.4.3) to the GPR time series. Parameter fitting uses a similar process to the one used to find the time-dependent parameters of the total mosquito population.

Data to estimate the total number of birds in the GTA was obtained from the North American Breeding Bird Survey (BBS) [11, 95]. The BBS was created in 1966 to monitor the effect of DDT usage on bird populations. It is now used to inform researchers and wildlife managers of significant changes in population levels due to habitat loss and land-use changes. It is jointly monitored by the US Geological Survey's Eastern Ecological Science Center, Environmental Canada's Canadian Wildlife Service, and Mexico's National Commission for the Knowledge and Use of Biodiversity. Data for BBS is collected via roadside surveys once a year by volunteers skilled in bird identification. Three-minute point counts are conducted by a single observer, and the same observer records data for each of the 50 stops spaced 0.5 miles apart on a survey route. There are over 4,100 survey routes across North America. At each stop, every bird seen or heard within a 0.25 mile (400m) radius is recorded and identified by species during the three-minute period. Population estimates are determined by multiplying the BBS bird density by the area of the GTA [7]. The BBS bird density is found by dividing the BBS value of the GTA by the area sampled from

Table 4.4: Breeding Bird Survey Estimates (BBS Values) of All Birds in the Greater Toronto Area (GTA) for Years 2010-2017 [11]. Total Birds Estimate the Entire Bird Population in the GTA as the Product of the BBS Bird Density and the GTA Region Area. Initial Relative Abundance (RA) is the Proportion of WNV-Competent Birds in the GTA from eBird Data After Applying the Gaussian Process Regressor [19, 62] on the First Day of the Mosquito Season. Initial Bird Population $N_b(0)$ is Found by Multiplying the Total Birds by the Initial RA.

Year	**BBS Value**	**Total Birds (Millions)**	**Initial RA**	$N_b(0)$ **(Millions)**
2010	3,043	8.67	0.25	2.19
2011	3,051	8.69	0.46	4.01
2012	3,098	8.83	0.36	3.19
2013	3,366	9.59	0.27	2.59
2014	3,064	8.73	0.26	2.24
2015	2,672	7.61	0.28	2.10
2016	2,259	6.43	0.33	2.13
2017	2,083	5.94	0.35	2.09

the survey stops. Five survey stops are located within the GTA region, and each stop has a 0.4km-radius, producing a sample area of 2.5km^2. This bird density is then multiplied by 7,123.64km^2, which is the total area of the GTA [26]. This yearly value provides the total number of birds in the GTA.

The initial condition of the WNV-competent bird population $N_b(0)$ is found by multiplying the relative abundance value of the GPR time series on the first day of the mosquito season by the estimated total number of all birds in the GTA from BBS. Values of yearly reported BBS values, estimated total bird population, initial relative abundance of WNV-competent birds, and initial WNV-competent birds are outlined in Table 4.4.

4.4.3 Total Human Population

Yearly estimates of human population data were not available for years 2010-2017. We therefore use the population data from the Canadian Census from 2021 to estimate the human population $N_h(t)$ for each year of simulation [27]. The sum of the human populations from the five Ontario public health units of the GTA result in over 6.7 million humans. This value is used as the initial condition $N_h(0)$ for each year of simulation.

4.5 Data Fitting and Simulations

4.5.1 Fitting to Human Case Data

Human case data from Public Health Ontario [61] is used to parameterize the per capita vector biting rate on humans α_h for years 2010-2017. We use a least squares method in Python to fit the yearly per capita vector biting rate on humans, similar to the method that was used to fit the demographic parameters for the vector and bird populations (Sections 4.4.1 and 4.4.2). The cost function is the sum of squared errors between the cumulative new human infections and the yearly estimated human cases. Yearly estimated human cases account for both symptomatic and asymptomatic infections. Nearly 80% of human WNV infections are assumed to be asymptomatic. We assume that the reported value of human cases reflects the true number of symptomatic infections. Dividing this value by 0.2 gives the estimated value of total infections. Therefore, the fitted vector biting rate on humans accounts for both symptomatic and asymptomatic cases. Similar to the process of finding yearly parameters for the vector and bird populations, we fit a separate α_h value for each year of fitting.

4.5.2 Numerical Method

Numerical simulations were carried out in Python. Simulations were run for years 2010-2017 using the GTA data. The duration of simulation T^* is the length between the first day of the mosquito season (Start Date indicated in Table B.2) and December 1st of the indicated year in days. We use a step size of $h = 0.1$ days and assume an equal number of time and infection-age steps for each year. That is, $T^* = A^*$, and the total number of time steps $N = T^*/h$ equals the total number of infection-age steps M. We use daily temperature data from ERA5 [32].

Yearly parameters for the time-varying vector and bird populations are found in Tables B.2 and D.3. For fitting α_h, we assume an initial estimate of 0.001 bites/(vector×day), and a vector-to-host ratio modification of $\eta = 25$. All other parameter values are described in Tables 4.1 and 4.3.

For each year, the total vector population $N_v(0)$ was initialized using values in Table B.2, the total bird population $N_b(0)$ was initialized using the relative abundance values in Table D.3 multiplied by the estimated number of all birds from BBS data [11], and the total human population $N_h(0)$ was initialized to 6.7×10^6 people.

For initializing infectious compartments, we assumed a single vector with infection-age zero to provide the initial infection. All other infection-age-structured compartments for the vector and birds were initialized to zero. Susceptible vectors were initialized to $S_v(0) = N_v(0) - 1$, $S_b(0) = N_b(0)$, and $S_h(0) = N_h(0)$. The recovered compartments were initialized to zero.

Trapezoidal quadrature is used to determine the force of infection at the time boundary $n = 0$. For notational convenience, let $U_j^n \approx i_b(jh, nh)$ denote the numerical approximation of infected bird infection-age-density at time nh with infection-age jh, and let $V_j^n \approx i_v(jh, nh)$ denote the numerical approximation of infected vector infection-age-density at time nh with infection-age jh. Further, let $x^n \approx \lambda_{vb}(nh)$ be the force of infection from vectors to birds at time nh, $y^n \approx \lambda_{bv}(nh)$ be the force of infection from birds to vectors at time nh, and $z^n \approx \lambda_{vh}(nh)$ be the force of infection from vectors to humans at time nh. Then the following estimates x, y, and z at the time boundary $n = 0$:

$$x^0 = \frac{\alpha_b \eta}{N_b^0 + N_h^0}\left(\frac{h}{2}\right)\left(\beta_{vb,0}^0 V_0^0 + 2\sum_{j=1}^{M-1} \beta_{vb,j}^0 V_j^0 + \beta_{vb,M}^0 V_M^0\right), \tag{4.7.1}$$

$$y^0 = \frac{\alpha_b \eta}{N_b^0 + N_h^0}\left(\frac{h}{2}\right)\left(\beta_{bv,0} U_0^0 + 2\sum_{j=1}^{M-1} \beta_{bv,j} U_j^0 + \beta_{bv,M} U_M^0\right), \tag{4.7.2}$$

$$z^0 = \frac{\alpha_h \eta}{N_b^0 + N_h^0}\left(\frac{h}{2}\right)\left(\beta_{vh,0}^0 V_0^0 + 2\sum_{j=1}^{M-1} \beta_{vh,j}^0 V_j^0 + \beta_{vh,M}^0 V_M^0\right), \tag{4.7.3}$$

where $\beta_{vb,j}^n$ is the per-bite vector-to-bird transmission probability at time nh (temperature T^n) and vector infection-age jh, $\beta_{bv,j}$ is the per-bite bird-to-vector transmission probability at bird infection-age jh, and $\beta_{vh,j}^n$ is the per-bite vector-to-human transmission probability at time nh (temperature T^n) and vector infection-age jh.

The following steps occur for the remaining time steps $n = 1, 2, ...N - 1$. Next, U and V are updated at the infection-age boundary $j = 0$,

$$U_0^{n+1} = x^n S_b^n, \quad \text{for } n = 0, 1, ..., N-1, \tag{4.7.4}$$

$$V_0^{n+1} = y^n S_v^n, \quad \text{for } n = 0, 1, ..., N-1. \tag{4.7.5}$$

A first-order explicit Euler method updates U and V for infection ages $j = 1, 2, ..., M-1$ according to the rules outlined in Appendix D.3:

$$V_{j+1}^{n+1} = \begin{cases} \frac{V_j^n}{1+h\mu_v^n} & r_v^n \geq 0 \\ h r_v^n V_j^n \left(1 - \frac{N_v^n}{K_v^n}\right) + V_j^n & r_v^n < 0 \end{cases}, \tag{4.7.6}$$

$$U_{j+1}^{n+1} = \begin{cases} \frac{U_j^n}{1+h(\gamma_{b,j}+\delta_b+\mu_b)} & r_b^n \geq 0 \\ \frac{(1+hr_b^n)U_j^n}{1+h\gamma_{b,j}} & r_b^n < 0 \end{cases}, \tag{4.7.7}$$

where μ_v^n is the vector per capita mortality rate at time nh (temperature T^n), r_v^n is the vector per capita net growth rate at time nh, K_v^n is the vector carrying capacity at time nh, N_v^n is the total vector population size at time nh, r_b^n is the bird per capita net growth rate at time nh, and $\gamma_{b,j}$ is the bird per capita recovery rate at bird infection-age jh.

Updating S_b, S_v, R_b, S_h, I_h, and R_h also uses a first-order explicit Euler method for the next time step according to the rules outlined in Appendix D.3 (equations omitted). Finally, trapezoidal quadrature is again used to update x, y, and z at the next time point:

$$x^{n+1} = \frac{\alpha_b \eta}{N_b^n + N_h^n}\left(\frac{h}{2}\right)\left(\beta_{vb,0}^n V_0^{n+1} + 2\sum_{j=1}^{M-1} \beta_{vb,j}^n V_j^{n+1} + \beta_{vb,M}^n V_M^{n+1}\right), \tag{4.7.8}$$

$$y^{n+1} = \frac{\alpha_b \eta}{N_b^n + N_h^n}\left(\frac{h}{2}\right)\left(\beta_{bv,0} U_0^{n+1} + 2\sum_{j=1}^{M-1} \beta_{bv,j} U_j^{n+1} + \beta_{bv,M} U_M^{n+1}\right), \tag{4.7.9}$$

$$z^{n+1} = \frac{\alpha_h \eta}{N_b^n + N_h^n}\left(\frac{h}{2}\right)\left(\beta_{vh,0}^n V_0^{n+1} + 2\sum_{j=1}^{M-1} \beta_{vh,j}^n V_j^{n+1} + \beta_{vh,M}^n V_M^{n+1}\right). \tag{4.7.10}$$

4.5.3 Simulations from Parameter Fitting

Simulations of infected vectors, birds, and humans for the GTA for years 2010-2017 are displayed in Figure 4.3 along with the monthly reported human infections from Public Health Ontario [61]. Table 4.5 contains the fitted α_h values for each year along with the respective residual between yearly cumulative human infections and estimated human cases from the data [61]. Note that although α_h is fitted to the yearly estimated human cases (symptomatic and asymptomatic), visualizations display results for only symptomatic (true reported) cases.

There is generally a single peak of vector infections that occurs between August and September. The largest peak of vector infections occurs for year 2010 with 412 infected vectors. For the infected birds, there is typically a large peak between August and September each year, with smaller oscillations of infections both preceding and proceeding the main peak. This behavior is observed from the temperature-varying nature of the vector-to-bird transmission probability, which is not considered in the bird-to-vector transmission probability. The largest peak for infected birds occurs for September 2012 with 2,500 bird infections.

Year 2012 also shows the most human infections, with a peak of 64 symptomatic human infections in September and 164 symptomatic infections overall. Interestingly, this year has the second highest fitted α_h value. The largest α_h value occurred for year 2017, which was also the year with the smallest peak values for vector and bird infections. Moreover, there is only a 0.42 Pearson correlation between the fitted α_h values and yearly reported human cases from the GTA. This demonstrates that α_h alone is not enough to determine the magnitude of peak human cases, but that the vector, bird host, and temperature also significantly impact the outbreak.

Despite only fitting yearly α_h values to a single data point (yearly estimated human cases), the model can accurately predict the peak occurrence of observed human cases. This highlights the power of incorporating the temperature-varying processes in the PDE model. Exceptions include year 2017, where our model shows two smaller peaks for human infections in mid August and early October, compared to the monthly data showing a single, larger peak in September. Despite this, the yearly total number of human cases still aligns with the fitted model simulations. The magnitude of the human infection peaks for 2011, 2012, and 2016 do not quite match the values from the human case data, but this is likely due to the differing temporal resolution between the model (daily infections) versus the data (monthly infections).

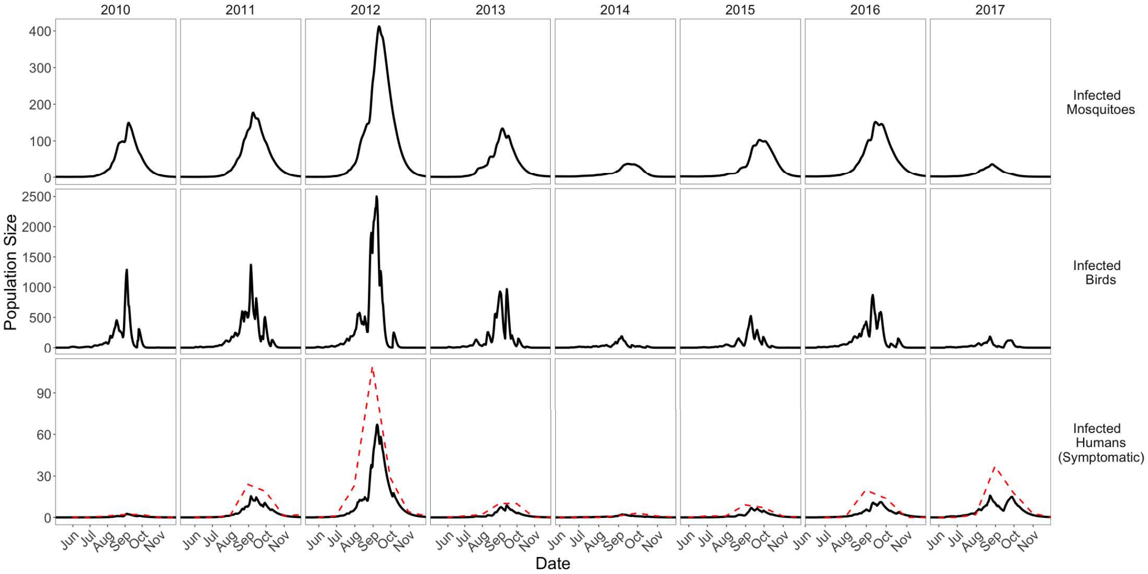


Figure 4.3: Simulations of Infected Mosquito Vectors (Top Row), Birds (Middle Row), and Symptomatic Humans (Bottom Row) for the Greater Toronto Area for Years 2010-2017. The Red Dashed Line in the Bottom Row Shows the Reported Monthly Human Cases from Public Health Ontario [61].

Next, we apply projections of future weather data under two climate change scenarios to investigate possible future dynamics of WNV transmission in the GTA.

4.6 Climate Change Scenario Analysis

4.6.1 Data Sources for Climate Change Projections

Data from NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) provided information to estimated projected temperature and precipitation time series for representative concentration pathways (RCP) 4.5 and 8.5 [55]. Recall that RCP4.5 refers to the “middle of the road” scenario, where socioeconomic factors are expected to follow historical trends, and that RCP8.5 is known as the “worst case scenario” with increased fossil fuel driven development [46]. The NEX-GDDP data was derived from 21 general circulation models (GCMs) developed from the 5th assessment report of the Intergovernmental Panel on Climate Change. The NEX-

Table 4.5: Fitted Values of the Vector Per Capita Biting Rate on Humans α_h for the Greater Toronto Area for Years 2010-2017. The Parameter α_h was Fitted to the Yearly Estimated Human WNV Cases Based on Data from Public Health Ontario [61]. Residual Values Indicate the Difference Between the Cumulative New Human Infections from the Model and the Yearly Estimated Cases.

Year	**Fitted** α_h	**Residual**
2010	0.0188	2.84×10^{-14}
2011	0.1439	-2.84×10^{-14}
2012	0.2741	-3.41×10^{-13}
2013	0.0863	0.00
2014	0.1047	1.06×10^{-14}
2015	0.1168	-9.95×10^{-14}
2016	0.0933	0.00
2017	0.6786	3.18×10^{-12}

GDDP data catalog provides daily maximum temperature, minimum temperatures, and precipitation for years 1950-2100 under RCP4.5 and RCP8.5.

We use the Delta Method on a monthly timescale to estimate projected time series of daily mean temperature and precipitation for an average year between 2080-2100 for the GTA [17]. Specifically, we calculated anomalies (deltas) between the monthly averages of 2080-2100 and 2006-2020 NEX-GDDP data for both RCP4.5 and RCP8.5 scenarios. We add these monthly deltas to the monthly average of 2006-2020 data from ERA5 (baseline data) [32]. Note that the ERA5 data was used to construct the mosquito PBM time series of the historical data and for the daily temperature data for simulations. The maximum and minimum daily temperatures of the NEX-GDDP data were averaged to estimate the daily mean temperature, and all precipitation time series applies a 14-day moving average before calculating delta values.

Figure 4.4 shows the temperature and precipitation for and average year of current conditions (2006-2020) and future projections (2080-2100) under RCP4.5 and RCP8.5 for the GTA. Under RCP4.5, the average temperature of the GTA is expected to increase by 2.25°C, and average precipitation is expected to increase by 5.53%, with the greatest increase expected for winter and spring months (December - April). Average temperature is expected to increase by 5.06°C on average under RCP8.5, with the greatest increase during the winter. Precipitation under RCP8.5 is expected to increase by 11.38% on average, with relative increases over 20% for early spring.

The purpose of using projected GCM data specific for the GTA is because climate change is expected to have a nonuniform effect on different micro-climates [46]. Other studies that analyze climate change scenario impact on vectors and vector-borne disease models assume a uniform increase of temperature [82, 54], which may over or underestimate the true expected conditions under climate change. The following section explains how the projected time series data is used to estimate plausible mosquito population sizes under the RCP4.5 and RCP8.5 scenarios.

4.6.2 Projections for Total Population Sizes

Projected daily mean temperature and precipitation time series found in Section 4.6.1 were fed into the mosquito PBM model [74] to construct daily time series of mosquitoes for the GTA under RCP4.5 and RCP8.5. Since there is limited information about how bird populations will be altered under climate change, we estimate a relative abundance time series for WNV-competent birds using the average relative abundance of the 2010-2017 time series generated from the eBird data applied with the GPR [19, 62].

Projections of WNV-competent bird relative abundance, along with mosquito population sizes for the GTA under RCP4.5 and RCP8.5 are displayed in Figure 4.5 and assume conditions for an average year between 2080-2100. Parameter values from

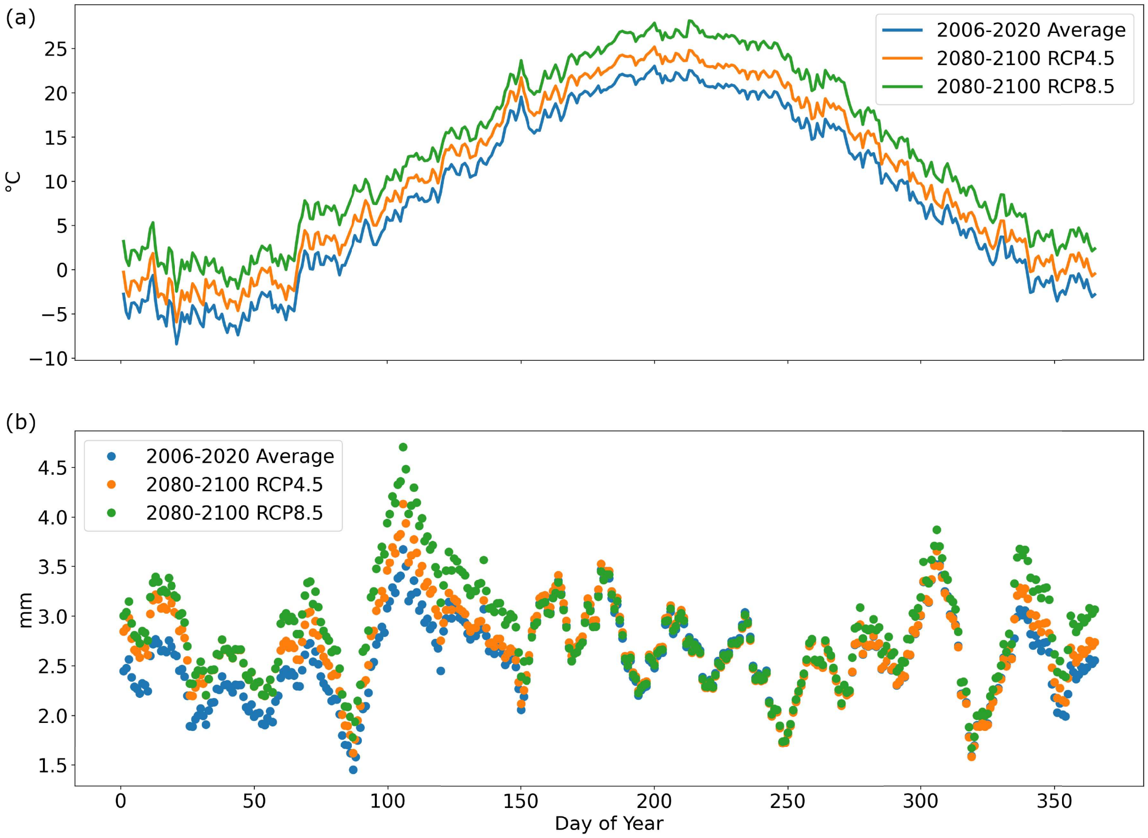


Figure 4.4: Projected Time Series for (a) Temperature and (b) Precipitation for an Average Year Between 2080-2100 Under Climate Scenarios RCP4.5 (Yellow) and RCP8.5 (Green) Compared to an Average Year Between 2006-2020 (Blue). Precipitation Time Series Represents a 14-Day Moving Average.

fitting the population time series are found in Table 4.6. Parameter values were fitted using the methods described in Section 4.4.

Notice that the mosquito PBM time series for RCP4.5 and RCP8.5 conditions are one and two orders of magnitude greater than the largest mosquito peak in the historical PBM time series (see Figure 2.3). These projections are understandable and plausible considering the overall 2.5 and 5°C temperature increases, and increased precipitation expected in the spring [54, 73].

We now use the parameters fitted from the projected population time series into the PDE model to simulate possible future dynamics of infected populations under climate change scenarios RCP4.5 and RCP8.5.

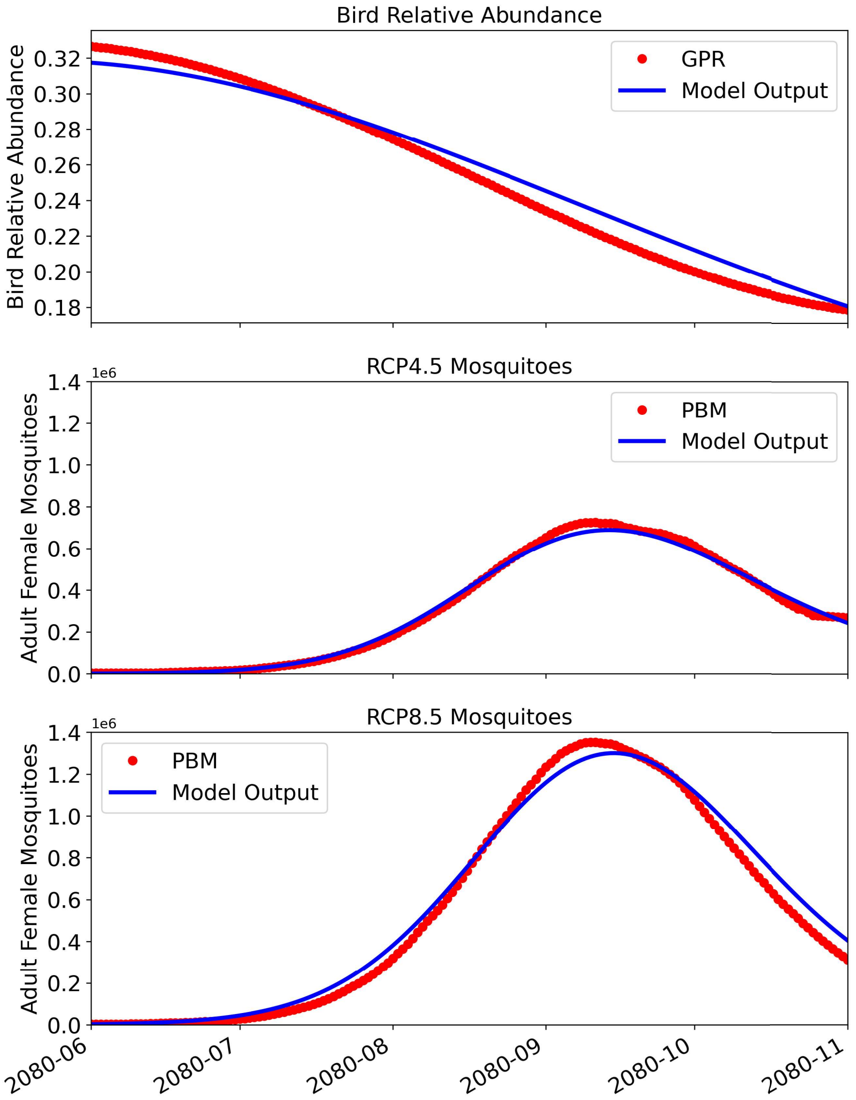


Figure 4.5: Projections of Total Bird and Vector Population Sizes. Red Dots Represent the (a) Average Time Series of WNV-Competent Bird Relative Abundance Based on eBird Data for Years 2010-2017 (Red Dots) [19], (b) Mosquito Process-Based Model Output [74] for Average Conditions Between 2080-2100 Under Climate Change Scenario RCP4.5 and (c) RCP8.5 Based on Projected Temperature and Precipitation Data [55]. Blue Curves Show the Simulations of Equation (4.3) from Fitting Parameters of Equation (4.4) to Their Respective Time Series.

Table 4.6: Fitted Parameter Values for Projected Vector and Bird Population Sizes Under Climate Change Scenarios RCP4.5 and RCP8.5. Note That "Start Day" Refers to the First Day of the Mosquito Season.

Parameter	**RCP4.5**	**RCP8.5**
Start Day	May 24	June 1
r	0.029	0.018
r_s	-0.080	-0.073
K	1.00×10^{7}	1.00×10^{8}
K_s	-9.5×10^{6}	-9.5×10^{7}
ϕ	2.10×10^{-4}	2.10×10^{-4}
ϕ_s	-5.38×10^{-3}	-5.38×10^{-3}
ω	0.01	0.01
Δt	165	173
$N_v(0)$	365.98	3215.85
$N_b(0)$	901.06	897.97

4.6.3 Simulations of Climate Change Scenarios

Simulations of infected vector, bird, and human population sizes for an average year between 2080-2100 for the GTA under climate change scenarios RCP4.5 and RCP8.5 are displayed in Figure 4.6. Simulations use parameters and initial conditions from Tables 4.1, 4.3, and 4.6. Due to the uncertainty of the vector per capita biting rate on humans, simulations use the mean and 95% confidence interval (CI) of the α_h values obtained from fitting historical human case data. This provides a range of possible human infection dynamics under the two climate change scenarios. Numerical simulations were conducted using the method outlined in Section 4.5.2. Simulations required that we reduce the vector-to-host tuning parameter from $\eta = 25$ to $\eta = 12.5$ for the RCP4.5 simulation and to $\eta = 6.25$ for the RCP8.5 simulation. These adjustments were necessary to avoid numerical instability, and we were unable to achieve numerically stable simulations under the same value of η across all three scenario conditions.

Peak infections for all populations are expected to occur around the same time as under current conditions (late September - early October). Compared to the greatest peak mosquito population from 2010-2017 simulations, peak mosquito population is estimated to increase by one order of magnitude under RCP4.5 conditions and by two orders of magnitude under RCP8.5 conditions. These estimates reflect the large impact of temperature increases from climate change on mosquito populations. However, when comparing the ratio of peak infected to total vectors across historical and scenario conditions, this value increases 14.5 times under RCP4.5 and 25 times under RCP8.5. This demonstrates the possible large effect of increased temperature on WNV transmission probability.

A difference of 162 human cases was observed between the greatest and least years of reported cases between 2010-2017 – years 2010 and 2014 had the least number of cases reported with 6 cases, and 2012 had the greatest number of reported cases with 168 cases. The range between overall symptomatic human infections is expected to increase 22-fold under RCP4.5 and 75-fold under RCP8.5. These range estimates were determined from taking the difference of overall symptomatic human infections produced by simulations using the upper and lower 95% CI of α_h values from historical data fitting. This shows how increased temperatures can lead to larger variation in overall human infections from year to year.

Under the most conservative estimation of the vector per capita biting rate on humans (lower 95% CI value), annual symptomatic human cases are expected to be similar to the greatest number observed from 2010-2017 (139 vs. 168 cases). However, RCP8.5 conditions could lead to a 17-fold increase of overall human cases

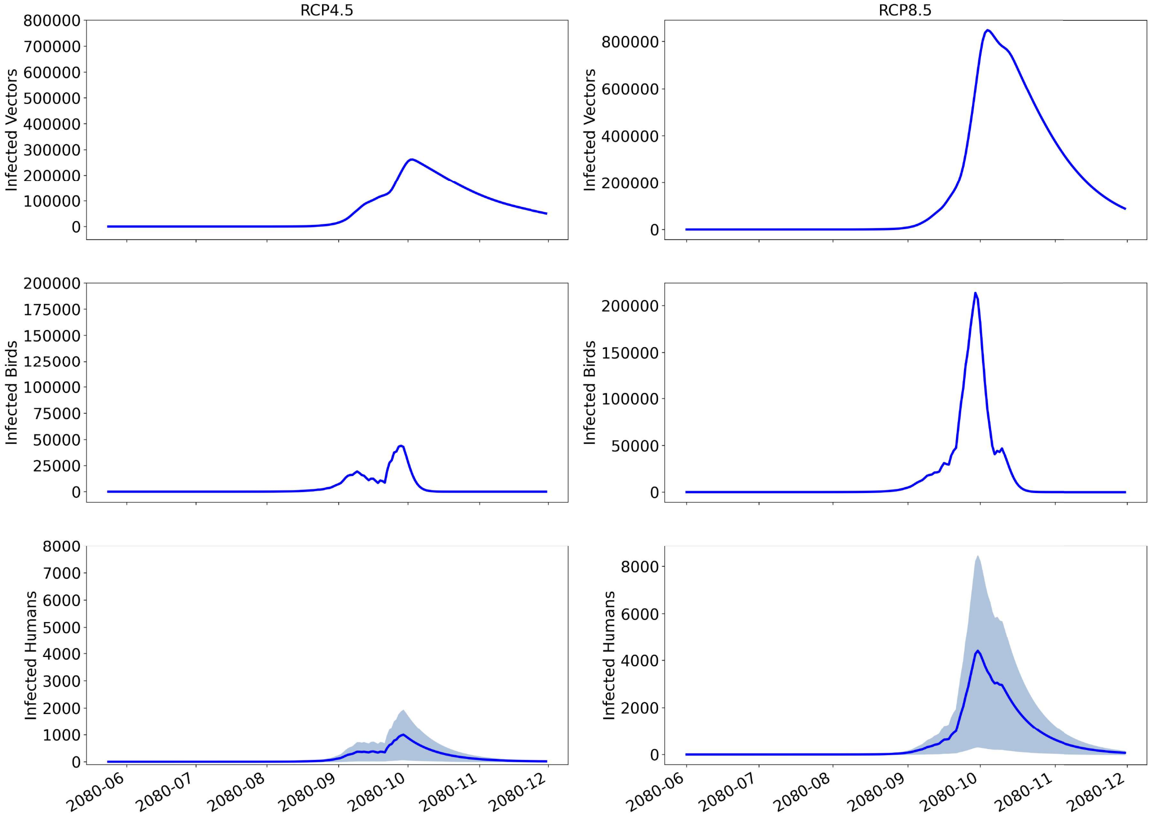


Figure 4.6: Simulations of Infected Vector (Top Row), Bird (Middle Row), and Symptomatic Human (Bottom Row) Population Sizes for an Average Year Between 2080-2100 in the Greater Toronto Area Under Climate Change Scenarios RCP4.5 (Left Side) and RCP8.5 (Right Side). For the Infected Symptomatic Human subfigures, the Blue Lines Represent Simulations Using the Mean Vector Per Capita Biting Rate on Humans $\alpha_h = 0.1896$ from Fitting to Historical Human Case Data, and Shaded Areas Reflect Simulations for the 95% Confidence Interval of Fitted α_h Values ([0.0136, 0.3654]).

from the most conservative α_h estimate. These estimates demonstrate the importance of limiting the warming of our planet. These results will, hopefully, encourage policy makers to make decisions that will protect their constituents from future WNV and other MBD risk.

4.7 Discussion

This chapter brings together many pieces and parts to investigate the possible impact of climate change on WNV transmission. We extend the simplified WNV

PDE model of Chapter 3 to include temperature-dependent processes for the vector per capita mortality rate and vector-to-host transmission probability, and incorporate inter-annual and intra-seasonal variation of vector and bird host demography through the data fusion technique described in Chapter 2. After parameterizing the model using human WNV case data from the Greater Toronto Area (GTA), we then use projected weather data from general circulation models to predict possible dynamics of future WNV transmission. We demonstrate the importance of including temperature dependence on demographic and epidemiological processes, as the model was able to accurately capture the peak timing of human infections, despite fitting the model to a single yearly data point.

Since there is limited information on how bird and human hosts population sizes will change for the GTA under climate change, this study only focuses on the contribution of changing mosquito populations affecting WNV transmission under climate change. We decided to fit the vector per capita biting rate on humans to the yearly estimated human WNV cases for two reasons. First, the per capita biting rate was shown to be one of the most sensitive parameters of the basic reproduction number of the simplified model. Second, there is likely temperature dependence on the biting rate, as studies of other mosquito species showed temperature-dependent biting rates [53, 69]. However, we do not know the functional form of this relationship for *Culex* mosquitoes, as the effect of temperature on *Culex* biting has yet to be studied. Future work may aim to extend the per capita biting rate to be temperature dependent instead of fitting separately per capita biting rates for each year.

There is still a lot of uncertainty about the exact translation from mosquito trap data to the true mosquito population. There is even more uncertainty about how the vector population size scales to the total host population size. Here we incorporate the parameter η to scale the vector-to-host population size ratio to a suitable level for generating a WNV outbreak with our assumed initial conditions. The selected η value affects the per capita vector-to-human biting rate α_h that is fitted from the historical human case data. It is challenging to select the best η for parameter fitting because we lack bird infection data for comparing simulations. We found that we cannot use the same η value in the climate scenario analysis as the one used in the historical data due to numerical stability issues. However, reducing the η value in the climate change scenarios may account for the likely increased human host population size in the next 60-80 years.

Analyses for climate change scenarios RCP4.5 and RCP8.5 estimate that peak and overall human WNV cases will substantially increase as temperature increases from climate change. Further, additional variation and uncertainty of peak case magnitude is expected under the worst case scenario of RCP8.5. These results show how crucial it is to limit warming of our planet in order to mitigate future WNV risk as much as possible. Although the results here are specific for WNV and for the GTA, this framework ultimately aids in our preparation for future vector-borne disease outbreaks.

It is worth revisiting the formulation of the force of infection equations, as different assumptions underlying the formulation may lead to differing conclusions. For this study, the force of infection considers the total vector biting rate α on the total host population $N_b + N_h$. We assume the biting rate on birds α_b and the biting rate on humans α_h to be proportions of the total biting rate on all hosts. We made this assumption to aid in the translation from the simplified model to the full model. Other studies formulate the force of infection by separating the host populations [6]. Under this modified assumption, the denominator of λ_{vb} and λ_{bv} would be N_b, and the denominator of λ_{vh} would be N_h. It will be interesting to see how simulation results under this alternative formulation will compare with our current model output, as it is possible that this alternative formulation may allow us to omit the vector-to-host scaling parameter η. It may even be worthwhile to incorporate a time-dependent per capita biting rate between birds and humans, which has been previously documented [75] and modeled for WNV [23]. Including these additional biological considerations to the model may further inform our climate change scenario analysis.

Chapter 5

CONCLUSION

As climate change continues to threaten our health and well being, and considering that mosquito-borne diseases take an enormous toll on human health and the world economy, it is crucial to understand the complex impacts our climate has on mosquito-borne disease (MBD) transmission. Understanding the role of climate change and demography on future West Nile Virus (WNV) risk involves a comprehensive model that incorporates numerous data streams and submodel outputs. The Climate Integrated Model for Mosquito-Borne Infectious Diseases (CIMMID) project at the Los Alamos National Laboratory aims to develop a continental scale model for analyzing MBD risk under climate change scenarios. This dissertation aims to answer two questions to enhance our knowledge about modeling WNV transmission. The first question asks about the impact of vector and host demography on WNV dynamics. The second question addresses the possible impacts of expected and likely climate change scenarios on demographic and epidemiological parameters for WNV transmission.

The second chapter prepares us to address the question of vector demography on WNV transmission. In that chapter, we develop a non-autonomous logistic model to capture the pattern of discrete mosquito time series data. Incorporating non-autonomous parameters allowed us to analyze the inter-annual and intra-seasonal variation of mosquito populations within a geographic region. We discovered that the parameters of the non-autonomous model have different sensitivity to the seasonal peak magnitude and timing for different types of mosquito populations. We also learned that we cannot simply connect the fitted parameters to weather data due to the complex, nonlinear interactions between mosquito processes and their environment. However, this data fusion framework is valuable for being a powerful tool in aiding the CIMMID initiative. The data fusion method using the non-autonomous logistic model is seamlessly integrated into the CIMMID framework. CIMMID uses earth systems models to partition land area into different units, called hydropop units, based on hydrology, vegetation, and population characteristics. There are tens of thousands of hydropop units for the North American continent, and each one delivers input into the mosquito process-based model (PBM) to generate time series of daily mosquito estimates. The data fusion framework of the non-autonomous logistic model facilitates the connection between the discrete time series data and continuous demographic parameters for the epidemiological models.

The third chapter addresses the question of vector and host demography by analyzing infection-age dynamics. We developed a three-population partial differential equations model that has infection-age heterogeneity for both mosquito vector and bird host populations. To our knowledge, this is the first PDE model for WNV that incorporates infection-age dependence on both vector and host populations. We establish the existence of a unique continuous solution under certain model assumptions. Results from sensitivity analysis indicate that control strategies for WNV mitigation should focus on increasing the vector mortality rate and decreasing the vector biting rate. We plan to submit the work in this chapter for publication in the Journal of Theoretical Biology, Mathematical Biosciences, or Mathematical Biosciences and Engineering.

The fourth chapter is the culmination of the dissertation and addresses our questions of interest. In this chapter, we connect the non-autonomous logistic model for mosquito populations from the second chapter with the infection-age structured WNV PDE model from the third chapter. We further extended the model to include temperature dependence on the vector per capita mortality rate (demographic parameter) and transmission probability from vectors to hosts (epidemiological parameter). Incorporating this temperature dependence allowed us to analyze possible future WNV dynamics under climate change scenarios using projected weather data from general circulation models. We fitted the vector per capita biting rate on humans to historical human case data, which was used to inform the range of likely future dynamics under climate change scenarios RCP4.5 and RCP8.5. The main takeaway from the scenario analysis is that temperature increases due to climate change will not only increase the overall number of human WNV infections for the GTA, but we can also expect a wider variation of outbreak magnitude from year-to-year. We plan to revisit the formulation of the force of infection before submitting the work for publication.

Future work will involve exploring the climate change impact on WNV transmission of other geographic locations with different regional characteristics. For example, Maricopa County, Arizona, which has a hot, dry climate compared to the GTA where it is temperate and wet. Despite these environmental differences, both areas have suffered WNV outbreaks in recent years. Since climate change is expected to have a heterogeneous impact across different types of micro-climates, it will be interesting to see if climate change will cause Maricopa County to be too hot for vector-borne disease transmission, or if the mosquito season will shift to earlier spring or later fall. By understanding the role that climate and demography have on MBD such as WNV, we can better prepare to mitigate future outbreaks that will be driven by the increased transmission risk that will result from climate change.

REFERENCES

[1] Abdelrazec, A., S. Lenhart and H. Zhu, "Transmission dynamics of West Nile virus in mosquitoes and corvids and non-corvids", Journal of Mathematical Biology **68**, 6, 1553–1582 (2014).

[2] Albrecht, L. and K. A. Kaufeld, "Investigating the impact of environmental factors on West Nile virus human case prediction in Ontario, Canada", Frontiers in Public Health **11**, URL `https://www.frontiersin.org/articles/10.3389/fpubh.2023.1100543` (2023).

[3] Anderson, J. F., A. J. Main, G. Cheng, F. J. Ferrandino and E. Fikrig, "Horizontal and Vertical Transmission of West Nile Virus Genotype NY99 by Culex salinarius and Genotypes NY99 and WN02 by Culex tarsalis", The American Journal of Tropical Medicine and Hygiene **86**, 1, 134–139, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3247122/` (2012).

[4] Banks, R. B., *Growth and Diffusion Phenomena: Mathematical Frameworks and Applications* (Springer, Berlin Heidelberg, 1993), 1994th edition edn.

[5] Bartlow, A. W., C. Manore, C. Xu, K. A. Kaufeld, S. Del Valle, A. Ziemann, G. Fairchild and J. M. Fair, "Forecasting Zoonotic Infectious Disease Response to Climate Change: Mosquito Vectors and a Changing Environment", Veterinary Sciences **6**, 2, 40, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6632117/` (2019).

[6] Bergsman, L. D., J. M. Hyman and C. A. Manore, "A mathematical model for the spread of west nile virus in migratory and resident birds", Mathematical biosciences and engineering: MBE **13**, 2, 401–424 (2016).

[7] Blancher, P. J., K. V. Rosenberg, A. Panjabi, B. Altman, J. Bart, C. Beardmore, G. Butcher, D. Demarest, R. Dettmers, E. Dunn, W. Easton, W. Hunter, E. Iñigo-Elias, D. Pashley, C. Ralph, T. Rich, C. Rustay, J. Ruth and T. Will, "Guide to the partners in flight population estimates database.", Partners in Flight Technical Series No 5 (2007).

[8] Blayneh, K. W., A. B. Gumel, S. Lenhart and T. Clayton, "Backward Bifurcation and Optimal Control in Transmission Dynamics of West Nile Virus", Bulletin of Mathematical Biology **72**, 4, 1006–1028, URL `https://doi.org/10.1007/s11538-009-9480-0` (2010).

[9] Bongard, J. and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems", Proceedings of the National Academy of Sciences **104**, 24, 9943–9948, URL `https://www.pnas.org/doi/10.1073/pnas.0609476104`, publisher: Proceedings of the National Academy of Sciences (2007).

[10] Bowman, C., A. B. Gumel, P. van den Driessche, J. Wu and H. Zhu, "A mathematical model for assessing control strategies against West Nile virus", Bulletin of Mathematical Biology **67**, 5, 1107–1133 (2005).

[11] Breeding Bird Survey, "BBS - USGS Patuxent Wildlife Research Center", URL `https://www.pwrc.usgs.gov/bbs/` (2023).

[12] Brunton, S., J. Proctor and J. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems | PNAS", PNAS URL `https://www.pnas.org/doi/10.1073/pnas.1517384113` (2016).

[13] CDC, "Adulticides | CDC", URL `https://www.cdc.gov/mosquitoes/mosquito-control/community/adulticides.html` (2023).

[14] Centers for Disease Control and Prevention, "Fighting the World's Deadliest Animal", URL `https://www.cdc.gov/globalhealth/stories/2019/world-deadliest-animal.html` (2023).

[15] Centers for Disease Control and Prevention, "West Nile Virus", URL `https://www.cdc.gov/westnile/index.html` (2023).

[16] Ciota, A. T., A. C. Matacchiero, A. M. Kilpatrick and L. D. Kramer, "The Effect of Temperature on Life History Traits of Culex Mosquitoes", Journal of medical entomology **51**, 1, 55–62, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3955846/` (2014).

[17] Climage Change, Agriculture, and Food Security, "Statistical Downscaling (Delta Method)", URL `https://www.ccafs-climate.org/statistical_downscaling_delta/` (2014).

[18] Coleman, B. D., "Nonautonomous logistic equations as models of the adjustment of populations to environmental change", Mathematical Biosciences **45**, 3, 159–173, URL `https://www.sciencedirect.com/science/article/pii/0025556479900579` (1979).

[19] Cornell Lab of Ornithology, "eBird: An online database of bird distribution and abundance", URL `https://ebird.org/home` (2023).

[20] Crutchfield, J. P. and M. B. S., "Equations of Motion from a Data Series", Complex Systems URL `https://www.complex-systems.com/abstracts/v01_i03_a03/` (1987).

[21] Cunniffe, N., F. Hamelin, A. Iggidr, A. Rapaport and G. Sallet, "Observability, Identifiability and Epidemiology – A survey", URL `http://arxiv.org/abs/2011.12202`, arXiv:2011.12202 [math] (2023).

[22] Daniels, B. C. and I. Nemenman, "Automated adaptive inference of phenomenological dynamical models", Nature Communications **6**, 1, 8133, URL `https://www.nature.com/articles/ncomms9133`, number: 1 Publisher: Nature Publishing Group (2015).

[23] DeFelice, N. B., E. Little, S. R. Campbell and J. Shaman, "Ensemble forecast of human West Nile virus cases and mosquito infection rates", Nature Communications **8**, 1, 14592, URL `https://www.nature.com/articles/ncomms14592`, number: 1 Publisher: Nature Publishing Group (2017).

[24] Di Gironimo, L., T. Bowering and D. Kellershohn, "Impact of july 9, 2013 storm on the city sewer and stormwater systems", URL `https://www.toronto.ca/legdocs/mmis/2013/ex/bgrd/backgroundfile-61502.pdf` (2013).

[25] Gorris, M., J. Randerson, S. Coffield, K. Treseder, C. Zender and C. Xu, "Assessing the Influence of Climate on the Spatial Pattern of West Nile Virus Incidence in the United States | Environmental Health Perspectives | Vol. 131, No. 4", Environmental Health Perspectives URL `https://ehp.niehs.nih.gov/doi/full/10.1289/EHP10986` (2023).

[26] Government of Canada, S. C., "Census Subdivision Boundary File", URL `https://www12.statcan.gc.ca/census-recensement/2021/geo/sip-pis/boundary-limites/index-eng.cfm`, last Modified: 2023-07-12 (2021).

[27] Government of Canada, S. C., "Census Profile, 2021 Census of Population", URL `https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E`, last Modified: 2023-03-29 (2022).

[28] Grozdanovski, T., J. J. Shepherd and A. Stacey, "Multi-scaling analysis of a logistic model with slowly varying coefficients", Applied Mathematics Letters **22**, 7, 1091–1095, URL `https://www.sciencedirect.com/science/article/pii/S0893965909000366` (2009).

[29] Gulbudak, H., "An Immuno-Epidemiological Vector–Host Model with Within-Vector Viral Kinetics | Journal of Biological Systems", URL `https://www.worldscientific.com/doi/abs/10.1142/S0218339020400021` (2020).

[30] Gulbudak, H., Z. Qu, F. Milner and N. Tuncer, "Sensitivity Analysis in an Immuno-Epidemiological Vector-Host Model", Bulletin of Mathematical Biology **84**, 2, 27, URL `https://doi.org/10.1007/s11538-021-00979-0` (2022).

[31] Herman, J. and W. Usher, "SALib: An open-source Python library for Sensitivity Analysis", Journal of Open Source Software **2**, 9, 97, URL `https://joss.theoj.org/papers/10.21105/joss.00097` (2017).

[32] Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee and J.-N. Thépaut, "The ERA5 global reanalysis", URL `https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.3803` (2018).

[33] Iannelli, M. and F. Milner, *The Basic Approach to Age-Structured Population Dynamics: Models, Methods and Numerics*, Lecture Notes on Mathematical Modelling in the Life Sciences (Springer Netherlands, Dordrecht, 2017), URL `http://link.springer.com/10.1007/978-94-024-1146-1`.

[34] Iwanaga, T., W. Usher and J. Herman, "Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses", Socio-Environmental Systems Modelling **4**, 18155–18155, URL `https://sesmo.org/article/view/18155` (2022).

[35] Jiang, D. and N. Shi, “A note on nonautonomous logistic equation with random perturbation”, Journal of Mathematical Analysis and Applications **303**, 164–172, URL `https://ui.adsabs.harvard.edu/abs/2005JMAA..303..164J`, aDS Bibcode: 2005JMAA..303..164J (2005).

[36] Kermack, W. and A. McKendrick, *A contribution to the mathematical theory of epidemics | Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* (Proceedings of the Royal Society of London, 1927), URL `https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118`.

[37] Kilpatrick, A. M., L. D. Kramer, M. J. Jones, P. P. Marra and P. Daszak, “West Nile Virus Epidemics in North America Are Driven by Shifts in Mosquito Feeding Behavior”, PLOS Biology **4**, 4, e82, URL `https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040082`, publisher: Public Library of Science (2006).

[38] Koenraadt, C. J. M., T. W. R. Möhlmann, N. O. Verhulst, J. Spitzen and C. B. F. Vogels, “Effect of overwintering on survival and vector competence of the West Nile virus vector Culex pipiens”, Parasites & Vectors **12**, 1, 147, URL `https://doi.org/10.1186/s13071-019-3400-4` (2019).

[39] Komar, N., “West Nile virus: epidemiology and ecology in North America”, Advances in Virus Research **61**, 185–234 (2003).

[40] Kramer, L. D., “Complexity of virus - vector interactions”, Current opinion in virology **21**, 81–86, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5138088/` (2016).

[41] Kuang, Y., J. Nagy and S. Eikenberry, *Introduction to Mathematical Oncology* (Chapman and Hall/CRC, London, 2016).

[42] Laboratory, L. A. N., “Laboratory Directed Research & Development | Science Programs”, URL `https://science-innovation.lanl.gov/science-programs/ldrd/` (2023).

[43] Lambrechts, L., K. P. Paaijmans, T. Fansiri, L. B. Carrington, L. D. Kramer, M. B. Thomas and T. W. Scott, “Impact of daily temperature fluctuations on dengue virus transmission by Aedes aegypti”, Proceedings of the National Academy of Sciences **108**, 18, 7460–7465, URL `https://www.pnas.org/doi/10.1073/pnas.1101377108`, publisher: Proceedings of the National Academy of Sciences (2011).

[44] Laperriere, V., K. Brugger and F. Rubel, “Simulation of the seasonal cycles of bird, equine and human West Nile virus cases”, Preventive Veterinary Medicine **98**, 2-3, 99–110 (2011).

[45] Lazebnik, T., L. Shami and S. Bunimovich-Mendrazitsky, “Spatio-Temporal influence of Non-Pharmaceutical interventions policies on pandemic dynamics and

the economy: the case of COVID-19", Economic Research-Ekonomska Istrazivanja **35**, 1, 1833–1861, URL `https://cris.haifa.ac.il/en/publications/spatio-temporal-influence-of-non-pharmaceutical-interventions-pol`, publisher: Taylor and Francis Ltd. (2022).

[46] Lee, J.-Y., J. Marotzke, G. Bala, L. Cao, S. Corti, J. Dunne, F. Engelbrecht, E. Fischer, J. Fyfe, C. Jones, A. Maycock, J. Mutemi, O. Ndiaye, S. Panickal and T. Zhou, "Future Global Climate: Scenario-based Projections and Near-term Information", in "Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change", edited by Intergovernmental Panel on Climate Change (IPCC), pp. 553–672 (Cambridge University Press, Cambridge, 2021), URL `https://www.cambridge.org/core/books/climate-change-2021-the-physical-science-basis/future-global-climate-scenariobased-projections-and-nearterm-information/309359EDDCFABB031C078AE20CEE04FD`.

[47] Macdonald, G., *The Epidemiology and Control of Malaria* (Oxford University Press, 1957), google-Books-ID: xMu2AAAAIAAJ.

[48] Maidana, N. A. and H. M. Yang, "Spatial spreading of West Nile Virus described by traveling waves", Journal of Theoretical Biology **258**, 3, 403–417, URL `https://linkinghub.elsevier.com/retrieve/pii/S0022519308006322` (2009).

[49] Mancuso, M., K. M. Martinez, C. A. Manore, F. A. Milner, M. Barnard and H. Godinez, "Fusing time-varying mosquito data and continuous mosquito population dynamics models", Frontiers in Applied Mathematics and Statistics **9**, URL `https://www.frontiersin.org/articles/10.3389/fams.2023.1207643` (2023).

[50] Manore, C. A., K. S. Hickmann, S. Xu, H. J. Wearing and J. M. Hyman, "Comparing dengue and chikungunya emergence and endemic transmission in A. aegypti and A. albopictus", Journal of Theoretical Biology **356**, 174–191 (2014).

[51] Martcheva, M., *An Introduction to Mathematical Epidemiology* (Springer, New York Heidelberg Dordrecht London, 2015), 1st ed. 2015 edition edn.

[52] Marx, V., "The big challenges of big data | Nature", Nature URL `https://www.nature.com/articles/498255a` (2013).

[53] Mordecai, E. A., K. P. Paaijmans, L. R. Johnson, C. Balzer, T. Ben-Horin, E. de Moor, A. McNally, S. Pawar, S. J. Ryan, T. C. Smith and K. D. Lafferty, "Optimal temperature for malaria transmission is dramatically lower than previously predicted", Ecology Letters **16**, 1, 22–30 (2013).

[54] Moser, S., J. Spencer, M. Barnard, R. Frantz, K. Rodarte and I. Crooker, "Scoping review of Culex mosquito life history trait heterogeneity in response to temperature", Parasites & Vectors URL `https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-023-05792-3` (2023).

[55] NASA Center for Climate Simulation, "NEX-GDDP", URL `https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp` (2023).

[56] Nesslage, G. and M. Wilberg, "Performance of Surplus Production Models with Time-Varying Parameters for Assessing Multispecies Assemblages", North American Journal of Fisheries Management **32** (2012).

[57] Nipa, K. F., S. R.-J. Jang and L. J. S. Allen, "The effect of demographic and environmental variability on disease outbreak for a dengue model with a seasonally varying vector population", Mathematical Biosciences **331**, 108516 (2021).

[58] Nisbet, R. M. and W. S. Gurney, "Population dynamics in a periodically varying environment", Journal of Theoretical Biology **56**, 2, 459–475 (1976).

[59] Pan American Health Organization, "Zika", URL `https://www.paho.org/en/topics/zika` (2023).

[60] Perez-Ramirez, E., F. Llorente and M. A. Jimenez-Clavero, "Experimental Infections of Wild Birds with West Nile Virus", Viruses **6**, 2, 752–781, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3939481/` (2014).

[61] Public Health Ontario, "West Nile Virus Surveillance", URL `https://www.publichealthontario.ca/en/Data-and-Analysis/Infectious-Disease/West-Nile-Virus` (2023).

[62] Rasmussen, C. E. and C. K. I. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, 2005), URL `https://direct.mit.edu/books/book/2320/Gaussian-Processes-for-Machine-Learning`.

[63] Reisen, W., Y. Fang and V. Martinez, "Effects of Temperature on the Transmission of West Nile Virus by Culex tarsalis (Diptera: Culicidae)", Journal of Medical Entomology URL `https://academic.oup.com/jme/article/43/2/309/1061884` (2006).

[64] Richard, Q., M. Choisy, T. Lefèvre and R. Djidjou-Demasse, "Human-vector malaria transmission model structured by age, time since infection and waning immunity", Nonlinear Analysis: Real World Applications **63**, 103393, URL `https://linkinghub.elsevier.com/retrieve/pii/S146812182100105X` (2022).

[65] Richards, S. L., S. L. Anderson, C. C. Lord, C. T. Smartt and W. J. Tabachnick, "Relationships Between Infection, Dissemination, and Transmission of West Nile Virus RNA in Culex pipiens quinquefasciatus (Diptera: Culicidae)", Journal of Medical Entomology **49**, 1, 132–142, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3320798/` (2012).

[66] Richards, S. L., B. D. Byrd, M. H. Reiskind and A. V. White, "Assessing Insecticide Resistance in Adult Mosquitoes: Perspectives on Current Methods", Environmental Health Insights **14**, 1178630220952790, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7477762/` (2020).

[67] Rock, K. S., D. A. Wood and M. J. Keeling, "Age- and bite-structured models for vector-borne diseases", Epidemics **12**, 20–29, URL `https://www.sciencedirect.com/science/article/pii/S1755436515000286` (2015).

[68] Ross, R., *The prevention of malaria* (New York, E.P. Dutton & company, 1910), URL `http://archive.org/details/pr00eventionofmalarossrich`.

[69] Rubel, F., K. Brugger, M. Hantel, S. Chvala-Mannsberger, T. Bakonyi, H. Weissenböck and N. Nowotny, "Explaining Usutu virus dynamics in Austria: model development and calibration", Preventive Veterinary Medicine **85**, 3-4, 166–186 (2008).

[70] Rund, S. S. C., I. K. Moise, J. C. Beier and M. E. Martinez, "Rescuing Troves of Hidden Ecological Data to Tackle Emerging Mosquito-Borne Diseases", Journal of the American Mosquito Control Association **35**, 1, 75–83, URL `https://doi.org/10.2987/18-6781.1` (2019).

[71] Saltelli, A., "A short comment on statistical versus mathematical modelling", Nature Communications URL `https://www.nature.com/articles/s41467-019-11865-8` (2019).

[72] SciPy, "scipy.optimize.least_squares — SciPy v1.11.3 Manual", URL `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least_squares.html` (2022).

[73] Shaman, J. and J. F. Day, "Reproductive Phase Locking of Mosquito Populations in Response to Rainfall Frequency", PLOS ONE **2**, 3, e331, URL `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000331`, publisher: Public Library of Science (2007).

[74] Shutt, D. P., D. W. Goodsman, K. Martinez, Z. J. L. Hemez, J. R. Conrad, C. Xu, D. Osthus, C. Russell, J. M. Hyman and C. A. Manore, "A Process-based Model with Temperature, Water, and Lab-derived Data Improves Predictions of Daily Culex pipiens/restuans Mosquito Density", Journal of Medical Entomology **59**, 6, 1947–1959, URL `https://doi.org/10.1093/jme/tjac127` (2022).

[75] Simpson, J. E., P. J. Hurtado, J. Medlock, G. Molaei, T. G. Andreadis, A. P. Galvani and M. A. Diuk-Wasser, "Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile virus as a model system", Proceedings. Biological Sciences **279**, 1730, 925–933 (2012).

[76] Smithburn, K. C., T. P. Hughes, A. W. Burke and J. H. Paul, "A Neurotropic Virus Isolated from the Blood of a Native of Uganda", The American Journal of Tropical Medicine and Hygiene **s1-20**, 4, 471–492, URL `https://www.ajtmh.org/view/journals/tpmd/s1-20/4/article-p471.xml`, publisher: The American Society of Tropical Medicine and Hygiene Section: The American Journal of Tropical Medicine and Hygiene (1940).

[77] Sobol, I., "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates", Mathematics and Computers in Simulation URL `https://www.semanticscholar.org/paper/Global-sensitivity-indices-for-nonlinear-models-and-Sobol/f788d77d621fdc72e469d694fbdc61fa1dc0a914` (2001).

[78] Solomon, S., I. P. on Climate Change and I. P. on Climate Change, eds., *Climate change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge ; New York, 2007), oCLC: ocn132298563.

[79] Styer, L. M., J. R. Carey, J.-L. Wang and T. W. Scott, "Mosquitoes do senesce: departure from the paradigm of constant mortality", The American Journal of Tropical Medicine and Hygiene **76**, 1, 111–117 (2007).

[80] Swaddle, J. P. and S. E. Calos, "Increased avian diversity is associated with lower incidence of human West Nile infection: observation of the dilution effect", PloS One **3**, 6, e2488 (2008).

[81] Temple, S. D., C. A. Manore and K. A. Kaufeld, "Bayesian time-varying occupancy model for West Nile virus in Ontario, Canada", Stochastic Environmental Research and Risk Assessment **36**, 8, 2337–2352, URL `https://doi.org/10.1007/s00477-022-02257-4` (2022).

[82] Trejo, I., M. Barnard, J. A. Spencer, J. Keithley, K. M. Martinez, I. Crooker, N. Hengartner, E. O. Romero-Severson and C. Manore, "Changing temperature profiles and the risk of dengue outbreaks", PLOS Climate **2**, 2, e0000115, URL `https://journals.plos.org/climate/article?id=10.1371/journal.pclm.0000115`, publisher: Public Library of Science (2023).

[83] Vance, R. R. and E. A. Coddington, "A nonautonomous model of population growth", Journal of Mathematical Biology **27**, 5, 491–506 (1989).

[84] VanDalen, K., J. Hall, L. Clark, R. McLean and C. Smeraski, "West Nile Virus Infection in American Robins: New Insights on Dose Response", URL `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0068537` (2013).

[85] Verhulst, P., "Notice sur la loi que la population suit dans son accroissement", Correspondance mathématique et physique URL `https://www.scienceopen.com/document?vid=b51b58ae-889e-49ac-bc75-759cffcdc238` (1838).

[86] Vinogradova, E. B., *Culex Pipiens Pipiens Mosquitoes: Taxonomy, Distribution, Ecology, Physiology, Genetics, Applied Importance and Control* (Pensoft Publishers, 2000), URL `https://www.nhbs.com/culex-pipiens-pipiens-mosquitoes-taxonomy-distribution-ecology-physiology-genetics-applied-importance-and-control-book`.

[87] Wan, H. and H. Zhu, "The backward bifurcation in compartmental models for West Nile virus", Mathematical Biosciences **227**, 1, 20–28 (2010).

[88] Wang, Y., W. Pons, J. Fang and H. Zhu, "The impact of weather and storm water management ponds on the transmission of West Nile virus", Royal Society Open Science **4**, 8, 170017, URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5579078/` (2017).

[89] Wimberly, M. C., J. K. Davis, M. V. Evans, A. Hess, P. M. Newberry, N. Solano-Asamoah and C. C. Murdock, "Land cover affects microclimate and temperature suitability for arbovirus transmission in an urban landscape", PLOS Neglected Tropical Diseases **14**, 9, e0008614, URL `https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0008614`, publisher: Public Library of Science (2020).

[90] Wonham, M. J., M. A. Lewis, J. Rencławowicz and P. van den Driessche, "Transmission assumptions generate conflicting predictions in host-vector disease models: a case study in West Nile virus", Ecology Letters **9**, 6, 706–725 (2006).

[91] World Health Organization, "Chikungunya", URL `https://www.who.int/news-room/fact-sheets/detail/chikungunya` (2020).

[92] World Health Organization, "Dengue and severe dengue", URL `https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue` (2023).

[93] World Health Organization, "Fact sheet about malaria", URL `https://www.who.int/news-room/fact-sheets/detail/malaria` (2023).

[94] Wu, X., K. Shirvan and T. Kozlowski, "On the connection between sensitivity and identifiability for inverse uncertainty quantification: 2018 Transactions of the American Nuclear Society, ANS 2018 and Embedded Topical Meeting Nuclear Fuels and Structural Materials", Transactions of the American Nuclear Society **118**, 411–414, URL `http://www.scopus.com/inward/record.url?scp=85062995468&partnerID=8YFLogxK` (2018).

[95] Ziolkowski, D., K. Parieck and S. J.R., "On the Road Again for a Bird Survey that Counts | U.S. Geological Survey", Birding URL `https://www.usgs.gov/media/files/road-again-a-bird-survey-counts` (2010).

APPENDIX A

JOURNAL PERMISSION

Some of the material in Chapter 2 was published in [49], which is published as an open access article distributed under the terms of the Creative Commons Attribution License in Frontiers in Applied Mathematics and Statistics.

APPENDIX B

LINKAGE MODEL

Table B.1: Candidate Models for the Model Selection Procedure. Models A–D Consider a Constant Carrying Capacity, While Models E–K Consider a Time-Varying Carrying Capacity. Entries with a Solid Line (“———”) Indicate that the Parameter was not Fitted for that Model, and Rather was Held Constant at the Initial Value. For Entries with Two Lines, the Upper Line Refers to the Values Used for the Total Mosquito Population, and the Lower Line Refers to the Values Used for the Active Mosquito Population. The Initial K or K_b Value for Each Model is the Maximum Value of the Mosquito Process-Based Model Time-Series used for Each Season's Fitting. ”Parameters Fitted” Refers to the Parameter Components which were Optimized for Each Season. ”Initial” Columns Indicate Initial Values of the Respective Parameter for Each Year. ”Range” Columns Refer to the Constraint Bounds ([Min, Max]) that Directed the Search Space for the Respective Parameter Optimization.

Model	r_b initial	r_b range	r_s initial	r_s range	K or K_b range	K_s initial	K_s range
A	0.01	———-	-0.07	[-0.35, 0]	———-	——	———-
B	0	[-0.2, 0.2]	-0.07	[-0.35, 0]	———-	——	———-
C	0.01	———-	-0.07	[-0.35, 0]	[1, 100,000] [1, 1,000]	——	———-
D	0	[-0.2, 0.2]	-0.07	[-0.35, 0]	[1, 100,000] [1, 1,000]	——	———-
E	0.01	———-	-0.07	[-0.35, 0]	———-	100 1	———-
F	0	[-0.2, 0.2]	-0.07	[-0.35, 0]	———-	100 1	———-
G	0.01	———-	-0.07	———-	[1, 100,000] [1, 1,000]	100 1	[0, 100,000] [0, 1,000]
H	0.01	———-	-0.07	[-0.35, 0]	———-	100 1	[0, 100,000] [0, 1,000]
J	0	[-0.2, 0.2]	-0.07	[-0.35, 0]	———-	100 1	[0, 100,000] [0, 1,000]
K	0	[-0.2, 0.2]	-0.07	[-0.35, 0]	[1, 100,000] [1, 1,000]	100 1	[0, 100,000] [0, 1,000]

Table B.2: Fitted Parameter Values of the Model (2.2) for Greater Toronto Area's Total Mosquito Population from Years 2005–2019. For Each Year, the Mosquito Fitting Season Lasts from the Indicated Start Date Through End Date (mm/dd/yyyy Format).

Year	**Start Date**	**End Date**	r_b	r_s	K_b	K_s	$P(0)$	**RMSE**
2005	05/11/2005	10/01/2005	-0.003	-0.056	32174	0	627.31	1242
2006	05/20/2006	10/01/2006	-0.065	-0.328	26771	24996	187.51	1195
2007	05/20/2007	10/01/2007	-0.007	-0.137	7749	2408	444.91	684
2008	05/18/2008	10/01/2008	-0.014	-0.110	63880	63278	499.13	1229
2009	06/01/2009	11/01/2009	-0.024	-0.350	18456	16695	57.00	1247
2010	05/01/2010	11/01/2010	0.001	-0.076	51744	0	176.74	903
2011	05/07/2011	10/01/2011	0.001	-0.069	51421	6004	255.78	383
2012	05/09/2012	10/01/2012	-0.007	-0.075	15458	0	721.52	938
2013	05/01/2013	11/01/2013	-0.005	-0.068	100000	0	529.20	3524
2014	05/25/2014	10/09/2014	-0.031	-0.182	8248	0	248.95	1097
2015	05/18/2015	10/17/2015	-0.004	-0.114	9452	0	402.98	647
2016	06/01/2016	10/03/2016	-0.015	-0.067	51915	46951	3293.05	580
2017	05/18/2017	10/01/2017	-0.027	-0.091	100000	0	803.38	1535
2018	06/01/2018	10/16/2018	-0.013	-0.048	100000	93746	5379.55	798
2019	05/04/2019	10/01/2019	-0.008	-0.101	98758	98758	1916.85	1221

Table B.3: Fitted Parameter Values of the Model's (2.2) Mosquito Fitting Season for Greater Toronto Area's Active Mosquito Population from Years 2005–2019. For Each Year, the Mosquito Fitting Season Lasts from the Indicated Start Date through End Date (mm/dd/yyyy Format).

Year	**Start Date**	**End Date**	r_b	r_s	K_b	K_s	$P(0)$	**RMSE**
2005	05/09/2005	05/26/2006	-0.012	-0.094	826.4	826.0	1.22	0.59
2006	05/27/2006	05/12/2007	-0.101	-0.350	43.2	40.6	0.25	0.79
2007	05/13/2007	05/30/2008	0.002	-0.250	11.3	6.2	0.01	0.62
2008	05/31/2008	05/30/2009	-0.009	-0.350	37.3	34.3	0.42	0.75
2009	05/31/2009	05/06/2010	-0.036	-0.350	23.7	22.2	0.12	0.98
2010	05/07/2010	05/02/2011	-0.011	-0.101	182.9	0.0	0.16	0.57
2011	05/03/2011	05/29/2012	0.026	-0.198	22.2	22.1	0.01	0.44
2012	05/30/2012	05/19/2013	-0.083	-0.175	18.4	2.1	2.03	0.53
2013	05/20/2013	04/30/2014	-0.091	-0.21	1000	0.0	0.17	1.70
2014	05/01/2014	05/19/2015	-0.002	-0.154	14.5	0.0	0.01	0.87
2015	05/20/2015	05/14/2016	-0.037	-0.134	16.2	0.0	0.53	0.39
2016	05/15/2016	04/30/2017	0.027	-0.329	18.5	14.8	0.01	0.59
2017	05/01/2017	05/12/2018	-0.003	-0.145	46.4	0.0	0.01	0.49
2018	05/13/2018	04/30/2019	0.034	-0.317	19.7	13.4	0.01	0.82
2019	05/01/2019	04/30/2020	-0.004	-0.139	94.4	0.0	0.01	0.60

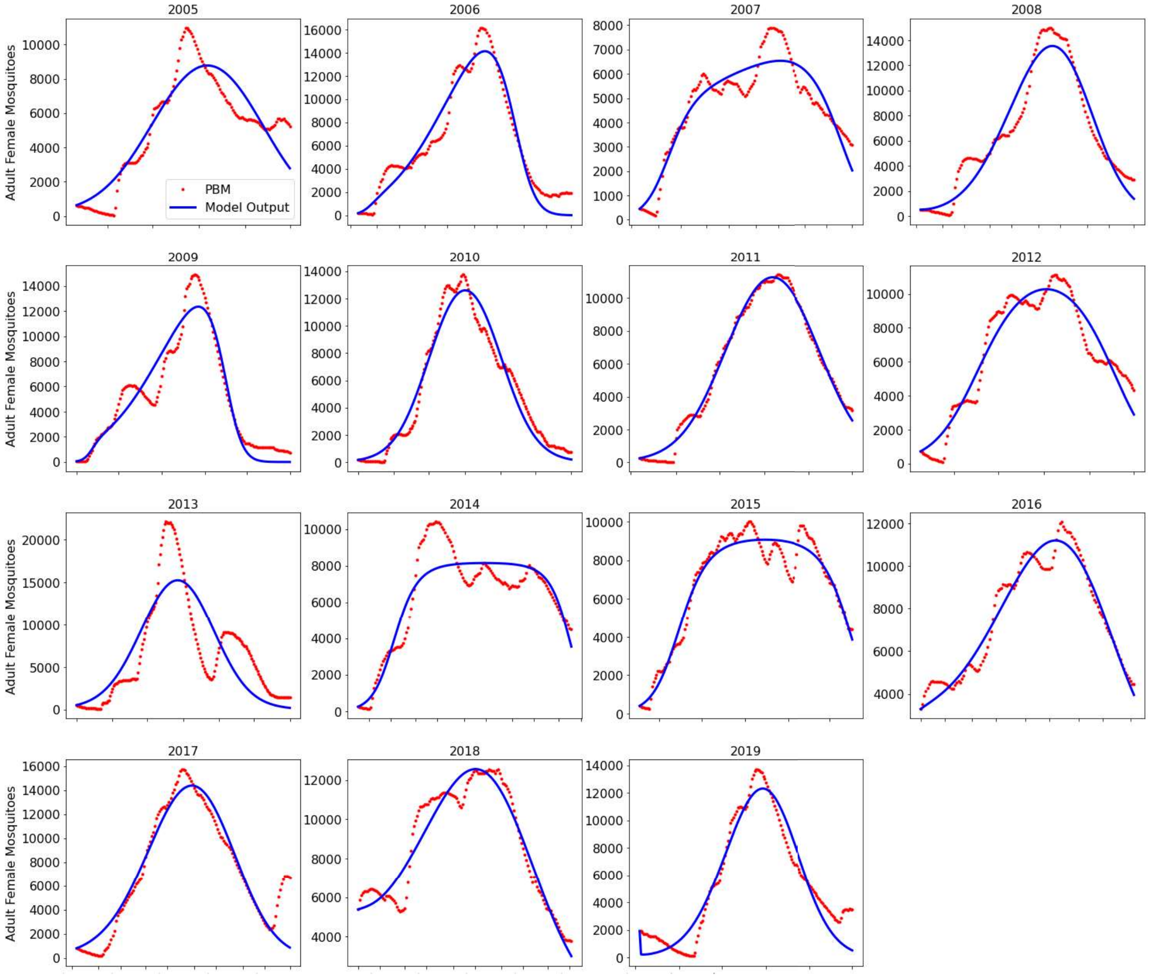


Figure B-1: Simulations of the Optimal Model Fits for the Total Population in the GTA (Blue Curves) and PBM Time Series (Red Dots) for Years 2005–2019.

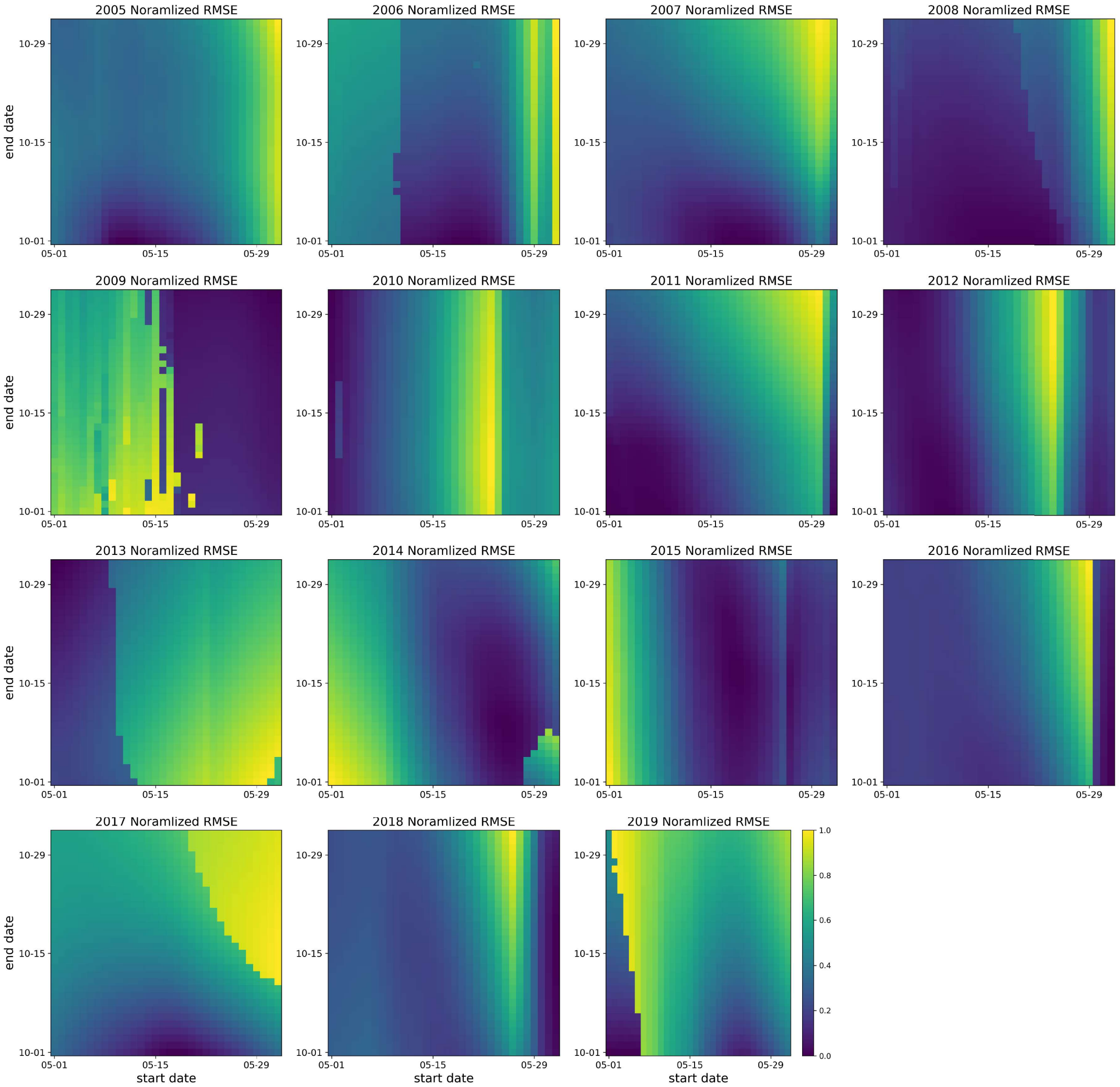


Figure B-2: Heatmaps of the Normalized Root Mean Squared Error (RMSE) Values from Total Population Fittings with Respect to Candidate Start and End Dates for Years 2005–2019. Dark Purple Regions Denote the Candidate Start and End Date Combinations with the Lowest Normalized Values.

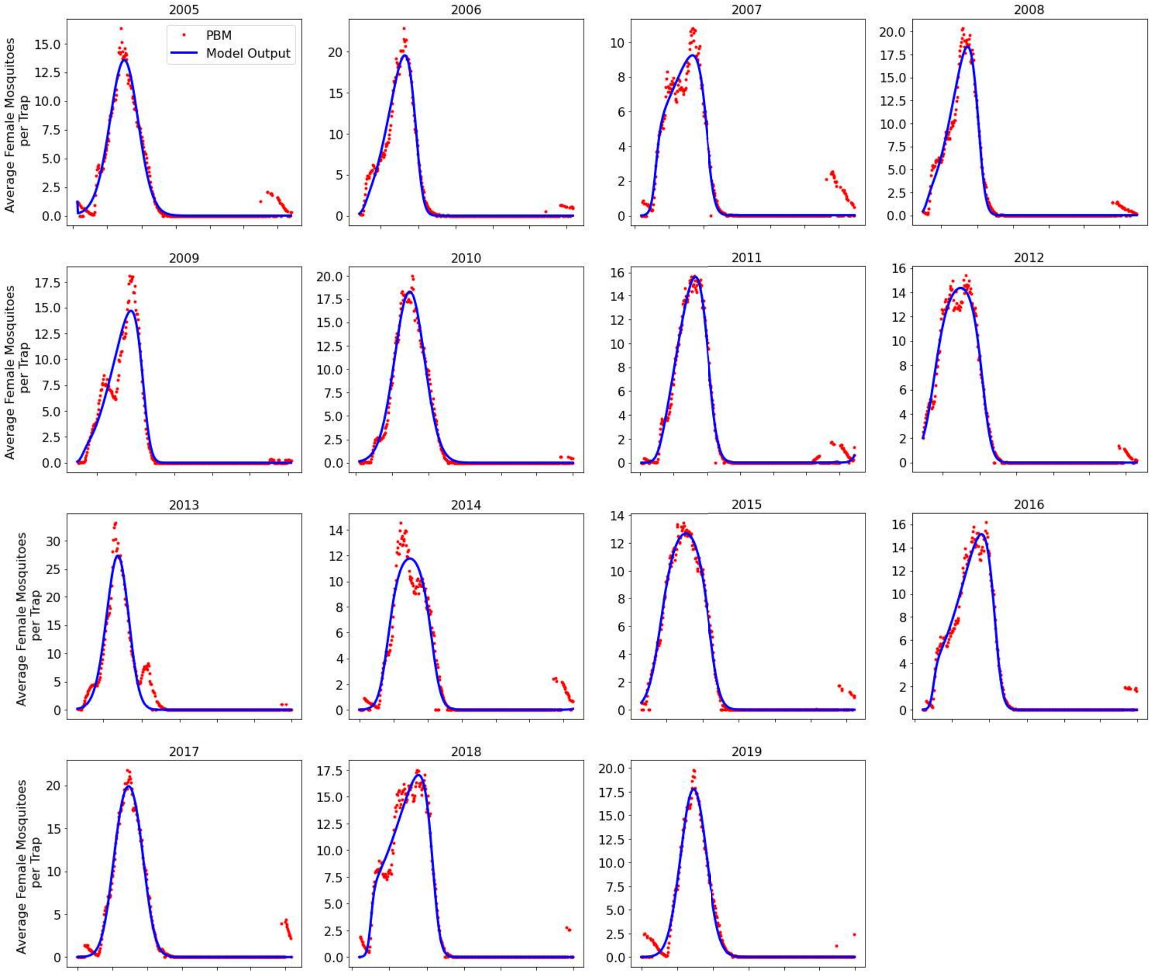


Figure B-3: Simulations of the Optimal Model fits for the Active Population in the GTA (Blue Curves) and PBM Time Series (Red Dots) for Years 2005–2019.

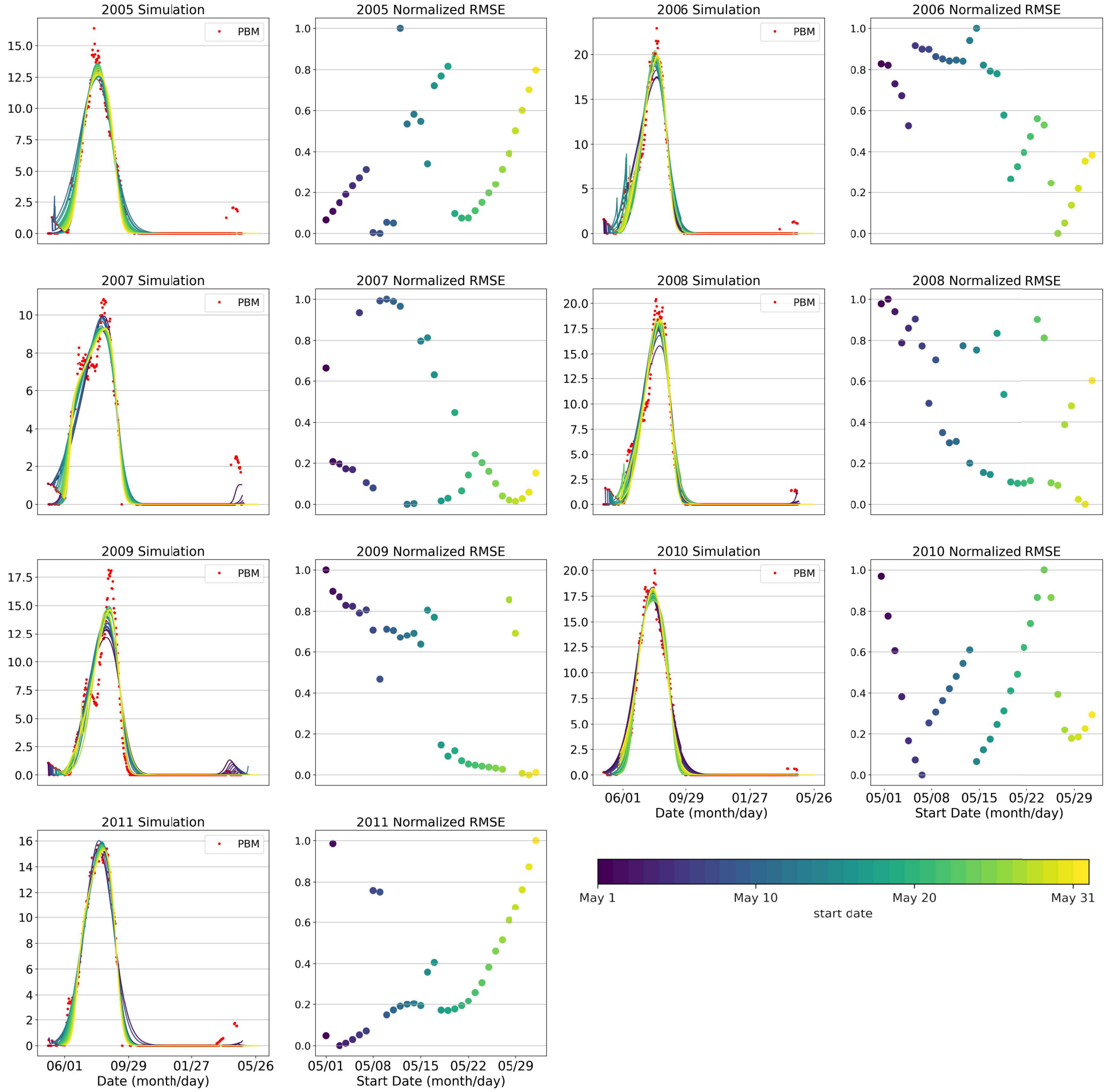


Figure B-4: Simulations (Curves) and Normalized Root Mean Squared Error Values of the Active Population in the Greater Toronto Area for Fitting Start Dates May 1st–June 1st for Years 2005–2011.

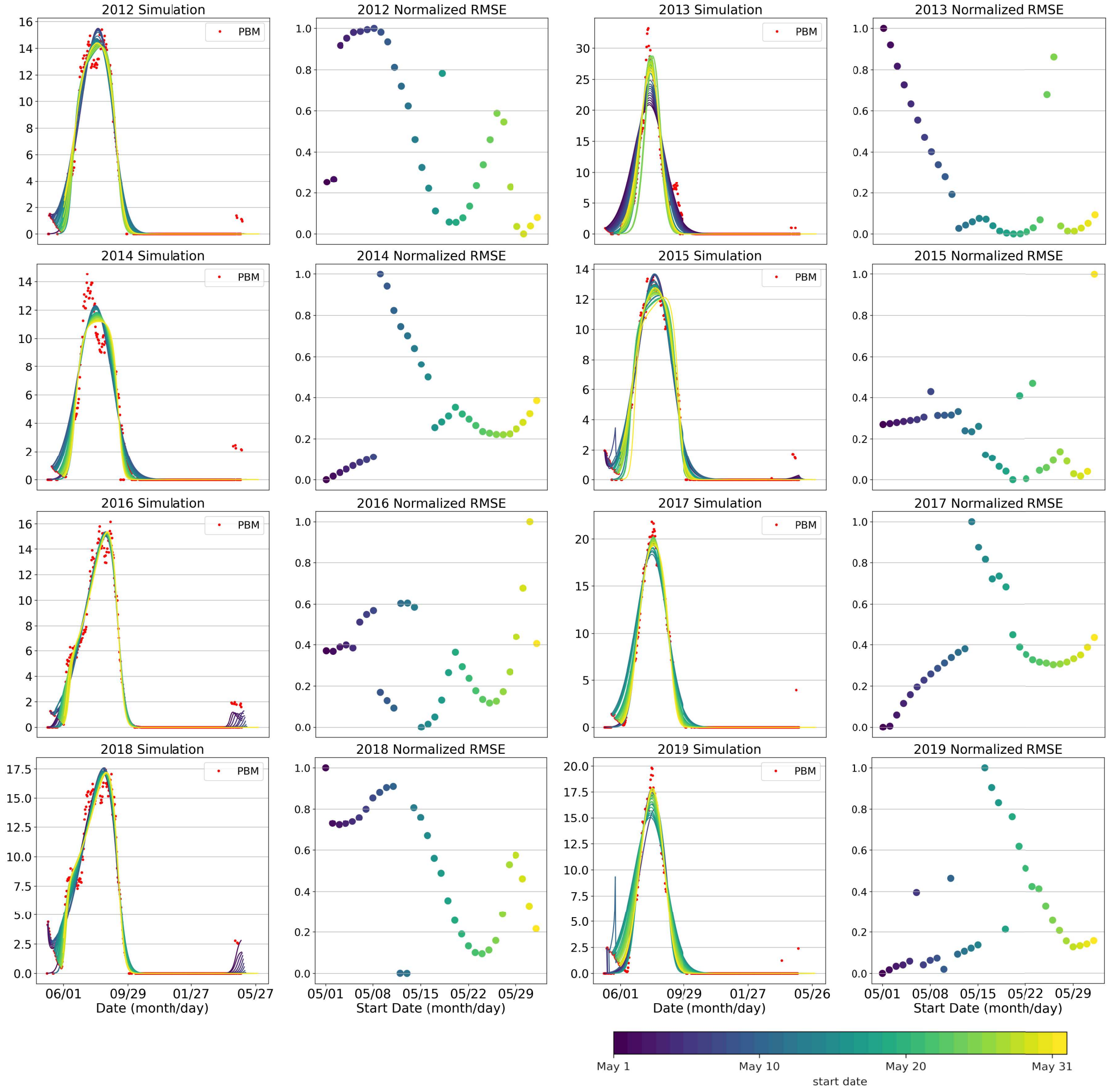


Figure B-5: Simulations (Curves) and Normalized Root Mean Squared Error Values of the Active Population in the Greater Toronto Area for Fitting Start Dates May 1st–June 1st for Years 2012–2019.

APPENDIX C

SIMPLIFIED WNV MODEL

C.1 Method of Characteristics for Infected Birds

Let $\xi_1, \xi_2, v : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be the characteristics of (3.3.4), where $\tau = \xi_1$, $t = \xi_2$, and $i_b = v$. We aim to find solutions $z \in \mathbb{R}_+$ and $t \in \mathbb{R}$ such that,

$$\xi(z,y) = \big(\xi_1(z,y), \xi_2(z,y)\big), \tag{C.1.1}$$

$$i_b\big(\xi(z,y)\big) = v(z,y). \tag{C.1.2}$$

Case 1: Suppose $t < \tau$. Then the system (C.1) transforms into the following initial value problem (IVP):

$$\frac{d\xi_1}{dy} = 1, \qquad \xi_1(z,0) = z, \tag{C.2.1}$$

$$\frac{d\xi_2}{dy} = 1, \qquad \xi_1(z,0) = 0, \tag{C.2.2}$$

$$\frac{dv}{dy} = -\gamma_b\big(\xi(z,y)\big)v(z,y), \qquad v(z,0) = i_b^0(z). \tag{C.2.3}$$

Integrating (C.2.1) and (C.2.2) gives:

$$\xi_1(z,y) = y + c_1(z), \tag{C.3.1}$$

$$\xi_2(z,y) = y + c_2(z), \tag{C.3.2}$$

where $c_1(z)$ and $c_2(z)$ are functions of z. Applying initial conditions gives:

$$c_1(z) = z, \tag{C.3.3}$$

$$c_2(z) = 0. \tag{C.3.4}$$

Substituting (C.3.3) and (C.3.4) into (C.3.1) and (C.3.2) gives the expression for $\xi(z,y)$:

$$\xi(z,y) = \big(\xi_1(z,y), \xi_2(z,y)\big) = \big(y + c_1(z), y + c_2(z)\big) = (y + z, y). \tag{C.4}$$

Substituting (C.4) into (C.2.3) yields:

$$\frac{dv}{dy} = -\gamma_b(y+z)v(z,y),$$

and solving by separation of variables gives

$$v(z,y) = v(z,0)e^{-\int_0^y \gamma_b(s+z)ds}. \tag{C.5}$$

Noting that $(\tau, t) = \xi(z,y)$, we can solve for z and y as functions of τ and t:

$$\tau = \xi_1(z,y) = z + y,$$

$$t = \xi_2(z,y) = y,$$

implies,

$$\begin{aligned} y &= t, \\ z &= \tau - t. \end{aligned}$$

Transforming (C.5) back to original variables gives:

$$i_b(\tau, t) = i_b^0(\tau - t)e^{-\int_0^t \gamma_b(\tau - t + s)ds}, \quad \text{if } t < \tau. \tag{C.6}$$

Let $\Pi_b(\tau) = e^{-\int_0^\tau \gamma_b(s)ds}$. Using the change of variable, $a = \tau - t + s$,

$$\begin{aligned} e^{-\int_0^t \gamma_b(\tau - t + s)ds} &= e^{-\int_{\tau - t}^{\tau} \gamma_b(a)da} \\ &= e^{-\int_0^\tau \gamma_b(a)da + \int_0^{\tau - t} \gamma_b(a)da} \\ &= \frac{e^{-\int_0^\tau \gamma_b(a)da}}{e^{-\int_0^{\tau - t} \gamma_b(a)da}} \\ &= \frac{\Pi_b(\tau)}{\Pi_b(\tau - t)}. \end{aligned}$$

Therefore, (C.6) becomes,

$$i_b(\tau, t) = i_b^0(\tau - t)\tfrac{\Pi_b(\tau)}{\Pi_b(\tau - t)}, \quad \text{if } t < \tau. \tag{C.7}$$

Case 2: Suppose now that $\tau \leq t$. Then the system (C.1) transforms into the following IVP:

$$\frac{d\xi_1}{dz} = 1, \qquad \xi_1(0, y) = 0, \tag{C.8.1}$$

$$\frac{d\xi_2}{dz} = 1, \qquad \xi_1(0, y) = y, \tag{C.8.2}$$

$$\frac{dv}{dz} = -\gamma_b\big(\xi(z, y)\big)v(z, y), \qquad v(0, y) = \lambda_{vb}(y)S_b(y). \tag{C.8.3}$$

Integrating (C.8.1) and (C.8.2) gives:

$$\xi_1(z, y) = z + c_3(y), \tag{C.9.1}$$

$$\xi_2(z, y) = z + c_4(y), \tag{C.9.2}$$

where $c_3(y)$ and $c_4(y)$ are functions of y. Applying initial conditions gives:

$$c_3(y) = 0, \tag{C.9.3}$$

$$c_4(y) = y. \tag{C.9.4}$$

Substituting (C.9.3) and (C.9.4) into (C.9.1) and (C.9.2) gives the expression for $\xi(z, y)$:

$$\xi(z, y) = \big(\xi_1(z, y), \xi_2(z, y)\big) = \big(z + c_3(y), z + c_4(y)\big) = (z, y + z). \tag{C.10}$$

Substituting (C.10) into (C.8.3) yields:

$$\frac{dv}{dz} = -\gamma_b(z)v(z,y),$$

and solving by separation of variables gives,

$$v(z,y) = \lambda_{vb}(y)S_b(y)e^{-\int_0^z \gamma_b(s)ds}. \tag{C.11}$$

Noting that $(\tau, t) = \xi(z, y)$, we can solve for z and y as functions of τ and t:

$$\begin{aligned} \tau &= \xi_1(z,y) = z, \\ t &= \xi_2(z,y) = z + y, \end{aligned}$$

implies,

$$\begin{aligned} z &= \tau, \\ y &= t - \tau. \end{aligned}$$

Transforming (C.11) back to original variables gives:

$$i_b(\tau,t) = \lambda_{vb}(t-\tau)S_b(t-\tau)\Pi_b(\tau), \quad \text{if } \tau \le t. \tag{C.12}$$

To verify that the initial and boundary conditions are compatible, suppose $t = \tau = a$. Then,

$$i_b(a,a) = i_b^0(0)\Pi_b(a), \quad \text{from (C.7)}, \tag{C.13.1}$$

$$i_b(a,a) = \lambda_{vb}(0)S_b^0\Pi_b(a), \quad \text{from (C.12)}. \tag{C.13.2}$$

Thus, in order for (C.13.1) and (C.13.2) to be equivalent, it follows that $i_b^0(0) = \lambda_{vb}(0)S_b^0$.

C.2 Integral Form of the Model

The integral form of (3.3.1) is found using variation of parameters:

$$S_v(t) = S_v^0 \Pi_v(t) e^{-\int_0^t \lambda_{bv}(s)ds} + \Lambda_v \int_0^t e^{-\int_0^{t-s} \lambda_{bv}(a)da} \Pi_v(t-s) ds. \tag{C.14.1}$$

Applying the change of variable, $\eta = t - s$, the integral form of (3.3.1) becomes,

$$S_v(t) = S_v^0 \Pi_v(t) e^{-\int_0^t \lambda_{bv}(s)ds} + \Lambda_v \int_0^t e^{-\int_0^{\eta} \lambda_{bv}(a)da} \Pi_v(\eta) d\eta. \tag{C.14.2}$$

To obtain the integral form of (3.3.2), integrate (3.6.1) with respect to τ:

$$\begin{aligned} I_v(t) &= \int_0^\infty i_v(\tau, t) d\tau \\ &= \int_0^t \lambda_{bv}(t-\tau) S_v(t-\tau) \Pi_v(\tau) d\tau + \underbrace{\int_t^\infty i_v^0(\tau - t) \frac{\Pi_v(\tau)}{\Pi_v(\tau - t)} d\tau}_{F_1(t)}. \end{aligned}$$

Applying the change of variable, $\eta = t - \tau$ gives:

$$I_v(t) = \int_0^t \lambda_{bv}(\eta) S_v(\eta) \Pi_v(t - \eta) d\eta + F_1(t).$$

By substituting (C.14.2) for $S_v(\eta)$, the expression for $I_v(t)$ becomes,

$$\begin{aligned} I_v(t) = & S_v^0 \Pi_v(t) \int_0^t \lambda_{bv}(\eta) e^{-\int_0^{\eta} \lambda_{bv}(a)da} d\eta \\ & + \Lambda_v \int_0^t \int_0^{\eta} \lambda_{bv}(\eta) \Pi_v(t-\eta) \Pi_v(\eta - s) e^{-\int_0^{\eta - s} \lambda_{bv}(a)da} ds d\eta + F_1(t). \end{aligned}$$

Changing the order of integration for the second term of $I_v(t)$ simplifies the expression:

$$\begin{aligned} I_v(t) = & S_v^0 \Pi_v(t) \int_0^t \lambda_{bv}(\eta) e^{-\int_0^{\eta} \lambda_{bv}(a)da} d\eta \\ & + \Lambda_v \int_0^t \Pi_v(t-s) \int_s^t \lambda_{bv}(\eta) e^{-\int_0^{\eta - s} \lambda_{bv}(a)da} d\eta ds + F_1(t). \end{aligned} \tag{C.15}$$

The integral form of (3.3.3) is found using separation of variables:

$$S_b(t) = S_b^0 e^{-\int_0^t \lambda_{vb}(s)ds}. \tag{C.16}$$

To get the integral form of (3.3.4), integrate (3.6.2) with respect to τ:

$$I_b(t) = \int_0^\infty i_b(\tau, t) d\tau$$

$$= \int_0^t \lambda_{vb}(t-\tau)S_b(t-\tau)\Pi_b(\tau)d\tau + \underbrace{\int_t^\infty i_b^0(\tau - t)\frac{\Pi_b(\tau)}{\Pi_b(\tau - t)}d\tau}_{F_2(t)}.$$

Applying the change of variable, $\eta = t - \tau$ gives:

$$I_b(t) = \int_0^t \lambda_{vb}(\eta)S_b(\eta)\Pi_b(t-\eta)d\eta + F_2(t).$$

Substituting (C.16) for $S_b(\eta)$ in the expression for $I_b(t)$ further gives:

$$I_b(t) = S_b^0 \int_0^t \lambda_{vb}(\eta)\Pi_b(t-\eta)e^{-\int_0^t \lambda_{vb}(s)ds}d\eta + F_2(t). \tag{C.17}$$

To obtain the integral form of (3.3.5), integrate with respect to t:

$$\begin{aligned} R_b(t) &= \int_0^t \int_0^\infty \gamma_b(\tau)i_b(\tau,s)d\tau ds \\ &= \int_0^t \int_0^s \gamma_b(\tau)i_b(\tau,s)d\tau ds + \int_0^t \int_s^\infty \gamma_b(\tau)i_b(\tau,s)d\tau ds. \end{aligned}$$

Substitute (3.6.2) to get,

$$R_b(t) = \int_0^t \int_0^s \gamma_b(\tau)\lambda_{vb}(s-\tau)S_b(s-\tau)d\tau ds + \underbrace{\int_0^t \int_s^\infty \gamma_b(\tau)i_b^0(\tau - s)\frac{\Pi_b(\tau)}{\Pi_b(\tau - s)}d\tau ds}_{F_3(t)}.$$

Substitute in (C.16) for S_b to obtain,

$$R_b(t) = S_b^0 \int_0^t \int_0^s \gamma_b(\tau)\lambda_{vb}(s-\tau)\Pi_b(\tau)e^{-\int_0^{s-\tau} \lambda_{vb}(a)da}d\tau ds + F_3(t). \tag{C.18}$$

The integral form of (3.3.6) is found using separation of variables:

$$S_h(t) = S_h^0 e^{-\int_0^t \lambda_{vh}(s)ds}. \tag{C.19}$$

To obtain the integral form of (3.3.7), first substitute (C.19) for $S_h(t)$:

$$\frac{dI_h}{dt} = S_h^0 \lambda_{vh}(t)e^{-\int_0^t \lambda_{vh}(s)ds} - \gamma_h I_h(t).$$

Then solve using variation of parameters to obtain:

$$I_h(t) = I_h^0 \underbrace{e^{-\gamma_h(t)}}_{\Pi_h(t)} + S_h^0 \int_0^t \lambda_{vh}(s)e^{-\int_0^s \lambda_{vh}(a)da}\underbrace{e^{-\gamma_h(t-s)}}_{\Pi_h(t-s)}ds. \tag{C.20}$$

Finally, to obtain the integral form of (3.3.8), substitute (C.20) for $I_h(t)$:

$$\frac{dR_h}{dt} = \gamma_h I_h^0 \Pi_h(t) + \gamma_h S_h^0 \int_0^t \lambda_{vh}(s)\Pi_h(t-s)e^{-\int_0^s \lambda_{vh}(a)da}ds.$$

Integrating with respect to t and simplifying terms gives:

$$R_h(t) = \gamma_h S_h^0 \int_0^t \int_0^\eta \lambda_{vh}(s)\Pi_h(\eta - s)e^{-\int_0^\eta \lambda_{vh}(a)da}dsd\eta - \frac{\gamma_h I_h^0 \Pi_h(t)}{\mu_v}. \tag{C.21}$$

To get the integral form of (3.4), we first substitute (3.6) for $i_v(\tau, t)$ and $i_b(\tau, t)$:

$$\lambda_{vb}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{vb}(\tau)\lambda_{bv}(t-\tau)S_v(t-\tau)\Pi_v(\tau)d\tau + \int_t^\infty \beta_{vb}(\tau)i_v^0(\tau - t)\frac{\Pi_v(\tau)}{\Pi_v(\tau - t)}d\tau\right),$$

$$\lambda_{vh}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{vh}(\tau)\lambda_{bv}(t-\tau)S_v(t-\tau)\Pi_v(\tau)d\tau + \int_t^\infty \beta_{vh}(\tau)i_v^0(\tau - t)\frac{\Pi_v(\tau)}{\Pi_v(\tau - t)}d\tau\right),$$

$$\lambda_{bv}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{bv}(\tau)\lambda_{vb}(t-\tau)S_b(t-\tau)\Pi_b(\tau)d\tau + \int_t^\infty \beta_{bv}(\tau)i_b^0(\tau - t)\frac{\Pi_b(\tau)}{\Pi_b(\tau - t)}d\tau\right).$$

Applying the changes of variables $\eta = t - \tau$ and $\eta = \tau - t$ to the first and second terms of each expression, respectively yields:

$$\lambda_{vb}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{vb}(t-\eta)\lambda_{bv}(\eta)S_v(\eta)\Pi_v(t-\eta)d\eta + \underbrace{\int_0^\infty \beta_{vb}(t+\eta)i_v^0(\eta)\frac{\Pi_v(t+\eta)}{\Pi_v(\tau - t)}d\eta}_{G_1(t)}\right), \tag{C.22.1}$$

$$\lambda_{vh}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{vh}(t-\eta)\lambda_{bv}(\eta)S_v(\eta)\Pi_v(t-\eta)d\eta + \underbrace{\int_0^\infty \beta_{vh}(t+\eta)i_v^0(\eta)\frac{\Pi_v(t+\eta)}{\Pi_v(\eta)}d\eta}_{G_2(t)}\right), \tag{C.22.2}$$

$$\lambda_{bv}(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{bv}(t-\eta)\lambda_{vb}(\eta)S_b(\eta)\Pi_b(t-\eta)d\eta + \underbrace{\int_0^\infty \beta_{bv}(t+\eta)i_b^0(\eta)\frac{\Pi_b(t+\eta)}{\Pi_b(\eta)}d\eta}_{G_3(t)}\right). \tag{C.22.3}$$

Substituting (C.14.2) and (C.16) for $S_v(\eta)$ and $S_b(\eta)$ further yields:

$$\lambda_{vb}(t) = \frac{\alpha}{N_b + N_h}\Bigg(S_v^0\Pi_v(t)\int_0^t \beta_{vb}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^\eta \lambda_{bv}(s)ds}d\eta$$

$$
\begin{aligned}
&\quad + \Lambda_v \int_0^t \int_0^\eta \beta_{vb}(t-\eta)\lambda_{bv}(\eta)\Pi_v(\eta-s)\Pi_v(t-\eta)e^{-\int_0^{\eta-s}\lambda_{bv}(a)da}dsd\eta + G_1(t)\Bigg),\\
\lambda_{vh}(t) &= \frac{\alpha}{N_b+N_h}\Bigg(S_v^0\Pi_v(t)\int_0^t \beta_{vh}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^\eta \lambda_{bv}(s)ds}d\eta\\
&\quad + \Lambda_v \int_0^t \int_0^\eta \beta_{vh}(t-\eta)\lambda_{bv}(\eta)\Pi_v(\eta-s)\Pi_v(t-\eta)e^{-\int_0^{\eta-s}\lambda_{bv}(a)da}dsd\eta + G_2(t)\Bigg),\\
\lambda_{bv}(t) &= \frac{\alpha}{N_b+N_h}\Bigg(S_b^0\int_0^t \beta_{bv}(t-\eta)\lambda_{vb}(\eta)e^{-\int_0^\eta \lambda_{vb}(s)ds}\Pi_b(t-\eta)d\eta + G_3(t)\Bigg).
\end{aligned}
\tag{C.23.1}
$$

Finally, simplify the expressions for $\lambda_{vb}(t)$ and $\lambda_{vh}(t)$ by changing the order of integration for their second terms:

$$
\begin{aligned}
\lambda_{vb}(t) = \frac{\alpha}{N_b+N_h}\Bigg(&S_v^0\Pi_v(t)\int_0^t \beta_{vb}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^\eta \lambda_{bv}(s)ds}d\eta\\
&+ \Lambda_v \int_0^t \Pi_v(t-s)\int_s^t \beta_{vb}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^{\eta-s}\lambda_{bv}(a)da}d\eta ds + G_1(t)\Bigg),
\end{aligned}
\tag{C.23.2}
$$

$$
\begin{aligned}
\lambda_{vh}(t) = \frac{\alpha}{N_b+N_h}\Bigg(&S_v^0\Pi_v(t)\int_0^t \beta_{vh}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^\eta \lambda_{bv}(s)ds}d\eta\\
&+ \Lambda_v \int_0^t \Pi_v(t-s)\int_0^\eta \beta_{vh}(t-\eta)\lambda_{bv}(\eta)e^{-\int_0^{\eta-s}\lambda_{bv}(a)da}d\eta ds + G_2(t)\Bigg).
\end{aligned}
\tag{C.23.3}
$$

$$
\lambda_{bv}(t) = \frac{\alpha}{N_b+N_h}\Bigg(S_b^0\int_0^t \beta_{bv}(t-\eta)\lambda_{vb}(\eta)e^{-\int_0^\eta \lambda_{vb}(s)ds}\Pi_b(t-\eta)d\eta + G_3(t)\Bigg).
$$

C.3 Proof of Existence and Uniqueness of the Model

For notational convenience, let $x(t) = \lambda_{vb}(t)$, $y(t) = \lambda_{vh}(t)$, and $z(t) = \lambda_{bv}(t)$. Then from (C.22),

$$x(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{vb}(t-\eta) z(\eta) S_v(\eta) \Pi_v(t-\eta) d\eta + G_1(t) \right),$$

$$z(t) = \frac{\alpha}{N_b + N_h}\left(\int_0^t \beta_{bv}(t-\eta) x(\eta) S_b(\eta) \Pi_b(t-\eta) d\eta + G_3(t) \right).$$

Let,

$$B_v(t) = \tfrac{\alpha}{N_b+N_h}\big(\beta_{vb}(t)\Pi_v(t)\big), \quad F_v(t) = \tfrac{\alpha}{N_b+N_h}\big(G_1(t)\big),$$
$$B_b(t) = \tfrac{\alpha}{N_b+N_h}\big(\beta_{bv}(t)\Pi_b(t)\big), \quad F_b(t) = \tfrac{\alpha}{N_b+N_h}\big(G_3(t)\big).$$

Then,

$$x(t) = \int_0^t B_v(t-\eta) z(\eta) S_v(\eta) d\eta + F_v(t), \tag{C.26.1}$$

$$z(t) = \int_0^t B_b(t-\eta) x(\eta) S_b(\eta) d\eta + F_b(t). \tag{C.26.2}$$

Substituting (3.8.1) and (3.8.3) into (C.26) yields,

$$x(t) = \int_0^t B_v(t-s) z(s) \Big(S_v^0 \Pi_v(s) e^{-\int_0^s z(a)da} + \Lambda_v \int_0^s e^{-\int_0^\eta z(a)da} \Pi_v(\eta) d\eta \Big) ds + F_v(t), \tag{C.27}$$

$$z(t) = \int_0^t B_b(t-s) x(s) S_b^0 e^{-\int_0^s x(a)da} ds + F_b(t). \tag{C.28}$$

Note that $B_v(t), B_b(t), F_v(t)$, and $F_b(t)$ are nonnegative continuous functions that vanish beyond $t > A^*$, and are therefore bounded. Further, note that $S_v(t)$ and $S_b(t)$ are nonnegative for $t \geq 0$. Let,

$$\mathfrak{F}_1(z(s)) = z(s) \Big(S_v^0 \Pi_v(s) e^{-\int_0^s z(a)da} + \Lambda_v \int_0^s e^{-\int_0^\eta z(a)da} \Pi_v(\eta) d\eta \Big), \tag{C.29}$$

$$\mathfrak{F}_2(x(s)) = x(s) S_b^0 e^{-\int_0^s x(a)da}, \tag{C.30}$$

Then,

$$x(t) = \int_0^t B_v(t-s) \mathfrak{F}_1(z(s)) ds + F_v(t), \tag{C.31.1}$$

$$z(t) = \int_0^t B_b(t-s) \mathfrak{F}_2(x(s)) ds + F_b(t). \tag{C.31.2}$$

$\mathfrak{F}_1$ is continuously differentiable in z and $\mathfrak{F}_2$ is continuously differentiable in x. Moreover, $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are uniformly continuous in $s \in [0, A_T]$. Therefore, $\mathfrak{F}_1$ is Lipschitz continuous in z and $\mathfrak{F}_2$ is Lipschitz continuous in x. I.e., there exists constants $L_1, L_2 > 0$ such that,

$$|\mathfrak{F}_1(z(s)) - \mathfrak{F}_1(\tilde{z}(s))| \leq L_1 |z(s) - \tilde{z}(s)|, \tag{C.32}$$
$$|\mathfrak{F}_2(x(s)) - \mathfrak{F}_2(\tilde{x}(s))| \leq L_2 |x(s) - \tilde{x}(s)|. \tag{C.33}$$

Let $\Omega_T = [0, A_T]$ and $X = C(\Omega_T) \times C(\Omega_T)$, where $C(\Omega_T)$ is the space of continuous functions on Ω_T. Further, X is endowed with the norm $\|x\|_X = \|x_1\| + \|x_2\|$, where $\|.\|$ is the usual norm on $C(\Omega_T)$. We define the operator $\Phi(x, z) = \big(\phi_x(x, z), \phi_z(x, z)\big)$ of X onto itself where the operators $\phi_x(x, z)$, and $\phi_z(x, z)$ are given by:

$$\phi_x(z)(t) = \int_0^t B_v(t - s)\mathfrak{F}_1(z(s))ds + F_v(t), \tag{C.34.1}$$
$$\phi_z(x)(t) = \int_0^t B_b(t - s)\mathfrak{F}_2(x(s))ds + F_b(t). \tag{C.34.2}$$

We prove the convergence of Φ to a fixed point by the iterative process:

$$\Phi^{(k)}(x, y, z) = \big(x^{(k)}, z^{(k)}\big) = \big(\phi_x(z^{(k-1)}), \phi_z(x^{(k-1)})\big), \quad \text{for } k = 1, 2, ... \tag{C.35}$$

I.e.,

$$x^{(k)}(t) = \phi_x(z^{(k-1)})(t) = \int_0^t B_v(t - s)\mathfrak{F}_1(z^{(k-1)}(s))ds + F_v(t), \tag{C.36}$$
$$z^{(k)}(t) = \phi_z(x^{(k-1)})(t) = \int_0^t B_b(t - s)\mathfrak{F}_2(x^{(k-1)}(s))ds + F_b(t), \tag{C.37}$$

where we take $\Phi^{(0)}(x, z) = (x^{(0)}, z^{(0)}) = (0, 0)$. Let,

$$\Phi_k(t) = \sup_{\tau \in [0,t]} |\Phi^{(k)}(\tau) - \Phi^{(k-1)}(\tau)|, \tag{C.38}$$

where,

$$|\Phi^{(k)}(\tau) - \Phi^{(k-1)}(\tau)| = |x^{(k)}(\tau) - x^{(k-1)}(\tau)| + |z^{(k)}(\tau) - z^{(k-1)}(\tau)|. \tag{C.39}$$

Then $\Phi_k(t)$ is a nonnegative, nondecreasing function of t and,

$$\Phi_k(A^*) = \|\Phi^{(k)} - \Phi^{(k-1)}\|_X. \tag{C.40}$$

Notice that,

$$\Phi_1(t) = \sup_{\tau \in [0,t]} ||x^{(1)}(\tau) - x^{(0)}(\tau)| + |z^{(1)}(\tau) - z^{(0)}(\tau)||, \tag{C.41}$$
$$= \sup_{\tau \in [0,t]} |F_v(\tau) + F_b(\tau)|, \tag{C.42}$$

and that,

$$\Phi_1(A^*) = \overline{F_v} + \overline{F_b}, \text{ where } \overline{F_i} = \sup_{t\in[0,A^*]} F_i(t), \text{ for } i = v, b. \tag{C.43}$$

By the Lipschitz continuity of $\mathfrak{F}_1$ and $\mathfrak{F}_2$,

$$\begin{aligned}\Phi_{k+1}(t) = &\int_0^t B_v(t-s)|\mathfrak{F}_1(z^{(k)}(s)) - \mathfrak{F}_1(z^{(k-1)}(s))|ds \\ &+ \int_0^t B_b(t-s)|\mathfrak{F}_2(x^{(k)}(s)) - \mathfrak{F}_2(x^{(k-1)}(s))|ds, \\ &\leq \overline{B_v}\int_0^t L_1|z^{(k)}(s) - z^{(k-1)}(s)|ds + \overline{B_b}\int_0^t L_2|x^{(k)}(s) - x^{(k-1)}(s)|ds, \end{aligned} \tag{C.44}$$

where,

$$\overline{B_i} = \sup_{t\in[0,A^*]} B_i, \text{ for } i = v, b. \tag{C.45}$$

Let $BL = \max\{\overline{B_v}L_1, \overline{B_b}L_2\}$. Then,

$$\Phi_{k+1}(t) \leq BL\int_0^t \Phi_k(s)ds. \tag{C.46}$$

By induction, it follows that,

$$\Phi_k(t) \leq \Phi_1(t)\frac{(BLt)^k}{k!} \leq (\overline{F_v} + \overline{F_b})\frac{(BLt)^k}{k!} \to 0 \text{ as } k \to \infty. \tag{C.47}$$

Therefore, the sequence $\{\Phi^{(k)}\}$ generated by (C.35) is a Cauchy sequence in X, and thus convergent. Hence, there exists a $(x(t), z(t))$ in X which is the limit of the sequence and is a fixed point for the operator Φ.

We prove the uniqueness of solutions by contradiction. That is, assume $(x(t), z(t))$ and $(\bar{x}(t), \bar{z}(t))$ are two solutions of (C.31). Then,

$$\begin{aligned}|x(t) - \bar{x}(t)| + |z(t) - \bar{z}(t)| = &\int_0^t B_v(t-s)|\mathfrak{F}_1(z(s)) - \mathfrak{F}_1(\bar{z}(s))|ds \\ &+ \int_0^t B_b(t-s)|\mathfrak{F}_2(x(s)) - \mathfrak{F}_2(\bar{x}(s))|ds.\end{aligned} \tag{C.48}$$

By the Lipschitz continuity of $\mathfrak{F}_1$ and $\mathfrak{F}_2$ and BL as defined before,

$$|x(t) - \bar{x}(t)| + |z(t) - \bar{z}(t)| \leq BL\int_0^t \big(|z(s) - \bar{z}(s)| + |x(s) - \bar{x}(s)|\big)ds. \tag{C.49}$$

By Grownwall's Lemma, it follows that,

$$|x(t) - \bar{x}(t)| + |z(t) - \bar{z}(t)| \leq 0, \tag{C.50}$$

Hence $x(t) \equiv \bar{x}(t)$ and $z(t) \equiv \bar{z}(t)$ on Ω_T. The existence of a unique solution for $y(t)$ follows from the existence of a unique solution for $z(t)$.

C.4 Local Asymptotic Stability of Disease-Free Equilibrium

The vector-host model (3.3)-(3.4) is linearized about the equilibrium,

$$(S_v^+, \bar{i}_v(\tau), S_b^+, \bar{i}_b(\tau), R_b^+, S_h^+, I_h^+, R_h^+),$$

by taking:

$$\begin{aligned}
S_v(t) &= S_v^+ + x_v(t), && \text{(C.51.1)}\\
i_v(\tau,t) &= \bar{i_v}(\tau) + y_v(\tau,t), && \text{(C.51.2)}\\
S_b(t) &= S_b^+ + x_b(t), && \text{(C.51.3)}\\
i_b(\tau,t) &= \bar{i}_b(\tau) + y_b(\tau,t), && \text{(C.51.4)}\\
R_b(t) &= R_b^+ + z_b(t), && \text{(C.51.5)}\\
S_h(t) &= S_h^+ + x_h(t), && \text{(C.51.6)}\\
I_h(t) &= I_h^+ + y_h(t), && \text{(C.51.7)}\\
R_h(t) &= R_h^+ + z_h(t). && \text{(C.51.8)}
\end{aligned}$$

To look for eigenvalues of the linear operator, we look for solutions of the form:

$$\begin{aligned}
x_v(t) &= \bar{x}_v e^{\lambda t}, && \text{(C.52.1)}\\
i_v(\tau,t) &= \bar{y}_v(\tau) e^{\lambda t}, && \text{(C.52.2)}\\
x_b(t) &= \bar{x}_b e^{\lambda t}, && \text{(C.52.3)}\\
i_b(\tau,t) &= \bar{y}_b(\tau) e^{\lambda t}, && \text{(C.52.4)}\\
z_b(t) &= \bar{z}_b e^{\lambda t}, && \text{(C.52.5)}\\
x_h(t) &= \bar{x}_h e^{\lambda t}, && \text{(C.52.6)}\\
y_h(t) &= \bar{y}_h e^{\lambda t}, && \text{(C.52.7)}\\
z_h(t) &= \bar{z}_h e^{\lambda t}. && \text{(C.52.8)}
\end{aligned}$$

Here, $\bar{x}_v, \bar{x}_b, \bar{z}_b, \bar{x}_h, \bar{y}_h, \bar{z}_h$ are arbitrary non-zero constants, and $\bar{y}_v(\tau), \bar{y}_b(\tau)$ are arbitrary functions of τ that are not identically zero. This results in the following system (the bars have been omitted):

$$\lambda x_v = -\frac{\alpha}{N_b^+ + N_h^+}\left(S_v^+ \int_0^\infty \beta_{bv}(\tau) y_b(\tau) d\tau + x_v \int_0^\infty \beta_{bv}(\tau)\bar{i}_b(\tau) d\tau\right) - \mu_v x_v, \tag{C.53.1}$$

$$\lambda y_v(\tau) + \frac{dy_v(\tau)}{d\tau} = -\mu_v y_v(\tau), \tag{C.53.2}$$

$$\lambda x_b = -\frac{\alpha}{N_b^+ + N_h^+}\left(S_b^+ \int_0^\infty \beta_{vb}(\tau) y_v(\tau) d\tau + x_b \int_0^\infty \beta_{vb}(\tau)\bar{i}_v(\tau) d\tau\right), \tag{C.53.3}$$

$$\lambda y_b(\tau) + \frac{dy_b(\tau)}{d\tau} = -\gamma_b(\tau) y_b(\tau), \tag{C.53.4}$$

$$\lambda z_b = \int_0^\infty \gamma_b(\tau) y_b(\tau) d\tau, \tag{C.53.5}$$

$$\lambda x_h = -\frac{\alpha}{N_b^+ + N_h^+}\left(S_h^+ \int_0^\infty \beta_{vh}(\tau) y_v(\tau) d\tau + x_h \int_0^\infty \beta_{vh}(\tau) \bar{i}_v(\tau) d\tau\right), \tag{C.53.6}$$

$$\lambda y_h = \frac{\alpha}{N_b^+ + N_h^+}\left(S_h^+ \int_0^\infty \beta_{vh}(\tau) y_v(\tau) d\tau + x_h \int_0^\infty \beta_{vh}(\tau) \bar{i}_v(\tau) d\tau\right) - \gamma_h y_h, \tag{C.53.7}$$

$$\lambda z_h = \gamma_h y_h, \tag{C.53.8}$$

with initial conditions:

$$y_v(0) = \frac{\alpha}{N_b^+ + N_h^+}\left(S_v^+ \int_0^\infty \beta_{bv}(\tau) y_b(\tau) d\tau + x_v \int_0^\infty \beta_{bv}(\tau) \bar{i}_b(\tau) d\tau\right), \tag{C.53.9}$$

$$y_b(0) = \frac{\alpha}{N_b^+ + N_h^+}\left(S_b^+ \int_0^\infty \beta_{vb}(\tau) y_v(\tau) d\tau + x_b \int_0^\infty \beta_{vb}(\tau) \bar{i}_v(\tau) d\tau\right). \tag{C.53.10}$$

Here, $N_b^+ = S_b^+ + \int_0^\infty \bar{i}_b(\tau) d\tau + R_b^+$ and $N_h^+ = S_h^+ + I_h^+ + R_h^+$ are the steady-state total populations of birds and humans, respectively. Solutions of the system (C.53) give the eigenvalues (λ) and eigenvectors of the differential operator. Knowing the distribution of the eigenvalues is sufficient to determine the stability of a given equilibrium for PDE operators – if all eigenvalues have negative real parts, then the corresponding equilibrium is locally asymptotically stable [33, 29]. On the other hand, if there is an eigenvalue with a positive real part, then the equilibrium is unstable. The equilibrium of interest is the disease-free equilibrium (DFE), $\mathcal{E}_0 = (S_v^0, 0, S_b^0, 0, 0, S_h^0, 0, 0)$, where $S_v^0 = \frac{\Lambda_v}{\mu_v}$. Note that at $\mathcal{E}_0$, $N_b = S_b^0$ and $N_h = S_h^0$. Under the DFE, the system (C.53) simplifies to:

$$\lambda x_v = -\frac{\alpha}{S_b^0 + S_h^0}\left(S_v^0 \int_0^\infty \beta_{bv}(\tau) y_b(\tau) d\tau\right) - \mu_v x_v, \tag{C.54.1}$$

$$\lambda y_v(\tau) + \frac{dy_v(\tau)}{d\tau} = -\mu_v y_v(\tau), \tag{C.54.2}$$

$$\lambda x_b = -\frac{\alpha}{S_b^0 + S_h^0}\left(S_b^0 \int_0^\infty \beta_{vb}(\tau) y_v(\tau) d\tau\right), \tag{C.54.3}$$

$$\lambda y_b(\tau) + \frac{dy_b(\tau)}{d\tau} = -\gamma_b(\tau) y_b(\tau), \tag{C.54.4}$$

$$\lambda z_b = \int_0^\infty \gamma_b(\tau) y_b(\tau) d\tau, \tag{C.54.5}$$

$$\lambda x_h = -\frac{\alpha}{S_b^0 + S_h^0}\left(S_h^0 \int_0^\infty \beta_{vh}(\tau) y_v(\tau) d\tau\right), \tag{C.54.6}$$

$$\lambda y_h = \frac{\alpha}{S_b^0 + S_h^0}\left(S_h^0 \int_0^\infty \beta_{vh}(\tau) y_v(\tau) d\tau\right) - \gamma_h y_h, \tag{C.54.7}$$

$$\lambda z_h = \gamma_h y_h, \tag{C.54.8}$$

with initial conditions:

$$y_v(0) = \frac{\alpha}{S_b^0 + S_h^0}\left(S_v^0 \int_0^\infty \beta_{bv}(\tau) y_b(\tau) d\tau\right), \tag{C.54.9}$$

$$y_b(0) = \frac{\alpha}{S_b^0 + S_h^0}\left(S_b^0 \int_0^\infty \beta_{vb}(\tau) y_v(\tau) d\tau\right). \tag{C.54.10}$$

Using variation of parameters to solve the differential equations of (C.54) gives:

$$y_v(\tau) = y_v(0) e^{-\lambda\tau} \Pi_v(\tau), \tag{C.55.1}$$

$$y_b(\tau) = y_b(0) e^{-\lambda\tau} \Pi_b(\tau). \tag{C.55.2}$$

Substituting (C.55.2) into (C.54.10) gives:

$$y_v(0) = \frac{\alpha}{S_b^0 + S_h^0}\left(S_v^0 \int_0^\infty \beta_{bv}(\tau) y_b(0) e^{-\lambda\tau} \Pi_b(\tau) d\tau\right). \tag{C.55.3}$$

Further substituting in (C.54.10) yields:

$$y_v(0) = \left(\frac{\alpha}{S_b^0 + S_h^0}\right)^2 S_v^0 S_b^0 \left(\int_0^\infty \beta_{vb}(\tau) y_v(\tau) d\tau\right)\left(\int_0^\infty \beta_{bv}(\tau) e^{-\lambda\tau} \Pi_b(\tau) d\tau\right). \tag{C.55.4}$$

By substituting in (C.55.1), dividing both sizes by $y_v(0)$, and substituting $S_v^0 = \frac{\Lambda_v}{\mu_v}$, we get:

$$1 = \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}\left(\int_0^\infty \beta_{vb}(\tau) e^{-\lambda\tau} \Pi_v(\tau) d\tau\right)\left(\int_0^\infty \beta_{bv}(\tau) e^{-\lambda\tau} \Pi_b(\tau) d\tau\right). \tag{C.55.5}$$

Equation (C.55.5) is a transcendental equation, and can therefore have many solutions. It is enough to show that all solutions λ of (C.55.5) have negative real parts to show the stability of $mathcalE_0$. We denote,

$$G(\lambda) = \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}\left(\int_0^\infty \beta_{vb}(\tau) e^{-\lambda\tau} \Pi_v(\tau) d\tau\right)\left(\int_0^\infty \beta_{bv}(\tau) e^{-\lambda\tau} \Pi_b(\tau) d\tau\right). \tag{C.56}$$

Notice that $G(0) = \mathcal{R}_0$. Suppose that $\mathcal{R}_0 > 1$. Then by (H2), $G(\lambda)$ is a decreasing function of λ. Since $G(0) > 1$ and $\lim_{\lambda \Longrightarrow \infty} G(\lambda) = 0$, then there exists a $\lambda^+ > 0$ such that $G(\lambda^+) = 1$. Therefore, $\mathcal{E}_0$ is unstable when $\mathcal{R}_0 > 1$.

Suppose now that $\mathcal{R}_0 < 1$. Then for all $\lambda = a + ib$ with $a < 0$,

$$\begin{aligned}|G(\lambda)| &\leq \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}\left(\int_0^\infty \beta_{vb}(\tau) \left|e^{-\lambda\tau}\right| \Pi_v(\tau) d\tau\right)\left(\int_0^\infty \beta_{bv}(\tau) \left|e^{-\lambda\tau}\right| \Pi_b(\tau) d\tau\right) \\ &\leq \frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}\left(\int_0^\infty \beta_{vb}(\tau) e^{-a\tau} \Pi_v(\tau) d\tau\right)\left(\int_0^\infty \beta_{bv}(\tau) e^{-a\tau} \Pi_b(\tau) d\tau\right) \leq \mathcal{R}_0 < 1.\end{aligned}$$

The equation $G(\lambda) = 1$ cannot be satisfied for λ's with $a < 0$. Therefore, $\mathcal{E}_0$ is locally asymptotically stable when $\mathcal{R}_0 < 1$.

C.5 Forward Normalized Sensitivity Indices

Recall that the basic reproduction number $\mathcal{R}_0$ can be written as:

$$\mathcal{R}_0 = \underbrace{\frac{\alpha^2 \Lambda_v S_b^0}{\mu_v (S_b^0 + S_h^0)^2}}_{\text{A}} \underbrace{\int_0^\infty \beta_{vb}(\tau)\Pi_v(\tau)d\tau}_{\text{B}} \underbrace{\int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau)d\tau}_{\text{C}},$$

and that the partial derivatives of $\mathcal{R}_0$ with respect to parameter p are found from Equation (3.12). The forward normalized sensitivity indices $\gamma_p^{\mathcal{R}_0}$ for the parameters of $\mathcal{R}_0$ are as follows:

$$\gamma_{M_1}^{\mathcal{R}_0} = 1, \tag{C.57.1}$$

$$\gamma_{k_1}^{\mathcal{R}_0} = \frac{k_1}{M_1 B} \int_0^\infty (\tau - \tau_{0,1}) e^{-k_1(\tau-\tau_{0,1})} \beta_{vb}^2(\tau)\Pi_v(\tau)d\tau, \tag{C.57.2}$$

$$\gamma_{\tau_{0,1}}^{\mathcal{R}_0} = \frac{-k_1 \tau_{0,1}}{M_1 B} \int_0^\infty e^{-k_1(\tau-\tau_{0,1})} \beta_{vb}^2(\tau)\Pi_v(\tau)d\tau, \tag{C.57.3}$$

$$\gamma_{a}^{\mathcal{R}_0} = 1, \tag{C.57.4}$$

$$\gamma_{b}^{\mathcal{R}_0} = \frac{b}{c^2 C} \int_0^\infty (\tau - b)\beta_{bv}(\tau)\Pi_b(\tau)d\tau, \tag{C.57.5}$$

$$\gamma_{c}^{\mathcal{R}_0} = \frac{1}{c^2 C} \int_0^\infty (\tau - b)^2\beta_{bv}(\tau)\Pi_b(\tau)d\tau, \tag{C.57.6}$$

$$\gamma_{M_2}^{\mathcal{R}_0} = \frac{-1}{C} \int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau) \int_0^\tau \gamma_b(s) ds d\tau, \tag{C.57.7}$$

$$\gamma_{k_2}^{\mathcal{R}_0} = \frac{-k_2}{M_2 C} \int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau) \int_0^\tau (s - \tau_{0,2}) e^{-k_2(s-\tau_{0,2})} \gamma_b^2(s) ds d\tau, \tag{C.57.8}$$

$$\gamma_{\tau_{0,2}}^{\mathcal{R}_0} = \frac{k_2 \tau_{0,2}}{M_2 C} \int_0^\infty \beta_{bv}(\tau)\Pi_b(\tau) \int_0^\tau e^{-k_2(s-\tau_{0,2})} \gamma_b^2(s) ds d\tau, \tag{C.57.9}$$

$$\gamma_{\Lambda_v}^{\mathcal{R}_0} = 1, \tag{C.57.10}$$

$$\gamma_{\alpha}^{\mathcal{R}_0} = 2, \tag{C.57.11}$$

$$\gamma_{\mu_v}^{\mathcal{R}_0} = -\left(\frac{\mu_v}{B} \int_0^\infty \beta v b(\tau)\Pi_v(\tau)\tau d\tau + 1 \right). \tag{C.57.12}$$

APPENDIX D

FULL WNV MODEL

D.1 Parameter Values and Sources

Table D.1: Values, Equations, and Sources for Parameters of the Full Model (4.5).

Parameter	**Value/Equation**	**Dimension**	**Source**
$r_v(t)$	Eq. (4.4.1)	days^{-1}	[49]
$K_v(t)$	Eq. (4.4.2)	vectors	[49]
$r_b(t)$	Eq. (4.4.3)	days^{-1}	Assumed from [19]
Λ_h	1/(365×75)	days^{-1}	[10]; [44]
α_b	1	bites/(vector×day)	[1]; [8]
α_h	Varies	bites/(vector×day)	Fitted from [61]
η	Varies	—	Assumed
$\beta_{vb}(\tau, T)$	Eq. (4.2)	bites^{-1}	Assumed from [63]; [43]
$\beta_{vh}(\tau, T)$	Eq. (4.2)	bites^{-1}	Assumed from [63]; [43]
$\beta_{bv}(\tau, t)$	Eq. (3.11.2)	bites^{-1}	Assumed from [39]
$\gamma_b(\tau)$	Eq. (3.11.3)	days^{-1}	Assumed from [39]
γ_h	1/14	days^{-1}	[44]
$\mu_v(T)$	Eq. (4.1.1)	days^{-1}	[16]
μ_b	1/(365×4)	days^{-1}	[44]
μ_h	1/(365×75)	days^{-1}	[10]; [44]
δ_b	0.001	days^{-1}	[88]
δ_h	1.00×10^{-5}	days^{-1}	[8]

Table D.2: Values, Equations, and Sources for Parameters of Infection-Age-Dependent Processes.

Parameter	**Value/Equation**	**Dimension**	**Source**
$M_1(T)$	Eq. (4.1.3)	—	[43]
k_1	0.495	days^{-1}	Fitted from [63]
$\tau_{0,1}(T)$	Eq. (4.1.2)	days^{-1}	[63]
a	1	—	Fitted from [63] [39]
b	3.06	days	Fitted from [63]; [39]
c	1.41	days	Fitted from [63]; [39]
M_2	0.5	days^{-1}	[39]
k_2	0.5	days^{-1}	Assumed from [39]
$\tau_{0,2}$	4.5	days	Assumed from [39]

D.2 Parameter Fittings for Bird Net Growth Rate

Table D.3: Fitted Parameter Values of Equation (4.4.3) for the WNV-Competent Bird Relative Abundance in the Greater Toronto Area for Years 2010–2017. For Each Year, the Bird Ecological Season Lasts from the Indicated Start Date Until the Start Date of the Following Year.

Year	**Start Date**	ϕ	ϕ_s	ω	$N_b(0)$	**RMSE**
2010	12/16/2009	-2.56×10^{-5}	-5.45×10^{-3}	0.01	0.149	0.016
2011	12/1/2010	1.20×10^{-3}	-9.68×10^{-3}	0.01	0.145	0.014
2012	12/4/2011	3.39×10^{-4}	-5.31×10^{-3}	0.01	0.202	2.52×10^{-3}
2013	12/31/2012	-8.30×10^{-4}	-5.34×10^{-3}	8.61×10^{-3}	0.175	0.011
2014	12/31/2013	2.77×10^{-4}	-4.77×10^{-3}	0.01	0.153	6.62×10^{-3}
2015	12/17/2014	4.90×10^{-4}	-5.18×10^{-3}	0.01	0.152	0.013
2016	12/1/2015	4.39×10^{-4}	-6.02×10^{-3}	0.01	0.171	0.015
2017	12/1/2016	3.03×10^{-6}	-5.29×10^{-3}	0.01	0.208	0.014

D.3 Numerical Implementation

Numerical implementation of the model is delicate due to the time-varying nature of both mosquito and bird populations. The sign of the net growth rate for mosquitoes $r_v(t)$ and birds $r_b(t)$ determines how the populations should be stratefied across the infection classes. When $r_v(t) > 0$ ($r_b(t) > 0$), then more mosquitoes (birds) are being recruited to the population than are dying. Note that since $r_v(t)$ and $r_b(t)$ are *net* growth rates, they represent the difference in individuals entering and leaving the population. Actual recruitment rates for mosquitoes $\psi_v(t)$ and birds $\psi_b(t)$ may be back-calculated from the net growth rates using the natural mortality rates:

$$r_v(t) = \psi_v(t) - \mu_v(T), \tag{D.1.1}$$

$$r_b(t) = \psi_b(t) - \mu_b, \tag{D.1.2}$$

where T is the temperature at time t. We assume the disease-induced mortality rate for WNV infected birds is small enough to neglect when back-calculating the true bird recruitment rate.

We assume all new individuals are recruited to their respective susceptible class and that individuals only become infected through the biting transmission process. Thus, when $r_v(t) > 0$, vector equations follow:

$$\frac{dS_v}{dt} = r_v(t)N_v(t)\left(1 - \frac{N_v(t)}{K_v(t)}\right) - \lambda_{bv}(t)S_v(t), \tag{D.2.1}$$

$$\frac{\partial i_v}{\partial t} + \frac{\partial i_v}{\partial \tau} = -\mu_v(T)i_v(\tau, t), \tag{D.2.2}$$

with infection-age boundary condition $i_v(0,t) = \lambda_{bv}(0)S_v(0)$. When $r_b(t) > 0$, bird equations follow:

$$\frac{dS_b}{dt} = r_b(t)N_b(t) - \lambda_{vb}(t)S_b(t), \tag{D.3.1}$$

$$\frac{\partial i_b}{\partial t} + \frac{\partial i_b}{\partial \tau} = -(\gamma_b(\tau) + \delta_b + \mu_b)i_b(\tau, t). \tag{D.3.2}$$

with infection-age boundary condition $i_b(0,t) = \lambda_{vb}(0)S_b(0)$.

However, when $r_v(t) \leq 0$ ($r_b(t) \leq 0$), then individuals are dying at the same rate or greater than those recruited. We must ensure that the appropriate number of individuals are removed from each infection class to avoid biologically unreasonable (negative population) simulations. For example, under equations (D.2) with $r_v(t) \leq 0$, then $N_v(t)$ is numerically less than $S_v(t) + I_v(t)$ and causes numerical instability. Therefore, when $r_v(t) \leq 0$, vector equations follow:

$$\frac{dS_v}{dt} = r_v(t)S_v(t)\left(1 - \frac{N_v(t)}{K_v(t)}\right) - \lambda_{bv}(t)S_v(t), \tag{D.4.1}$$

$$\frac{\partial i_v}{\partial t} + \frac{\partial i_v}{\partial \tau} = r_v(t)i_v(\tau, t)\left(1 - \frac{N_v(t)}{K_v(t)}\right), \tag{D.4.2}$$

and when $r_b(t) \leq 0$, bird equations follow:

$$\frac{dS_b}{dt} = r_b(t)S_b(t) - \lambda_{vb}(t)S_b(t), \tag{D.5.1}$$

$$\frac{\partial i_b}{\partial t} + \frac{\partial i_b}{\partial \tau} = r_b(t)i_b(\tau, t). \tag{D.5.2}$$

The infection-age boundary conditions for equations (D.4) and (D.5) are the same as the ones for (D.2) and (D.3), respectively. Equations (D.4) and (D.5) are not valid for the case where $r_v(t) > 0$ or $r_b(t) > 0$ since this would add individuals to the infected compartments that would otherwise come from the biting process. Further, notice that for both cases $r_v(t) > 0$ and $r_v(t) \leq 0$, equations sum up to the equation for the total mosquito population (4.3.1) (similar for bird population). We can therefore ensure that the populations stratified across infection classes add up to the total population values that align with the PBM and GPR data.