To the Graduate Council:

I am submitting herewith a dissertation written by Rachael Lynn Miller Neilan entitled "Optimal Control Applied to Population and Disease Models". I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Suzanne Lenhart, Major Professor

We have read this dissertation
and recommend its acceptance:

Louis Gross

Charles Collins

Agricola Odoi

Accepted for the Council:

Caroline R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Optimal Control Applied to Population and Disease Models

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Rachael Lynn Miller Neilan

August 2009

Copyright © 2009 by Rachael Lynn Miller Neilan.

All rights reserved.

Dedication

I dedicate this dissertation and completion of the doctorate program to my parents, Larry and Susie Miller. I will never be able to thank them enough for their unending support, encouragement, and advice. They have believed in me, my mathematical abilities, and my drive to succeed since I can remember. I love them tremendously.

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible. I am especially grateful to my advisor, Dr. Suzanne Lenhart for her encouragement from the beginning. She provides inspiration, sound advice, and many good ideas. I would have not accomplished this dissertation without her.

I would like to thank the other members of my committee, Dr.Lou Gross, Dr. Charles Collins, and Dr. Agricola Odoi for their time and critical review of my work. I thank them all for not only being on my thesis committee, but for their time and efforts teaching me in the classroom. They all contributed to my graduate studies. I would further like to thank Dr. Lou Gross for the invaluable teaching opportunities granted to me during my graduate studies.

The financial support during my graduate study came from a teaching assistantship which I held in the Department of Mathematics at the University of Tennessee. I am most grateful to this department and the staff who make it so enjoyable.

Lastly, and most importantly, I would like to thank my wonderful husband, Michael Neilan. Truly passionate about mathematics, his drive and determination are inspirational. I thank him for his mathematical advice and encouragement. Above all, I thank him for his unconditional love and support.

Abstract

This dissertation considers the use of optimal control theory in population models for the purpose of characterizing strategies of control which minimize an invasive or infected population with the least cost. Three different models and optimal control problems are presented. Each model describes population dynamics via a system of differential equations and includes the effects of one or more control methods.

The first model is a system of two ordinary differential equations describing dynamics between a native population and an invasive population. Population growth terms are functions of the control, constructed so that the value of the control may affect each population differently. A novel existence result is presented for the case of quadratic growth functions. With parameters chosen to mimic the competition between cottonwood and salt cedar plants, optimal schedules of controlled flooding are displayed.

The second model, a system of six ordinary differential equations, describes the spread of cholera in a human population through ingestion of *Vibrio cholerae*. Equations track movement of susceptible individuals to either an asymptomatic infected class or a symptomatic infected class through ingestion of bacteria, both in a hyperinfectious state and a less-infectious state. Recovered individuals temporarily move to an immune class before being placed back in the susceptible class. A new result quantifies contributions to the basic reproductive number from multiple infectious classes. Within the model, three control functions represent rehydration and antibiotic treatment, vaccination, and sanitation. The cost-effective balance of multiple cholera intervention methods is compared for two endemic populations.

The third model describes the spread of disease in both time and space using a system of

three parabolic partial differential equations with convection-diffusion movement terms and no-flux boundary conditions. A control function representing vaccination is incorporated. State variables track the number of susceptible, infected, and immune individuals. Detailed analysis for the characterization of the optimal control is provided. The model and optimal control results are applied to the spread of rabies among raccoons with the control function determining the timing and placement of oral vaccine baits. Results illustrate cost-effective vaccine distribution strategies for both regular and irregular patterns of rabies propagation.

Contents

1 Introduction 1

1.1 Mathematical Models of Invasion 1

1.2 Optimal Control for Ordinary Differential Equations 4

1.3 Optimal Control for Partial Differential Equations 7

1.4 Numerical Approximations to Solutions 8

2 Optimal Control Applied to Native-Invasive Population Dynamics . . . 10

2.1 Background 10

2.2 Model Formulation 12

2.3 Existence and Optimality System 15

2.4 Numerical Simulations 24

2.4.1 Quadratic Growth Functions 25

2.4.2 Quadratic Growth Functions: Alternative Scenario 28

2.4.3 Linear Growth Functions 28

2.5 Concluding Remarks and New Contributions 30

3 Optimal Intervention Strategies for a Cholera Outbreak 33

3.1 Background 33

3.2 Modeling the Spread of Cholera 36

3.2.1 A Brief History of Cholera Models 36

3.2.2 Model Formulation 40

3.2.3 The Basic Reproduction Number, $\mathcal{R}_0$ 43

3.2.4 Parameter Sensitivity Analysis 46

3.3 Optimal Control Problem with Multiple Controls 49

3.3.1 Introduction of Controls and Objective Functional 49
3.3.2 Existence and Optimality System 53
3.4 Population-Specific Strategies of Cholera Control 58
3.4.1 Parameter Values 59
3.4.2 Cholera Outbreaks in the Absence of Intervention 61
3.4.3 Optimal Strategies of Multiple Controls 63
3.4.4 Optimal Strategies of Single Control 74
3.4.5 Sensitivity Analysis of Control-Related Parameters 77
3.5 Concluding Remarks and New Contributions 80

4 Optimal Control of a Spatiotemporal Epidemic Model and Application to Rabies 81
4.1 Background 81
4.2 Model Formulation and Optimal Control Problem 83
4.3 Existence and Optimality System 86
4.3.1 Existence of State Solution and Optimal Control 86
4.3.2 Optimality System and Uniqueness 102
4.4 Application to Rabies and Raccoons 113
4.4.1 Spatial Homogeneity and Rabies Spread 117
4.4.2 Spatial Heterogeneity and Rabies Spread 120
4.4.3 Conclusions of Numerical Simulations 126
4.5 Concluding Remarks and New Contributions 128

5 Summary and Future Directions 129

Bibliography 131

Appendices 141

A Numerical Methods 142

B Previous Cholera Models 146

C Rabies Simulations 149

Vita 190

List of Tables

3.1 Summary of model notation and units 42

3.2 Sensitivity analysis parameters 47

3.3 Sensitivity analysis results 48

3.4 Additional notation for model with control 51

3.5 Population-specific parameter values 62

3.6 Fixed cost coefficients and control bounds 68

3.7 Sensitivity analysis parameters 78

3.8 Bogra-specific results of sensitivity analysis 78

3.9 Calcutta-specific results of sensitivity analysis 79

4.1 Parameters for homogeneous rabies spread and control 118

4.2 Parameters for heterogeneous rabies spread and control 123

B.1 Maximum likelihood estimates for King *et al.* two-path model 148

List of Figures

2.1 Quadratic case: Population densities without control 26

2.2 Quadratic case: Optimal control and corresponding population densities over timespan of three years. 27

2.3 Quadratic case: Optimal control and corresponding population densities over timespan of thirty years.l 27

2.4 Alternative quadratic case: Population densities without control 29

2.5 Alternative quadratic case: Optimal control and corresponding population densities 29

2.6 Linear case: Population densities without control 31

2.7 Linear case: Optimal control and corresponding population densities 31

3.1 Cholera model diagram 42

3.2 Cholera outbreaks for various values of β_H 60

3.3 Susceptible and recovered classes in absence of control 64

3.4 Asymptomatic and symptomatic infected classes in absence of control . . . 65

3.5 Hyperinfectious and less-infectious vibrio classes in absence of control . . . 66

3.6 Optimal balance of three controls 68

3.7 Susceptible and recovered classes with optimal balance of three controls . . 69

3.8 Asymptomatic and symptomatic infected classes with optimal balance of three controls 70

3.9 Hyperinfectious and less-infectious vibrio classes with optimal balance of three controls 71

3.10 Optimal strategies of one control alone 75

3.11 Infected classes with the optimal applications of one control alone and the optimal balance of all three controls 76

4.1 Homogeneous domain and site of initial infection 118

4.2 Heterogeneous domain 122

B.1 Diagram of Hartley *et al.* cholera model 146

B.2 Map of Bengal Bay districts from King *et al.* 147

B.3 Diagram of King *et al.* two-path cholera model 148

C.1 Spatial Homogeneity: Susceptible population over 52 weeks with no infection 150

C.2 Spatial Homogeneity: Susceptible population over 52 weeks in absence of vaccination 151

C.3 Spatial Homogeneity: Infected population over 52 weeks in absence of vaccination 152

C.4 Spatial Homogeneity: Optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 153

C.5 Spatial Homogeneity: Susceptible population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 154

C.6 Spatial Homogeneity: Infected population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 155

C.7 Spatial Homogeneity: Immune population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 156

C.8 Spatial Homogeneity: Optimal 15-week vaccination starting at $t = 27$, with $a = 0.02$ 157

C.9 Spatial Homogeneity: Susceptible population during optimal 15-week vaccination starting at $t = 27$, with $a = 0.02$ 158

C.10 Spatial Homogeneity: Infected population during optimal 15-week vaccination starting at $t = 27$, with $a = 0.02$ 159

C.11 Spatial Homogeneity: Immune population during optimal 15-week vaccination starting at $t = 27$, with $a = 0.02$ 160

C.12 Spatial Homogeneity: Optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 161
C.13 Spatial Homogeneity: Susceptible population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 162
C.14 Spatial Homogeneity: Infected population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 163
C.15 Spatial Homogeneity: Immune population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 164
C.16 Spatial Homogeneity: Optimal 20-week vaccination starting at $t = 21$, with $a = 0.03$ 165
C.17 Spatial Homogeneity: Susceptible population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.03$ 166
C.18 Spatial Homogeneity: Infected population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.03$ 167
C.19 Spatial Homogeneity: Immune population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.03$ 168
C.20 Spatial Heterogeneity: Susceptible population over 52 weeks with no infection 169
C.21 Spatial Heterogeneity: Susceptible population over 52 weeks, in absence of vaccination 170
C.22 Spatial Heterogeneity: Infected population over 52 weeks, in absence of vaccination 171
C.23 Spatial Heterogeneity: Optimal 10-week vaccination schedule starting at $t = 0$, with $a = 0.02$ 172
C.24 Spatial Heterogeneity: Susceptible population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 173
C.25 Spatial Heterogeneity: Infected population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 174
C.26 Spatial Heterogeneity: Immune population during optimal 10-week vaccination starting at $t = 0$, with $a = 0.02$ 175

C.27 Spatial Heterogeneity: Optimal 20-week vaccination schedule starting at $t = 21$, with $a = 0.01$ 176
C.28 Spatial Heterogeneity: Susceptible population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 177
C.29 Spatial Heterogeneity: Infected population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 178
C.30 Spatial Heterogeneity: Immune population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ 179
C.31 Spatial Heterogeneity and LDT: Susceptible population starting at $t = 21$ weeks, in absence of vaccination 180
C.32 Spatial Heterogeneity and LDT: Infected population starting at $t = 21$ weeks, in absence of vaccination 181
C.33 Spatial Heterogeneity and LDT: Optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.10$ 182
C.34 Spatial Heterogeneity and LDT: Susceptible population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.10$ 183
C.35 Spatial Heterogeneity and LDT: Infected population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.10$ 184
C.36 Spatial Heterogeneity and LDT: Immune population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.10$ 185
C.37 Spatial Heterogeneity and LDT: Optimal 20-week vaccination schedule starting at $t = 21$, with $a = 0.01$ and $c = 0.25$ 186
C.38 Spatial Heterogeneity and LDT: Susceptible population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.25$ 187
C.39 Spatial Heterogeneity and LDT: Infected population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.25$ 188
C.40 Spatial Heterogeneity and LDT: Immune population during optimal 20-week vaccination starting at $t = 21$, with $a = 0.01$ and $c = 0.25$ 189

Chapter 1

Introduction

1.1 Mathematical Models of Invasion

Mathematical models aid in understanding the dynamics and possible consequences of interacting populations. Models are particularly useful in developing strategies for the control of a harmful invasion. In the context of this dissertation, an invasion is perceived as an intruding population competing for resources with a native population or as an infectious disease permeating a susceptible population. Mathematical analysis, along with numerical simulations, can evaluate consequences of intervention strategies that would be unethical or cost-prohibitive to attempt under real world conditions.

To describe the interactions between population states, we use a system of either ordinary differential equations (ODEs) or partial differential equations (PDEs). Within the system, methods of intervention are modeled as control functions. The value of the control function affects the state differential equations and trajectories. We formulate an appropriate goal which depends on the control(s) and corresponding solutions of our system. An optimal control is one which achieves our goal.

In Chapter 2, the system of coupled ODEs

$$\frac{dN_1(t)}{dt} = (\Theta_1(t, u(t)) - a_{11}N_1(t) - a_{12}N_2(t))N_1(t),$$
$$\frac{dN_2(t)}{dt} = (\Theta_2(t, u(t)) - a_{21}N_1(t) - a_{22}N_2(t))N_2(t)$$

describes dynamics between a native population, N_1, and an invasive population, N_2. The control, u, appears within the growth function, Θ_i, of each species. A novel existence result allows us to choose the growth functions as quadratic functions of u. Parameter values for numerical simulations are chosen to represent the competition between the native cottonwood tree and the invasive salt cedar shrub in the American southwest. Here, the level of flooding permitted in nearby rivers is a control which affects each population's growth rate differently. The goal is to characterize the control schedule which minimizes the invasive population and the cost of the control while maximizing the native population at a final time.

In Chapter 3, the ODE system

$$\frac{dS}{dt} = -\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big]S(t) + \omega R(t),$$
$$\frac{dI_A}{dt} = p\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big]S(t) - (e_1 + \gamma_1)I_A(t)$$

$$\frac{dI_S}{dt} = (1-p)\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big]S(t) - (e_2 + \gamma_2)I_S(t)$$
$$\frac{dR}{dt} = \gamma_1 I_A(t) + \gamma_2 I_S(t) - \omega R(t),$$
$$\frac{dB_H}{dt} = \eta_1 I_A(t) + \eta_2 I_S(t) - \chi B_H(t),$$
$$\frac{dB_L}{dt} = \chi B_H(t) - \delta B_L(t)$$

is formulated to describe the spread of cholera within a human population. Population states include susceptible (S), asymptomatic infected (I_A), symptomatic infected (I_S), and recovered or immune (R) individuals. In addition, states (B_H and B_L) are added to track

the concentration of hyperinfectious and less-infectious bacteria within the environment. The environmental supply of bacteria is linked to the disease transmission terms within the human state equations. We extend the above model to include the effects of three controls (vaccination, sanitation, and rehydration therapy combined with antibiotics). The goal is to characterize the balance of the three controls which minimizes death due to disease and the cost of intervention over a finite time period. We also investigate optimal applications of one control.

In Chapter 4, we investigate a spatiotemporal model tracking the number of susceptible (S), infected (I), and immune (R) individuals within a population. To do this, we use the PDE system

$$\begin{aligned} L_1 S &= b(S+R) - \mu_1 S - \beta SI - avS, \\ L_2 I &= \beta SI - \mu_2 I, \\ L_3 R &= -\mu_1 R + avS \end{aligned}$$

where the operators L_k, $k = 1, 2, 3$ are defined as

$$L_k u \equiv \frac{\partial u}{\partial t} - \sum_{i,j=1}^{n} (a_{ij}^k u_{x_i})_{x_j} + \sum_{i=1}^{n} (b_i^k u)_{x_i}.$$

No-flux boundary conditions are also attached to the state system. The control v, considered a rate of vaccination, moves individuals from the S class to the R class. This system allows a more thorough investigation of disease dynamics by looking at both spatial and temporal spread. The goal is to characterize the timing and location of vaccination which minimizes the number of infected individuals and the cost of the control over a finite space and time domain. For numerical simulations, parameters are chosen according to the spread of rabies among a raccoon population. Here, vaccination is achieved through distribution of oral vaccine baits over a large area. We investigate optimal vaccination distribution for both regular and irregular patterns of disease propagation.

In designing the optimal control problem for all the scenarios described above, the

goal is formulated as minimizing an integral expression called the objective functional. To characterize the control(s) which minimize the objective functional, optimal control theory is used. The following is an introduction to optimal control theory and includes fundamental results regarding control of one state equation. Subsequential chapters provide detailed proofs for our state systems and optimal controls. Because numerical simulations will be used to approximate solutions and illustrate results, we also include a section in this introductory chapter summarizing the employed numerical methods.

1.2 Optimal Control for Ordinary Differential Equations

Although the models presented in this dissertation will be systems of two or more differential equations, we look first at optimal control of a single ODE. We denote $u(t)$ as the control and $x(t)$ as the state. The state function, $x(t)$, satisfies the ODE modeling the scenario. The control affects the state ODE by

$$x'(t) = g(t, x(t), u(t)).$$

Both $u(t)$ and $x(t)$ affect the goal, represented by the objective functional. The objective functional is typically an integral expression formulated terms of the state and control variables. We seek to find an optimal control and corresponding state that achieve the maximum (or minimum) of our objective functional.

The optimal control problem may be stated as

$$\max_{u \in U} \int_0^T f(t, x(t), u(t))dt \tag{1.1}$$

subject to

$$x'(t) = g(t, x(t), u(t)) \tag{1.2}$$

$$\text{where } x(0) = x_0 \text{ and } x(T) \text{ is free.} \tag{1.3}$$

We assume our control set U to be Lebesgue measurable functions.

An optimal control, denoted by $u^*(t)$, achieves the maximum. Assuming f and g are continuously differentiable in their arguments, one can state the first order necessary conditions in the simplest form by Pontryagin's Maximum Principle [66]. Optimal control theory for ODEs was developed by Pontryagin and his collaborators around 1950. They developed the key idea of introducing the adjoint function to attach the differential equation to the objective functional. For an introduction to optimal control theory, see [49].

Theorem 1. Pontryagin's Maximum Principle *If $u^*(t)$ and $x^*(t)$ are optimal for problem (1.1)-(1.3), then there exists adjoint variable $\lambda(t)$ such that*

$$H(t, x^*(t), u(t), \lambda(t)) \leq H(t, x^*(t), u^*(t), \lambda(t)),$$

at each time, where the Hamiltonian H is defined by

$$H(t, x(t), u(t), \lambda(t)) = f(t, x(t), u(t)) + \lambda(t) g(t, x(t), u(t))$$

and

$$\lambda'(t) = -\frac{\partial H(t, x(t), u(t), \lambda(t))}{\partial x}$$
$$\lambda(T) = 0.$$

The final time condition on the adjoint variable is called the transversality condition. This principle converted the problem of finding a control which maximizes the objective functional subject to the state ODE and initial condition to a the problem of optimizing the Hamiltonian pointwise. As stated above, the Hamiltonian is defined by

$$H(t,x,u,\lambda) = f(t,x,u) + \lambda g(t,x,u)$$
$$= (\text{integrand}) + (\text{adjoint}) \times (\text{RHS of ODE}).$$

One can generate the necessary conditions by maximizing H with respect to u at u^*. They are

$$\frac{\partial H}{\partial u} = 0 \Rightarrow f_u + \lambda g_u = 0 \text{ (optimality equation)},$$
$$\lambda' = -\frac{\partial H}{\partial x} \Rightarrow \lambda' = -\left(f_x + \lambda g_x\right) \text{ (adjoint equation), and}$$
$$\lambda(T) = 0 \text{ (transversality condition)}.$$

We can also consider second order conditions. For maximization, we have

$$\frac{\partial^2 H}{\partial u^2} \leq 0 \quad \text{at } u^*$$

and for minimization, we have

$$\frac{\partial^2 H}{\partial u^2} \geq 0 \quad \text{at } u^*.$$

We started with two unknowns, u^* and x^*, and we introduced an adjoint variable, λ. Now we have three unknowns, u^*, x^* and λ. Frequently, when the control enters into the problem in a nonlinear way, we can set $H_u|_{u=u^*} = 0$ and solve for u^*. This characterization of the optimal control will be in terms of x^* and λ. In many biological or physical problems, the controls will have bounds imposed due to the specific application. Pontryagin's Maximum Principle still holds when the controls are constrained within the bounds

$$a \leq u(t) \leq b.$$

We refer to the state and adjoint differential equations, along with control characterization, as the optimality system. Often, solutions of the optimality system cannot be solved explicitly, but can be approximated numerically. See Chapter 1.4 for a summary of the numerical methods used in this dissertation to solve optimality systems.

1.3 Optimal Control for Partial Differential Equations

No full generalization of Pontryagin's Maximum Principle has been made to infinite dimensional systems. However, in a system governed by PDEs, the ideas of Pontryagin can be used as an aid in characterizing an optimal control through an optimality system containing state and adjoint equations. The following is an outline of the solution technique as applied to the parabolic system in Chapter 4. See [51] for some specific results.

Let $u(x,t)$ denote the control function and $s(x,t)$ denote the state variable, both functions of space and time. Consider the optimal control problem of finding u^* such that

$$J(u^*) = \inf_{u \in U} J(u)$$

subject to

$$Ls = g(s,u)$$

where L is a known partial differential operator specific to the scenario and g is a continuous function of s and u. The set of admissible controls, U, is considered L^∞ functions of our space-time domain. The objective functional, specified here by J, is an integral expression containing a combination of both state and control variables.

To begin, the existence and uniqueness of the solution to the state equation are proven. For our system in Chapter 4, we construct a proof via a contraction mapping. The existence of an optimal control is then obtained through a minimizing sequence argument and continuous dependence of the states on the control. To obtain the necessary conditions for the optimal control, the objective functional J is differentiated with respect to the control u. Specifically, for u, $h \in U$, we take the Gateaux derivative of J with respect to u in the direction h,

$$\lim_{\epsilon \to 0} \frac{J(u + \epsilon h) - J(u)}{\epsilon}.$$

Since the objective functional usually contains the state variables, the states must first be differentiated with respect to the control. These directional derivatives are shown to converge to sensitivity functions. The sensitivities solve a linearized version of the state equations. At this point, *a priori* estimates are needed for the existence of the state system, the existence of an optimal control, and the convergence of the difference quotients to sensitivities.

Finally, the adjoint of the operator of the linearized equations is introduced with a final time condition and other appropriate boundary conditions. The components of the right-hand side of the adjoint system are the derivatives of the integrand of the objective functional with respect to each state.

Through standard optimality techniques, analyzing the objective functional and utilizing relationships between the state and adjoint equations, a characterization of the control is formulated. The optimality system consists of the state equations, adjoint equations, and control characterization. In the parabolic case, uniqueness for the optimality system, which characterizes the unique optimal control, holds only for sufficiently small time.

Next, we review the numerical scheme used to solve the optimality system.

1.4 Numerical Approximations to Solutions

It may not be possible to solve the optimality system analytically. Instead, we use numerical methods to approximate solutions and display results. To solve the optimality system with initial conditions for the states and final time conditions for the adjoints, an iterative scheme will be used. For systems of ODEs, we use a Runge Kutta method of the fourth order within the iterative scheme. For systems of PDEs, an explicit finite difference method accompanies the iterative scheme. We generalize the iterative scheme below, but refer the reader to Appendix A for details regarding the Runge Kutta and finite difference algorithms.

The iterative scheme can be generalized by the following steps.

1. Establish initial guess for control variable.

2. Given initial conditions for states, approximate solutions for state equations using either Runge Kutta or finite difference forward sweep method.

3. Given the state solutions from previous step and the final time conditions for adjoints, approximate solutions for adjoint equations using either Runge Kutta or finite difference backwards sweep method.

4. Update value of control variable by averaging the previous value and the new value arising from the control characterization.

5. Repeat steps 1-4 until successive values of all states, adjoints, and control(s) are sufficiently close.

The last step determining convergence requires the state, adjoint, and control values from two successive iterations to satisfy the relation

$$\frac{\|v - oldv\|}{\|v\|} \leq \epsilon$$

where ϵ is the accepted tolerance, v is the vector (or matrix) of current values, oldv is the vector (or matrix) from the previous iteration, and $\|\cdot\|$ refers to the sum of the absolute value of the elements within v.

Chapter 2

Optimal Control Applied to Native-Invasive Population Dynamics

2.1 Background

The problem of invasive, non-native species has become common in much of the world, in large part due to the introduction of new species by human beings, albeit sometimes accidentally. When developing successful plans for the eradication or control of invasives, the resulting dynamics must be considered [53, 59, 75]. Human alterations to a local environment can alter the interactions between two competing species, and should be taken into account when managing invasive species [64]. A natural disturbance in the environment, such as fire or flood, is often detrimental to businesses, homes and agriculture. Therefore such disturbance is often restricted or prevented. However, native species may depend on this natural disturbance to grow and propagate.

An illustrative example of this occurs in the American southwest. Native cottonwood trees (such as *Populus deltoides*) have to contend with a number of invasive species; among the most prevalent are *Tamarix ramosissima* and several other members of the genus

Tamarix, which are commonly known as tamarisks or salt cedars. Due to the low levels of rainfall in much of the region, cottonwood forests tend to dominate naturally only along rivers or other areas where there is surface water. Salt cedars, originally from central Asia and western China, were introduced into New Mexico in the 19th century, and to nearby regions at roughly the same time. Salt cedars were classified as an invasive species by the mid-twentieth century [89]. They are more tolerant of drought, salinity, and fire than cottonwoods, all of which can occur when overbank flooding is restricted or eliminated [73]. Currently, salt cedars occupy nearly every drainage system in arid and semi-arid areas west of the Great Plains, at minimum one million acres, and have been reported in about half of the eastern states of the United States [89].

For many years, the number of young cottonwood trees along rivers in the desert southwest has declined, with few if any young trees replacing the older ones. It has been suggested that salt cedars may competitively exclude cottonwood. However recent studies have indicated that densities of cottonwood are not affected by the presence of salt cedar unless the densities of cottonwood are very low [72, 73, 74, 80]. While in many cases disturbance of an ecosystem allows an invasive species to become established, flooding is critical to the continuance of cottonwood forests. Current flood controls on large Southwest rivers, such as the Rio Grande, have allowed for an opportunistic invasion of salt cedars, which are far less affected by the conditions. Some studies have suggested that when the environment is allowed to return to a more natural state (especially with regard to flood control), the two species will co-exist, or the cottonwood will become the dominant species once again (see Stromberg [80], and Sher *et al.* [72]).

Disturbances such as fire may be used to control an invasive species and to allow the native species to flourish [16, 39, 65]. Unlike flooding, however, the timing and extent of fire is not always as advantageous to the native species. In the case of frequent fires, an invasive species can take advantage of an open space that fire can bring, and push out native species. Prescribed burns may still be necessary to prevent wildfires, but consideration needs to be taken so that the ecosystem does not respond in the same way when invasives are present, and larger burns may now help invasive species substantially more than native ones [16]. In other circumstances, a prescribed burn at an appropriate time of the year

could be critical for preventing the spread of an invasive species, such as in Pollack and Kan's research on prescribed burns to control exotic annual grasses among native perennial grasses in California [65]. Also see the works by Beckage *et al.* [8] for the responses of pine savannas to fire disturbances, and by Bond *et al.* [13] documenting the importance of fire in promoting the spread of grasslands and savannas.

We present a system of two ordinary differential equations (ODEs) to represent our competition model with a control acting on the growth terms. A key point is that the control affects each species differently. We note that ODE models can represent such competition in plants or animals [24, 83, 85]. The primary goal here is to maximize the amount of the native species by applying the control action at appropriate times of year, while minimizing the "cost" of implementing the control. This "cost" may include the actual financial impact to carry out the management plan but also may consider the negative impact on the environment or economic development.

Motivated by the idea of "flooding," which may be detrimental at low and high levels and advantageous at medium levels, we consider growth functions which are quadratic in the controls. A new control analysis result presented here develops the existence of an optimal control in the case of quadratic functions of the control appearing in both ODEs, under a restriction on the coefficients of the quadratics.

The model and optimal control formulation are introduced in the Section 2.2. We present the existence of optimal control results in Section 2.3, along with the necessary conditions for solving this optimal control problem. The optimal control is characterized in terms of the adjoints and states. In the Section 2.4, we illustrate some numerical results for the cottonwood-salt cedar scenario and other situations and include some concluding remarks.

2.2 Model Formulation

The state variables in this model represent the population densities (seedlings per unit of area). The general model for the population dynamics is given by

$$\frac{dN_1(t)}{dt} = (\Theta_1(t, u(t)) - a_{11}N_1(t) - a_{12}N_2(t))N_1(t), \tag{2.1}$$

$$\frac{dN_2(t)}{dt} = (\Theta_2(t, u(t)) - a_{21}N_1(t) - a_{22}N_2(t))N_2(t), \tag{2.2}$$

$$N_1(0) = G_1, \qquad N_2(0) = G_2, \tag{2.3}$$

where $\Theta_i(t, u(t))$ are the intrinsic net growth rates (depending on the control $u(\cdot)$), and a_{ij} are the interaction coefficients indicating how species j affects species i. We are using the convention that the N_1 population is the (native) population that we want to grow by applying the control. Population N_2 is the competitor (perhaps invasive) population. The initial population values G_1 and G_2 are known positive constants. The interaction coefficients a_{11}, a_{12}, a_{21} and a_{22} are assumed to be positive.

The control variable $u(t)$ represents the amount of intervention at time t (in years). The control u has a pair of constraints. First, the control is scaled so that $0 \leq u \leq 1$. More notably, we only allow the control action to take place during certain times of the year. Connecting back with our motivating example, we would only allow flooding during the spring thaw when cottonwoods release their seeds. Starting the year in the spring, this is taken to occur during the first tenth of each year. Similarly, prescribed burns may be done in the late spring just before seed dispersal, or right after the start of seasonal rains. For instance, Pollack and Kan [65] found that after carefully timed burns, medusahead, an exotic grass present in California's Central Valley, was dramatically reduced or eliminated while native species returned to dominance in every site examined.

Thus we define our control set to be

$$U = \{u : [0, T] \to [0, 1] | u \text{ Lebesgue measurable}\},$$

and

$$\Theta_1(t,u) = (a_1u^2 + b_1u)I_\Omega(t) + c_1,$$
$$\Theta_2(t,u) = (a_2u^2 + b_2u)I_\Omega(t) + c_2.$$

Here, $I_\Omega(\cdot)$ is the characteristic function of the set Ω, which is given by $\Omega = \bigcup_{i=1}^{T}[\sigma_i, \tau_i]$, with T denoting the number of years of controls and interval $[\sigma_i, \tau_i]$ being in the i-th year. Then we see that on the set $[0,T] \setminus \Omega$, the control does not appear in the state equation, which amounts to saying that the control does not apply during those time intervals.

The growth rate parameters, a_1, b_1, c_1, a_2, b_2 and c_2, are chosen to fit a specific scenario. In the cottonwood-salt cedar case, ecological research influences the choice of growth rates. Tallent–Halsell and Walker [81] noted that both insufficient and excessive water levels can severely affect the establishment rate of salt cedars. In spite of this, salt cedars are better able to tolerate drought than cottonwoods; the population density still can increase without excessive flooding, though it more likely comes from vegetative growth rather than seedlings. Sher *et al.* [73, 74] found that flooding patterns similar to those present before the use of damming and other controls increased the establishment of cottonwood seedlings. Thus, the parameters are chosen so that with little or no flooding the salt cedar population, N_2, has a higher growth rate than the cottonwood population, N_1.

It should be noted that for other scenarios, the choice of growth rate parameters may be quite different. For instance, Brooks and Pyke [16] comment that regular, periodic fire in some ecosystems may allow for invasive species to become dominant, even in cases where fire is a natural occurrence and where native species are well-adapted to regular burns. In this circumstance, the growth parameters would instead be chosen so that the native population N_1 has a higher growth rate in the absence of fire, but not for extensive burns.

Our objective functional to be maximized is given by:

$$J(u(\cdot)) = AN_1(T) - KN_2(T) - \int_0^T [Bu(t) + \epsilon u^2(t)]I_\Omega(t)dt \tag{2.4}$$

where the native population is to be maximized at the final T subject to balancing the

minimization of the second population and the cost of implementing the control. Many times the cost of implementing a control would be nonlinear and as such a simple nonlinear case is taken here. As mentioned, the costs can include funds needed for control implementation as well as the negative financial impact on the environment. If one takes $K = 0$, then the emphasis is to promote the first population without regard to the second population.

2.3 Existence and Optimality System

The nonlinear state system with measurable coefficients as controls has a solution $N(\cdot) \equiv (N_1(\cdot), N_2(\cdot))$ taking values in $\mathbf{R}^2_+ \equiv \{(N_1, N_2) \in \mathbf{R}^2 \mid N_1, N_2 > 0\}$, due to a standard result in Lukes [54], and *a priori* L^∞ bounds hold for the states from the structure of the system.

Because the growth functions, Θ_1, Θ_2, are quadratic in the control with a coefficient restriction, we must do some analysis to get the existence of an optimal control. To begin with, for any $(t, N) \in [0, T] \times \mathbf{R}^2_+$, let us denote

$$f(t, N, u) = \begin{pmatrix} \Theta_1(t, u)N_1 \\ \Theta_2(t, u)N_2 \end{pmatrix}, \qquad g(N) = \begin{pmatrix} (a_{11}N_1 + a_{12}N_2)N_1 \\ (a_{21}N_1 + a_{22}N_2)N_2 \end{pmatrix},$$
$$f^0(t, u) = (\epsilon u^2 + Bu)I_\Omega(t), \qquad g^0(N) = AN_1 - KN_2.$$

Then the state equation reads

$$\frac{dN(t)}{dt} = f(t, N(t), u(t)) - g(N(t)),$$

and the payoff functional reads

$$J(u(\cdot)) = g^0(N(T)) - \int_0^T f^0(t, u(t))dt.$$

Our main result of this section is the following theorem.

Theorem 2. *Suppose there exists a* $\kappa \geq 0$ *such that*

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \kappa \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad B\kappa - \epsilon \leq 0, \tag{2.5}$$

then there exists an optimal control u^* *in* U *with corresponding states* N_1^*, N_2^* *that maximizes the objective functional* $J(u)$ *defined by (2.4).*

Proof: For any $(t, N) \in [0, T] \times \mathbf{R}_+^2$, define

$$\varepsilon(t, N) = \Big\{ (z^0, z) \in \mathbf{R}^{1+2} \;\Big|\; z^0 \geq f^0(t, u),\ z = f(t, N, u),\ \text{ for some u}\in [0, 1] \Big\}.$$

The proof consists of two steps. In *Step 1*, we will show that an optimal control exists if for almost all $t \in [0, T]$ and each N$\in \mathbf{R}_+^2$, the set $\varepsilon(t, N)$ is convex and closed. In *Step 2*, we will show that $\varepsilon(t, N)$ is convex and closed.

Step 1: Let $\{u^n(\cdot)\}_{n\geq 1}$ be a maximizing sequence and $N^n(\cdot)$ be the state trajectory corresponding to $u^n(\cdot)$. We can show that the sequence $\{N^n(\cdot)\}$ is uniformly bounded and equicontinuous. Thus, by the Ascoli-Arzela Theorem, we assume

$$N^n(\cdot) \to \bar{N}(\cdot) \text{ uniformly on [0,T] for some } \bar{N}(\cdot) \in \text{C([0,T];}\mathbf{R}_+^2).$$

By the definition of $\varepsilon(t, N)$,

$$(z_n^0(t), z_n(t)) \equiv (f^0(t, u^n(t)), f(t, N^n(t), u^n(t))) \in \varepsilon(t, N^n(t)) \quad \text{a.e. } t \in [0, T].$$

The uniform boundedness of $N^n(\cdot)$ leads to the uniform boundedness of $(z_n^0(\cdot), z_n(\cdot))$, thus implying the $L^2(0, T)$ weak convergence of (z_n^0, z_n) to some $(\bar{z}^0, \bar{z})$.

By Mazur's Theorem, there exists some function $\sigma : \mathbf{N} \to \mathbf{N}$ and a sequence of sets of real numbers $\{\alpha_{ij} | i = j, ..., \sigma(j)\}_{j=1}^{\infty}$ such that

$$\alpha_{ij} \geq 0, \quad \sum_{i=j}^{\sigma(j)} \alpha_{ij} = 1$$

and the sequence

$$\varphi_j(\cdot) \equiv \sum_{i=j}^{\sigma(j)} \alpha_{ij}(z_i^0(t), z_i(t)) \to (\bar{z}^0(t), \bar{z}(t)) \text{ strongly in } L^2(0,T).$$

Therefore, we may assume that on a subsequence $\varphi_j(\cdot) \to (\bar{z}^0(t), \bar{z}(t))$ pointwise almost everywhere on [0,T]. By Egorov's Theorem, we have for any $\epsilon > 0$, there exists an $E_\epsilon \subseteq [0,T]$, such that $|E_\epsilon| > T - \epsilon$ (where $|E_\epsilon|$ is the Lebesgue measure of E_ϵ) and

$$\varphi_j(\cdot) \to (\bar{z}^0(t), \bar{z}(t)) \quad \text{uniformly on } E_\epsilon \text{ as j}\to \infty,$$

Furthermore, because $N^n(\cdot) \to \bar{N}(\cdot)$ uniformly on E_ϵ and $\varepsilon(t,N)$ is convex, we have that for any $\delta > 0$

$$\varphi_j(t) \in N_\delta(\varepsilon(t, \bar{N}(t))) \quad \text{for } t \in E_\epsilon \text{ and j large,}$$

where $N_\delta(A)$ is the δ-neighborhood of set A.

The closedness of $\varepsilon(t,N)$ gives

$$(\bar{z^0}(t), \bar{z}(t)) \in \bigcap_{\delta>0} N_\delta(\varepsilon(t, \bar{N}(t))) = \varepsilon(t, \bar{N}(t)) \quad \text{for } t \in E_\epsilon.$$

Since $\epsilon > 0$ is arbitrary, we obtain

$$(\bar{z}^0(t), \bar{z}(t)) \in \varepsilon(t, \bar{N}(t)) \quad \text{a.e. } t \in[0,\text{T}].$$

This means that there exists $\bar{u}(\cdot)$ such that

$$\bar{z}^0(t) \geq f^0(t, \bar{u}(t)), \quad \bar{z}(t) = f(t, \bar{N}(t), \bar{u}(t)) \quad \text{a.e. } t \in[0,\text{T}].$$

By the definition of $\varepsilon(t,N)$ we obtain

$$\tfrac{d\bar{N}(t)}{dt} = \bar{z}(t) - g(\bar{N}(t)).$$

Thus $\bar{N}(\cdot)$ is the state trajectory corresponding to the control $\bar{u}(\cdot)$.

The known convergence of $N^n(T) \to \bar{N}(T)$ in $\mathbf{R}^2_+$ and $(z_n^0(t), z_n(t)) \to (\bar{z}^0, \bar{z})$ weakly in $L^2(0,T)$ gives

$$\begin{aligned}\sup_{u(\cdot)\in U} J(u(\cdot)) &= \lim_{n\to\infty} J(u^n(\cdot)) \\ &= \lim_{n\to\infty} g^0(N^n(T)) - \lim_{n\to\infty}\int_0^T z_n^0(t)dt \\ &= g^0(\bar{N}(T)) - \int_0^T \bar{z}^0(t)dt \\ &\le g^0(\bar{N}(T)) - \int_0^T f^0(t,\bar{u}(t))dt \\ &= J(\bar{u}(\cdot)).\end{aligned}$$

Thus, $\bar{u}(\cdot)$ is an optimal control.

Step 2: The closed property is obvious since $u \mapsto (f^0(t,u), f(t,N,u))$ is continuous. For the convexity, we note that (2.5) yields

$$f(t,N,u) = (\kappa u^2 + u)I_\Omega(t)\begin{pmatrix} b_1 N_1 \\ b_2 N_2 \end{pmatrix} + \begin{pmatrix} c_1 N_1 \\ c_2 N_2 \end{pmatrix}.$$

Clearly, when $t \in [0,T]\setminus\Omega$,

$$\varepsilon(t,N) = [0,\infty) \times \left\{ \begin{pmatrix} c_1 N_1 \\ c_2 N_2 \end{pmatrix} \right\}$$

is trivially convex. Hence we now consider $t \in \Omega$. Denote $\varphi(u) = \kappa u^2 + u$. Since $\kappa \ge 0$, $\varphi'(u) = 2\kappa u + 1 \ge 0$, for $u \in [0,1]$. Thus, $\varphi(\cdot)$ is convex, monotone increasing, and

$$\varphi([0,1]) = [\varphi(0), \varphi(1)] = [0, \kappa + 1], \tag{2.6}$$

which is convex. Consequently, $f(t,N,[0,1])$ is convex. Now, for any $(z^0,z),(\zeta^0,\zeta) \in \varepsilon(t,N)$ and $\lambda \in (0,1)$, we have some $u,v \in [0,1]$ such that

$$z = f(t, N, u), \quad z^0 \geq f^0(t, u),$$
$$\zeta = f(t, N, u), \quad \zeta^0 \geq f^0(t, u)$$

By the convexity of $\varphi(\cdot)$ and (2.6), we can find some $w \in [0, 1]$ such that

$$\varphi(\lambda u + (1-\lambda)v) \leq \lambda\varphi(u) + (1-\lambda)\varphi(v) = \varphi(w) \tag{2.7}$$

which leads to

$$\begin{aligned}
\lambda z + (1-\lambda)\zeta &\equiv \lambda f(t, N, u) + (1-\lambda) f(t, N, v) \\
&= [\lambda\varphi(u) + (1-\lambda)\varphi(v)] \begin{pmatrix} b_1 N_1 \\ b_2 N_2 \end{pmatrix} + \begin{pmatrix} c_1 N_1 \\ c_2 N_2 \end{pmatrix} \\
&= \varphi(w) \begin{pmatrix} b_1 N_1 \\ b_2 N_2 \end{pmatrix} + \begin{pmatrix} c_1 N_1 \\ c_2 N_2 \end{pmatrix} \\
&= f(t, N, w).
\end{aligned}$$

By the monotonicity of $u \mapsto \varphi(u)$, we have from the inequality in (2.7) that

$$\lambda u + (1-\lambda)v \leq w. \tag{2.8}$$

Further, from the equality in (2.7), we see that

$$\kappa[\lambda u^2 + (1-\lambda)v^2 - w^2] + \lambda u + (1-\lambda)v - w = 0. \tag{2.9}$$

Hence, in the case that $\kappa > 0$, (2.9), (2.5) and (2.8) imply

$$\begin{aligned}\lambda z^0 + (1-\lambda)\zeta^0 - f^0(t,w) &\geq \lambda f^0(t,u) + (1-\lambda) f^0(t,v) - f^0(t,w) \\ &= \epsilon[\lambda u^2 + (1-\lambda)v^2 - w^2] + B[\lambda u + (1-\lambda)v - w] \\ &= (B - \frac{\epsilon}{\kappa})[\lambda u + (1-\lambda)v - w] \\ &\geq 0.\end{aligned}$$

In the case that $\kappa = 0$, (2.9) becomes

$$w = \lambda u + (1-\lambda)v.$$

By the above equality and the convexity of $u \mapsto u^2$, we have

$$\begin{aligned}\lambda z^0 + (1-\lambda)\zeta^0 - f^0(t,w) &\geq \lambda f^0(t,u) + (1-\lambda) f^0(t,v) - f^0(t,w) \\ &= \epsilon(\lambda u^2 + (1-\lambda)v^2 - w^2) + B(\lambda u + (1-\lambda)v - w) \\ &= \epsilon(\lambda u^2 + (1-\lambda)v^2 - [\lambda u + (1-\lambda)v]^2) \\ &\geq 0.\end{aligned}$$

Hence, we obtain

$$\lambda(z^0, z) + (1-\lambda)(\zeta^0, \zeta) \in \varepsilon(t, N),$$

proving the convexity of $\varepsilon(t, N)$. □

We now use Pontryagin's Maximum Principle [66] to find the optimal control. Formulated from the cost functional (2.4) and the governing dynamics (2.1),(2.2), the Hamiltonian is

$$H = -Bu - \epsilon u^2 + \lambda_1(\Theta_1(u) - a_{11}N_1 - a_{12}N_2)N_1 + \lambda_2(\Theta_2(u) - a_{21}N_1 - a_{22}N_2)N_2 \tag{2.10}$$

where λ_1 and λ_2 are the associated adjoints for states N_1 and N_2 respectively. The adjoint variables will be needed to calculate the optimal control. The system of equations is found by taking the appropriate partial derivatives of the Hamiltonian (2.10) with respect to the associated state variable.

Theorem 3. *Suppose (2.5) holds with $B\kappa - \epsilon < 0$. Given an optimal control u^* and corresponding states N_1^*, N_2^*, there exist adjoint functions satisfying*

$$\lambda_1' = \lambda_1(-\Theta_1(u^*) + 2a_{11}N_1^* + a_{12}N_2^*) + \lambda_2 a_{21}N_2^* \tag{2.11}$$

$$\lambda_2' = \lambda_2(-\Theta_2(u^*) + 2a_{22}N_2^* + a_{21}N_1^*) + \lambda_1 a_{12}N_1^* \tag{2.12}$$

with transversality conditions

$$\lambda_1(T) = A, \;\; \lambda_2(T) = -K. \tag{2.13}$$

Furthermore, the optimal control is characterized by

$$u^*(t) = \min\left[1, \max(0, \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)})\right] \tag{2.14}$$

for $0 \le t \le T$ such that $\epsilon - a_1\lambda_1(t)N_1^*(t) - a_2\lambda_2(t)N_2^*(t) \neq 0$.

If there exists $0 \le \bar{t} \le T$ such that $\epsilon - a_1\lambda_1(\bar{t})N_1^*(\bar{t}) - a_2\lambda_2(\bar{t})N_2^*(\bar{t}) = 0$, then $u^*(\bar{t}) = 1$.

Proof: To calculate the adjoint differential equations (3.18) and (3.19), we use

$$\lambda_1' = -\frac{\partial H}{\partial N_1}$$
$$\lambda_2' = -\frac{\partial H}{\partial N_2}.$$

Final time values in (2.13) are derived from the transversality conditions

$$\frac{d(AN_1(T))}{dN_1(T)} = \lambda_1(T),$$
$$\frac{d(-KN_2(T))}{dN_2(T)} = \lambda_2(T).$$

On the set of all $0 \leq t \leq T$ such that $\epsilon - a_1\lambda_1(t)N_1^*(t) - a_2\lambda_2(t)N_2^*(t) \neq 0$, the Hamiltonian (2.10) is quadratic in the control, u. Because of the bounds on the control, we consider three cases in the characterization of the control.

1. On the set $\{t|0 < u^*(t) < 1\}$, we have

$$\begin{aligned} 0 &= \frac{\partial H}{\partial u} \mid_{u^*} \\ &= -B + -2\epsilon u^* + 2a_1u^*\lambda_1N_1^* + b_1\lambda_1N_1^* + 2a_2u^*\lambda_2N_2^* + b_2\lambda_2N_2^*. \end{aligned}$$

   Solving the above for u^* yields

$$u^*(t) = \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)}.$$

2. On the set $\{t|u^*(t) = 0\}$, we have

$$\begin{aligned} 0 &\geq \frac{\partial H}{\partial u} \mid_{u^*} \\ &= -B + b_1\lambda_1N_1^* + b_2\lambda_2N_2^*. \end{aligned} \tag{2.15}$$

   Multiplying the above inequality by $-\kappa$ and using relationships in (2.5) yields

$$\begin{aligned} 0 &\leq \kappa B - \kappa b_1\lambda_1N_1^* - \kappa b_2\lambda_2N_2^* \\ &\leq \epsilon - a_1\lambda_1N_1^* - a_2\lambda_2N_2^* \end{aligned} \tag{2.16}$$

The two inequalities (2.15),(2.16) force

$$\frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)} \leq 0$$

and therefore the characterization

$$u^*(t) = \max(0, \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)})$$

holds on this set.

3. On the set $\{t|u^*(t) = 1\}$, we have

$$\begin{aligned} 0 &\leq \frac{\partial H}{\partial u} \mid_{u^*} \\ &= -B + -2\epsilon + 2a_1\lambda_1 N_1^* + b_1\lambda_1 N_1^* + 2a_2\lambda_2 N_2^* + b_2\lambda_2 N_2^*. \end{aligned}$$

Rearranging the inequality gives

$$2\epsilon - 2a_1\lambda_1 N_1^* - 2a_2\lambda_2 N_2^* \leq -B + b_1\lambda_1 N_1^* + b_2\lambda_2 N_2^*. \tag{2.17}$$

Because $0 \geq \frac{\partial^2 H}{\partial u^2} \mid_{u^*} = -2\epsilon + 2a_1\lambda_1 N_1^* + 2a_2\lambda_2 N_2^*$ and because $2(\epsilon - a_1\lambda_1 N_1^* - a_2\lambda_2 N_2^*) \neq 0$ on this set, the left hand side of (2.17) is positive and therefore the right-hand side is also positive. Dividing both sides of (2.17) by $2\epsilon - 2a_1\lambda_1 N_1^* - 2a_2\lambda_2 N_2^*$ gives

$$1 \leq \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)}$$

and therefore the characterization

$$u^*(t) = \min(1, \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)})$$

holds on this set.

Considering these three cases we obtain the desired characterization,

$$u^*(t) = \min[1, \max(0, \frac{-B + b_1\lambda_1(t)N_1^*(t) + b_2\lambda_2(t)N_2^*(t)}{2\epsilon - 2a_1\lambda_1(t)N_1^*(t) - 2a_2\lambda_2(t)N_2^*(t)})].$$

If there exists $0 \leq \bar{t} \leq T$ such that

$$\epsilon - a_1\lambda_1(\bar{t})N_1^*(\bar{t}) - a_2\lambda_2(\bar{t})N_2^*(\bar{t}) = 0, \tag{2.18}$$

the Hamiltonian (2.10) becomes linear in the control. At $t = \bar{t}$ the coefficient of u in the Hamiltonian is

$$\Psi = -B + b_1\lambda_1(\bar{t})N_1(\bar{t}) + b_2\lambda_2(\bar{t})N_2(\bar{t}).$$

Due to (2.18) and the assumption that $B\kappa - \epsilon < 0$, we have

$$\begin{aligned} -\kappa\Psi &= \kappa B - \kappa b_1\lambda_1(\bar{t})N_1(\bar{t}) - \kappa b_2\lambda_2(\bar{t})N_2(\bar{t}) \\ &= \kappa B - a_1\lambda_1(\bar{t})N_1(\bar{t}) - a_2\lambda_2(\bar{t})N_2(\bar{t}) \\ &= \kappa B - \epsilon \\ &< 0. \end{aligned}$$

By definition, κ is non-negative. The strict inequality above implies $\kappa > 0$ and $\Psi > 0$. With $\Psi > 0$, we take $u^*(\bar{t}) = 1$. □

Thus the optimality system consists of the state system (2.1)-(2.3), together with the adjoint system (2.11),(2.12) and associated transversality boundary conditions (2.13), coupled with the characterization of the optimal control (2.14). We now examine the population changes for a variety of parameter values and growth functions, including our cottonwood-salt cedar scenario.

2.4 Numerical Simulations

Solutions to the optimality system are computed numerically for a variety of parameter values and growth functions (Θ_1, Θ_2), including those modeling the cottonwood-salt cedar

case. For the numerical simulations, we use an iterative scheme employing a Runge Kutta method of the fourth order. See Chapter 1 and Appendix A for details.

2.4.1 Quadratic Growth Functions

In the cottonwood-salt cedar scenario, quadratic growth functions can represent reasonable qualitative behavior for newly established seedling communities. The interaction coefficients, a_{ij}, express how species j affects species i. Sher *et al.* [74] found that during natural conditions, cottonwood density affects the population of both species (in a negative fashion) more than the salt cedar density. In fact, salt cedar density does not seem to affect the population size of cottonwoods except when the cottonwood density is much smaller than the salt cedar density. Thus we choose coefficients so that a_{12} is very small and $a_{21} > a_{22}$.

For the numerical illustration, the interaction parameter values are:

$$a_{11} = .01, a_{12} = .0001, a_{21} = .15, a_{22} = .01.$$

To model the cottonwood-salt cedar scenario, we choose $\Theta_1 < \Theta_2$ when $u = 0$ and we use our quadratic growth functions,

$$\begin{aligned} \Theta_1 &= .5u^2 + .5u \\ \Theta_2 &= -.1u^2 - .1u + .2. \end{aligned} \tag{2.19}$$

With the above interaction parameters and growth functions, if $u = 0$ (i.e. no flooding), we find that as $t \to \infty$, the native population N_1 becomes extinct and the invasive population N_2 approaches a positive carrying capacity. However, for mid values of u (i.e. natural flooding conditions), we find that as $t \to \infty$, the native population N_1 approaches a positive carrying capacity and the invasive population N_2 becomes extinct.

For appropriate initial conditions, we scale the population densities to match the ratio suggested by Sher *et al.* [74]. Because the mass difference between a newly germinated cottonwood and salt cedar is five-fold, all experimental densities in [74] include five salt

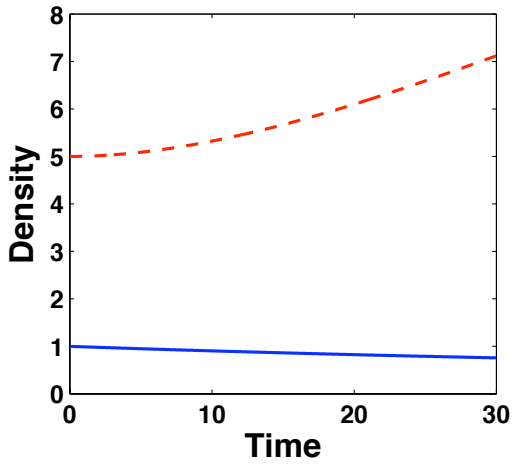


Figure 2.1: N_1 (solid) and N_2 (dotted) densities without control

cedars for every one cottonwood. This seedling density ratio is also comparable to that found in field situations [82]. Accordingly, we assume that the two seedling populations start with densities (seedlings per unit area) scaled so that $G_1 = 1, G_2 = 5$. When $u = 0$, the growth functions are 0 and .2 respectively. The cottonwood density decreases, while the salt cedar density increases without any flooding; see Figure 2.1.

Consider the case $A = 9, B = 2.5$, $\epsilon = 7$, and $K = 0$, in which we are concentrating on the cottonwood population and our goal does not directly involve the salt cedar population. To correspond with the spring thaw, flooding is only allowed during the first one-tenth of the year for a period of three years. In Figure 2.2, the optimal control does not get close to the maximum u value, due to the cost of implementing the control. If we decrease the cost coefficients, B and ϵ, the optimal control values are greater. The optimal control allows for a moderate increase in the cottonwood density.

Because of the lifespan of such trees, controlled flooding may need to be implemented for a much longer time period in order to achieve desired results. Consider the optimal flooding schedule for thirty years in Figure 2.3. Here, we again see increasing amounts of flooding until year 6, at which time we have reached maximum levels. The optimal strategy suggests continuing max flooding until year 18, at which time we discontinue flooding due to the associated costs. With 18 consecutive years of flooding in the spring, cottonwoods are able to dominate salt cedars. Taylor *et al.* [82] show that under such conditions, the cottonwoods

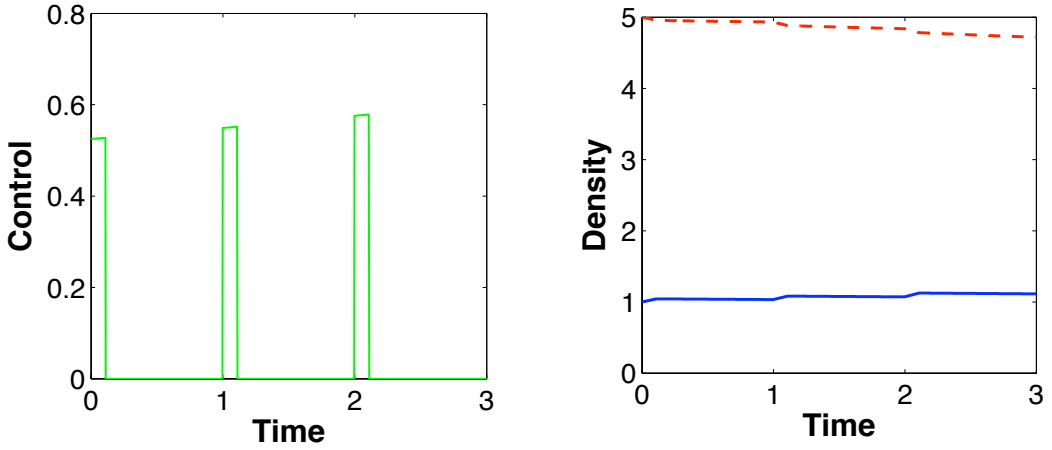


Figure 2.2: Left: Optimal Control Schedule for $T = 3$ years, Right: Corresponding N_1 (solid) and N_2 (dotted) densities

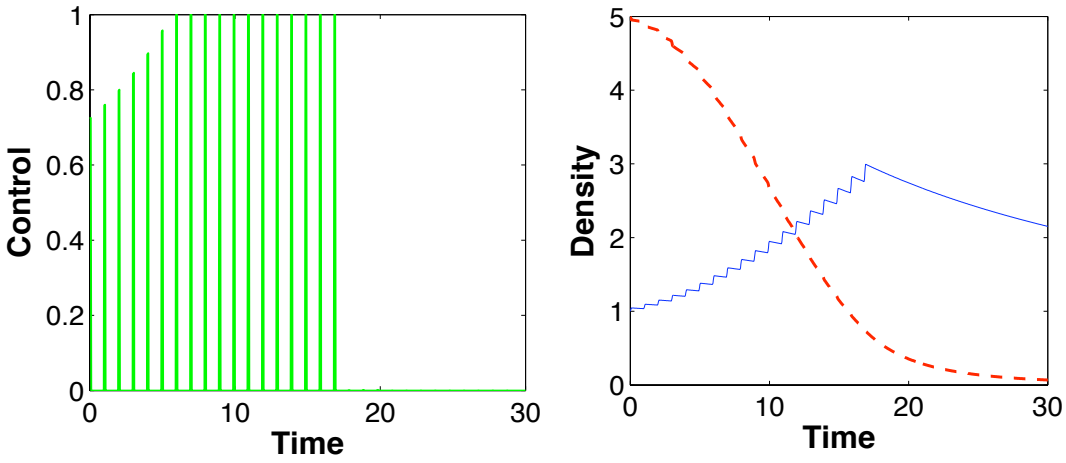


Figure 2.3: Left: Optimal Control Schedule for $T = 30$ years, Right: Corresponding N_1 (solid) and N_2 (dotted) densities

will continue to dominate beyond the sapling stage, thereby forming a dominant canopy species. A dominant cottonwood canopy should further limit the growth rate of the salt cedar.

2.4.2 Quadratic Growth Functions: Alternative Scenario

We now consider the case of two competing populations with equal initial densities. Again N_2 is considered a population which is typically dominated by the N_1 population, but because of the absence of some control, N_1 has diminished and N_2 has increased. In contrast to the cottonwood-salt cedar case, we let N_2 have a competitive advantage over N_1, meaning, $a_{12} > a_{21}$. Our interaction parameters are:

$$a_{11} = .01, a_{12} = .02, a_{21} = .01, a_{22} = .01.$$

We choose the same quadratic growth functions as that in Section 2.4.1. The initial population densities are scaled so that $G_1 = G_2 = 1$. With no control, the growth rates of N_1 and N_2 are 0 and .2 respectively. See Figure 2.4 for populations without control.

In Section 2.4.1, we allowed the control to be applied for one-tenth of a year. However, the control may have longer lasting effects. Suppose the control is applied during the first one-fourth of each year for three years. Let $A = 10, B = 0, K = 1, \epsilon = 12$. The numerical results are displayed in Figure 2.5. The presence of the control significantly increases the N_1 growth rate, and promotes a reasonable proportion between the two populations.

2.4.3 Linear Growth Functions

For some scenarios, linear growth functions may be a better representation. We consider a situation similar to that of the cottonwood (N_1) and salt cedars (N_2), but now assume flooding at higher levels has negative effects on the growth rate of the N_2 population. We choose linear growth functions such that $\Theta_1 < \Theta_2$ when $u = 0$. Furthermore, Θ_1 is an increasing function of u and Θ_2 is a decreasing function of u which becomes negative for larger values of u.

We assume a lack of control has allowed N_2 to become the dominant species. The

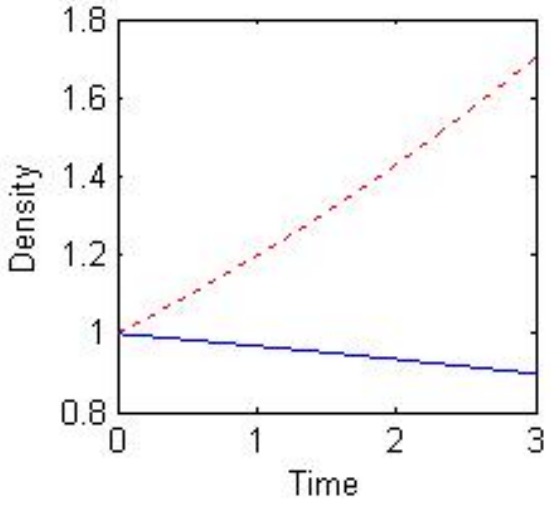


Figure 2.4: N_1 (solid) and N_2 (dotted) densities without control

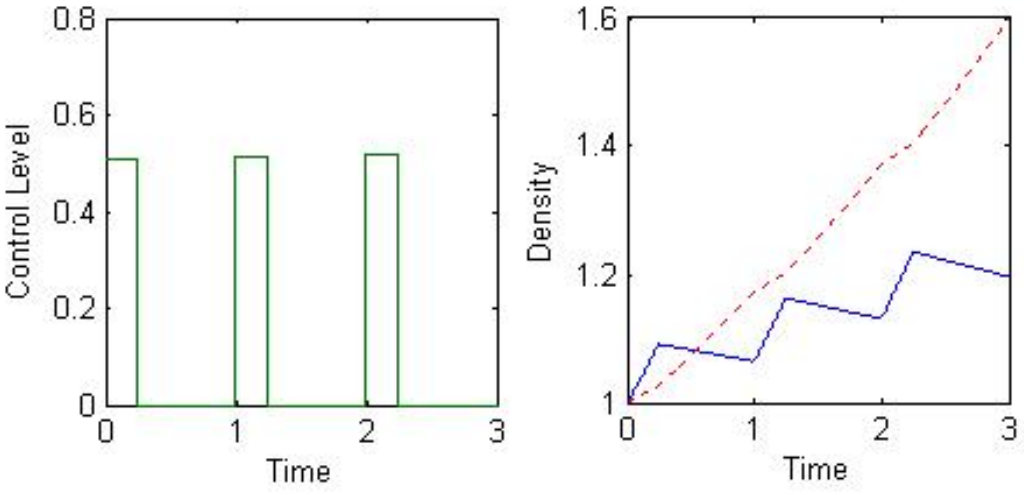


Figure 2.5: Left: Optimal Control Schedule, Right: Corresponding N_1 (solid) and N_2 (dotted) densities

interaction coefficients are the same as presented in Section 2.4.1. The initial conditions and growth rates are

$$G_1 = 1, G_2 = 2 \tag{2.20}$$

$$\Theta_1 = u$$

$$\Theta_2 = -u + .2. \tag{2.21}$$

Without control, N_2 increases and N_1 decreases; see Figure 2.6.

Because minimizing $N_2(T)$ is an important part of our goal, we allow $K > 0$. We use the parameters

$$A = .2, B = 1, K = 1, \epsilon = .5, T = 3.$$

Here we assume the control affects both species during the first one fourth of each year. This could correspond to flooding during the spring thaw but it assumes that such action affects the species' growth rates for several months.

The optimal control and corresponding states are displayed in Figure 2.7. After three years of control, N_1 is clearly the dominating species once again. Because we allow Θ_2 to take negative values for larger values of u, we are able to see a decrease in the N_2 population. This is essential if one not only wants to control an invasive species, but ultimately eliminate the species from the area.

2.5 Concluding Remarks and New Contributions

This work demonstrates that optimal control theory can be an appropriate tool for designing intervention strategies to control an invasive species. An interesting feature of our optimal control problem is that controls act in only part of each year. We proved existence results for an optimal control in certain cases of quadratic growth functions. In those cases, as well as cases for linear growth functions, numerical solutions illustrate optimal intervention for various scenarios with different interaction coefficients. In particular, we illustrate how

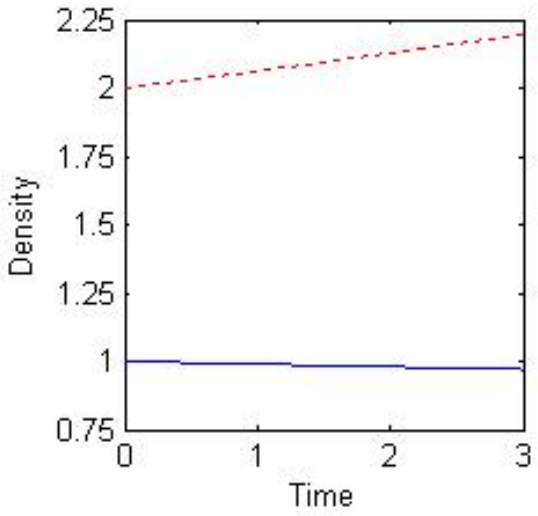


Figure 2.6: N_1 (solid) and N_2 (dotted) densities without control

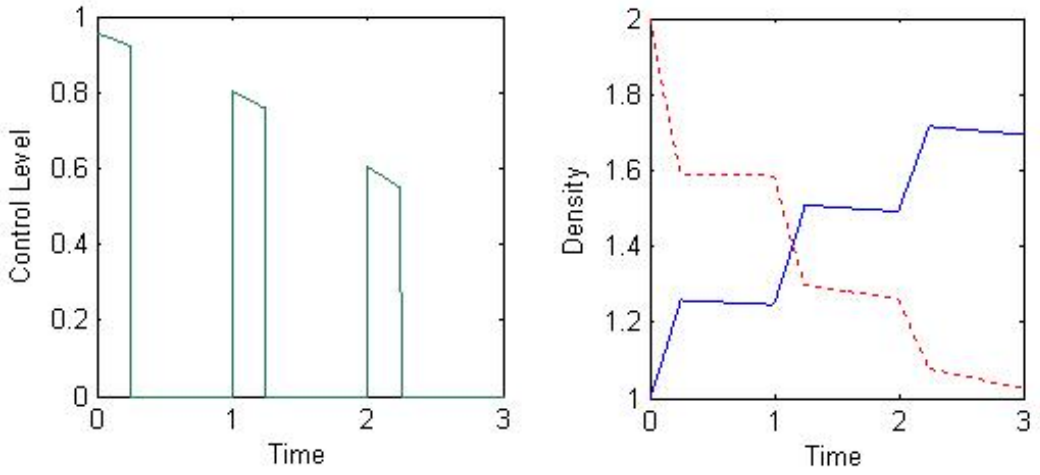


Figure 2.7: Left: Optimal Control Schedule, Right: Corresponding N_1 (solid) and N_2 (dotted) densities

optimal flooding regimes can aid cottonwood growth in the presence of a competing population, such as the salt cedar. Other forms of interaction terms (not involving the controls) may be used, including those for predator-prey scenarios and various functional responses.

We have obtained novel existence results of optimal controls for certain structures of growth functions. Difficulty in proving existence results arises in the quadratic terms in the ODEs involving the controls, due to the fact that maximizing sequences of controls only converge weakly. In the future, we may continue to explore what forms of growth functions will give further existence results.

Acknowledgements

This was done in collaboration with Daniel Kern, Suzanne Lenhart, and Jiongmin Yong. The corresponding article was published in *Journal of Biological Dynamics* 1(4): 413-426, 2007.

Chapter 3

Optimal Intervention Strategies for a Cholera Outbreak

3.1 Background

Cholera is a diarrheal illness caused by infection of the intestine with the bacterium *Vibrio cholerae*. While cholera has been a recognized disease for about two hundred years, controlling deadly outbreaks remains a challange. In fact, incidence has increased in recent years causing enormous loss of life and financial devastation to families and health care systems. In 2007, over 200,000 cases of cholera worldwide were reported [93]. Recent focus on cholera is attributed to a deadly outbreak in South Africa. In December of 2008, the Ministry of Health in Zimbabwe reported a total of 11,735 cholera cases with 484 deaths since August 2008 [92]. However, cholera incidence may in fact be far higher because fear of economic and social losses from an outbreak becoming known often cause countries to underreport outbreaks.

Populations lacking prior immunity to the organism can be devastated by the disease in a matter of weeks [56]. Transmission to humans occurs through eating food or drinking water contaminated with *V. cholerae* from other cholera patients. In the most severe cases dramatic fluid loss from the continuous diarrhea causes dehydration, leading to shock and death within hours [91]. The severity of the symptoms depend on the dose of bacteria

ingested by the infected individual. Outbreaks can be quite hard to predict because a large proportion of infected individuals do not exhibit symptoms or only experience mild symptoms. Although their infection may be inapparent, these individuals still contribute vibrios to environmental sources of bacteria [20, 50].

How does one treat this effectively? Oral rehydration therapy is a simple, cheap, and effective treatment for dehydration associated with mild cases of cholera. The solution consists of salts and sugars and is administered orally. Oral rehydration therapy greatly reduces the death toll from cholera when implemented in frequent small doses. It prevents dehydration from killing an individual with mild cholera symptoms before the disease runs its course [60]. For severe cases, effective treatment requires intravenous rehydration and is often administered in combination with antibiotics. Antibiotics decrease both the duration and volume of diarrhea caused by cholera [36, 86]. There are concerns related to antibiotics' effectiveness and side-effects and to their creation of drug-resistant organisms. In many areas, tetracycline and doxycycline are not used as a prophylactic treatment but as effective and low-cost treatment of individuals with severe cholera symptoms [1, 91].

Presently there is a two-dose oral cholera vaccine (OCV) consisting of killed whole-cell *Vibrio cholerae* that has received World Health Organization prequalification, and several other oral vaccines are under development or have limited availability [93]. The two-dose OCV provides 85-90% protection for adults for the first six months after administration [93], but there are logistical constraints in giving the OCV that limit its success in addressing emergency outbreaks. For example, the bulkiness of the vaccine and buffer solution required for a successful vaccination campaing can complicate shipment to refugee camps and storage at vaccine site [48]. In addition, it is not possible to administer the vaccine to children under the age of two, and its efficacy for older children is not known [93]. Thus, while OCV implementations would most likely be a part of any mitigation strategy, they are unable to address all of the complexities found in varied forms of cholera outbreaks.

Because the vibrios are mostly spread through contaminated water, an obvious strategy would be to address the sanitation of water supplies. This includes the sterilization of cooking, washing, and drinking water by means of boiling or use of cloth filters. It also includes treatment of general sewage water before it enters water supplies. However, there

have been many recent examples of humanitarian crises in which sanitation efforts could not be implemented. For example, in May 2004 the UN found that of the internationally displaced persons in Darfur, Sudan, only 20% had access to adequate water and only 5% had adequate sanitation [93]. More than a year later, approximately half still lacked an adequate water supply and one-fourth lacked adequate sanitation [93]. Thus, even an internationally-recognized crisis area may be unable to receive the needed intervention to provide a safe water supply for its inhabitants [22, 48, 60, 93].

The World Health Organization believes that the balance between OCV administration and water sanitation to achieve an effective mitigation strategy will depend largely on parameters specific to individual populations [93]. In determining strategies, it is important to note the epidemiologic differences in cholera in different settings. For example, in endemic situations, case fatality rates are fairly low and a large proportion of children become ill. On the other hand, in epidemic situations there will be a high case fatality rate unless rehydration facilities are available [84]. In refugee camps, it is suggested that provision of purified water, suitable defaecation facilities, rehydration therapy, and case management should be able to reduce the case fatality rate dramatically [60]. It has traditionally been thought that antibiotic treatment and vaccination therapies are not effective for such settings [60, 93]. In contrast, for settings in which cholera has become endemic, such as in Bangladesh, there have been studies suggesting that strategic vaccinations are practical and can infer "herd immunity" wherein nonvaccinated individuals receive indirect protection [2].

A complex system of interactions occurs between the human host, pathogen, and the environment, and the potential control mechanisms are strongly influenced by socio-economic pressures and realities. Our objective is to formulate a model for cholera that includes relevant biological detail, accounts for multiple intervention strategies, and allows optimal control methods to be applied. We begin by formulating an ODE model which incorporates the dynamics of the susceptible, infected, and recovered human classes and accounts for the vibrio concentrations in drinking water. For our compartment model, the basic reproduction number is calculated. A new result quantifies contributions to the reproductive number from both asymptomatic and symptomatic infections. The ODE system is extended to include the effects of the several controls mentioned above. We use optimal

control theory, parameter sensitivity analysis, and numerical simulations to illustrate the intervention strategies that minimize death due to disease and the cost of treatment during a cholera outbreak. Our mathematical results provide a design for cost-effective strategies of cholera control and provide a framework for control of diseases with multiple intervention methods.

3.2 Modeling the Spread of Cholera

Before formulating our cholera model, we present several previously proposed cholera models and discuss the importance of each in Section 3.2.1. We then present our model formulation in Section 3.2.2 and calculate the basic reproduction number in Section 3.2.3. Because several parameters in our model are difficult to determine or have shown to vary between populations, we conduct a sensitivity analysis in Section 3.2.4 to illustrate which parameters are most influential in determining the severity of a cholera outbreak.

3.2.1 A Brief History of Cholera Models

In 2001, Codeço proposed a model which explicitly accounts for the concentration of *V. cholerae* in the water supply [20]. She models the population sizes of the susceptible (S) and infectious (I) individuals, and the concentration of vibrios (B) by

$$\begin{aligned} \frac{dS}{dt} &= n(H-S) - \beta\lambda(B)S, \\ \frac{dI}{dt} &= \beta\lambda(B)S - rI, \\ \frac{dB}{dt} &= B(nb-mb) + \eta I. \end{aligned}$$

In the system above, H is a constant denoting the size of the total human population, n is the human birth rate, r is a recovery rate, and nb and mb are rates of vibrio growth and loss respectively. The parameter η measures contribution from an infected individual to the vibrio reservoir through excretion and β is the ingestion rate of untreated water, measured in liters per day. The incidence $\beta\lambda(B)S$ is such that

$$\lambda(B) = \frac{B}{K+B},$$

where K is a half saturation constant. Thus, when the concentration of vibrios $B = K$ and consumption of untreated water is $\beta = 1$ liter per day, the probability of an exposed susceptible individual becoming infected is $\frac{1}{2}$.

A term in the model that is difficult to quantify is the contribution of infected individuals to the concentration of *V. cholerae* in the environment. One can imagine that the volume of untreated water varies greatly in different environments and also the lack of sanitary latrines differs. In addition, water supplies are not well-mixed so a contribution of *V. cholerae* to the water would not be evenly distributed. Thus, while there is strong data quantifying the volume of fecal matter for infected individuals and the concentration of *V. cholerae* therein, it is problematic to suggest the correlation of this output to the contamination of the drinking water. Codeço conducted an equilibrium analysis to suggest the correlation between values of β and η which would lead to a cholera outbreak in a hypothetical community. While this approach may be too simple to yield quantitative predictions, the qualitative understanding of model outcomes as values of η change provide interesting dynamics. For example, environmental changes can increase or decrease (perhaps periodically) the parameters β and η and one can observe their impact on an outbreak.

In 2002, Pascual, Bouma, and Dobson [62] published the first of several extensions of the work by Codeço. Pascual *et al.* added a class that tracks the water volume, suggesting a strategy to quantify the relationship between vibrio concentration and the water volume that changes seasonally. The model accounted for precipitation, stream flow, and drainage, and the infectivity was modeled in a way that considers the community volume of water. This structure provided an alternate suggestion for how to show the influence of climatic effects on a cholera outbreak in a model that considers vibrio-to-human transmission.

In 2005, Koelle, Pascual, and Yunus suggested an SIRS ODE model for a disease with a seasonal driver that affects the transmission rate of the pathogen [44]. Separately, Koelle, Rodó, Pascual, Yunus, and Mostofa considered a difference equation model with a different

seasonal driver [45]. Both models considered all disease transmission to be human-to-human. The latter model included both a seasonal and a long-term transmission component, and the former model analyzed a trade-off between climate sensitivity and the maximum transmission rate.

In 2006, Hartley, Morris, and Smith [37] reconsidered the model posed by Codeço in 2001, adding a hyperinfectious (HI) state of the *V. cholerae* to the model. In a letter to *Nature* in 2002, Merrell *et al.* reported the results of a study in which it was found that freshly shed *V. cholerae* from human intestines outcompeted other *V. cholerae* by as much as 700-fold for at least the first 5 hours in the environment [55]. The model by Hartley *et al.* was similar to Codeço's with the exception of an additional state to account for the HI vibrio concentration. It is emphasized that the additional HI state provides a basis for the transmission pathway known as 'human to human'. The model follows the movement of susceptible individuals (S) to an infected class (I) through interaction with a sufficiently large concentration of vibrios, both in a non-HI state (B_L) and an HI state (B_H). The infectious individuals add to the concentration of HI vibrios in the drinking water, which transform to less-infectious in a matter of hours. In a number of days, infected individuals recover and move to the removed state (R). A diagram illustrating the dynamics of the Hartley *et al.* model is seen in Figure B.1.

Hartley *et al.* compute the basic reproductive number for this model, $R_0 = \frac{\xi N}{\gamma+b}\left(\frac{\beta_H}{\kappa_H \chi} + \frac{\beta_L}{\kappa_L \delta_L}\right)$. In this quantity, N is total population, b is the birth rate, γ is the recovery rate, ξ is the vibrio shedding rate, χ is the vibrio transition rate, δ is the vibrio death rate and κ_H and κ_L are half saturation constants for the HI and non-HI vibrios respectively. Of all the parameters, the HI and non-HI vibrio ingestion rates, β_H and β_L, are the most important for control and the least well known. The authors' analysis suggests that the HI state is extremely important in the transmission of epidemic cholera and responsible for 'explosive' outbreaks. When 'human to human' transmission is dominant (i.e. $\beta_H >> \beta_L$), Hartley *et al.* estimate $R_0 \approx 18$. In such a case, interventions should be targeted toward minimizing risk of transmission of the HI vibrios to have maximal impact on limiting spread of cholera.

Shortly afterwards, in June 2006, Pascual *et al.* comment in PLoS Medicine [63] that the formula for R_0 given by Hartley *et al.* is an overestimation of reproductive number for

cholera. A model incorporating 'human to human' transmission fitted with data collected from Matlab, Bangladesh indicates a more appropriate value is $R_0 \approx 3$. Pascual *et al.* also point out that while an epidemic declines from a depletion of susceptibles in Hartley *et al.*'s model, the seasonal outbreaks observed in Matlab are curtailed prior to a significant depletion of susceptibles indicating the presence of an environmental factor. The dynamics of vibrio-specific phage predation are a likely mechanism for the observed reduction in cholera transmission rate at the end of a seasonal outbreak. Despite discrepancies in the models of Pascual *et al.* and Hartley *et al.*, both accentuate the need to include some variant of 'human to human' transmission to explain cholera dynamics.

Recently, in August 2008, King, Ionides, Pascual and Bouma propose three mechanistic models of the SIRS form for cholera. The authors investigate the models' ability to explain monthly cholera death counts in the twenty-six districts of the former British East Indian province of Bengal during the period 1891 - 1940 [42]. Figure B.2 in Appendix B is from King *et al.* and displays the twenty-six districts. The models incorporate both transmission due to human prevalence via a mass action term and transmission from an environmental reservoir. One of the three models proposed is a two-path model (Figure B.3) which includes a class for severe infections as well as a class for mild, inapparent infections. In the two-path model, infected individuals with the mild symptoms experience short-term immunity to severe cholera infection and negligible death due to disease. The periodic force of infection depends on a seasonal component as well as the proportion of the population with severe infections.

Log-likelihood estimates were generated using the 1900's mortality data and indicate that six of the twenty-six districts are well-described by the two-path model. Maximum likelihood estimates for the six districts are displayed in Table B.1. Estimates for the six districts indicate considerable variation in parameter values between populations. For example, the percent of cholera infections resulting in inapparent symptoms ranges from 70 to 99 between populations. Variation is also observed in recovery rates, cholera-related death rates, and rates of waning immunity. A noteworthy prediction of these estimates is that immunity must wane on a timescale of weeks to months as opposed to the previously suspected timescale of 3-10 years [42].

3.2.2 Model Formulation

The inclusion of hyperinfectious *V. cholerae*, inapparent human infections, and waning immunity are essential in modeling the spread of cholera. Previously, no one cholera model contained all of these features. We formulate an S-I-R type model to include the following key components:

- **Hyperinfectious (HI) and less-infectious (non-HI) vibrios.** To capture both 'human-to-human' and 'environment-to-human' pathways of transmission, we acknowledge and differentiate between the concentration of HI vibrios (class B_H) and the concentration of less-infectious or non-HI vibrios (class B_L) in drinking water. The rate at which a susceptible individual ingests water recently contaminated with HI vibrios, perhaps by an infected member of the same or nearby household, is denoted by β_H. The rate at which a susceptible individual ingests water contaminated with non-HI vibrios, perhaps from an environmental reservoir, is denoted by β_L. As done in Hartley *et al.* [37], half saturation constants for the vibrio concentrations are denoted as κ_H and κ_L.

- **A class of asymptomatic infecteds**. Strategies for the control of cholera outbreaks must address severe infections as well as mild, inapparent infections. Individuals with severe infections are classified as symptomatic infecteds (class I_S) and those with mild or inapparent infections are classified as asymptomatic infecteds (class I_A) in our model. The proportion infections resulting in mild or inapparent symptoms is denoted by p and the proportion resulting in severe symptoms is denoted by $1 - p$. Both infected classes actively shed HI vibrios, contributing to the source of infection. For the asymptomatic and symptomatic infected classes respectively, e_1 and e_2 are cholera-related death rates, γ_1 and γ_2 are rates of recovery from cholera, and η_1 and η_2 are vibrio shedding rates.

- **Waning immunity.** We assume that the time an individual spends in the immune class (class R) is exponentially distributed with an average length of $\frac{1}{\omega}$ days. Once their immunity has waned, individuals re-enter the susceptible population (class S).

Summarized in Figure 3.1, state variables track susceptible, asymptomatic infected, symptomatic infected, and recovered human populations as well as HI and non-HI vibrio concentrations. The system of six ODEs describing the dynamics between states S, I_A, I_S, R, B_H, and B_L are

$$\frac{dS}{dt} = -\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big] S(t) + \omega R(t), \tag{3.1}$$

$$\frac{dI_A}{dt} = p\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big] S(t) - (e_1 + \gamma_1) I_A(t), \tag{3.2}$$

$$\frac{dI_S}{dt} = (1-p)\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big] S(t) - (e_2 + \gamma_2) I_S(t), \tag{3.3}$$

$$\frac{dR}{dt} = \gamma_1 I_A(t) + \gamma_2 I_S(t) - \omega R(t), \tag{3.4}$$

$$\frac{dB_H}{dt} = \eta_1 I_A(t) + \eta_2 I_S(t) - \chi B_H(t), \tag{3.5}$$

$$\frac{dB_L}{dt} = \chi B_H(t) - \delta B_L(t), \tag{3.6}$$

with initial conditions

$$S(0) = S_0, \quad I_A(0) = I_{A0}, \quad I_S(0) = I_{S0}, \quad R(0) = R_0, \quad B_L(0) = B_{L0}, \quad B_H(0) = B_{H0}. \tag{3.7}$$

Parameter notation, descriptions, and units are displayed in Table 3.1. Because we are interested in short-term analysis (less than 1 year), natural birth and death terms are excluded from our model. It should be noted that disease spread can occur during certain funeral rituals. In this model, we do not account for contamination of water supplies from vibrios of deceased individuals.

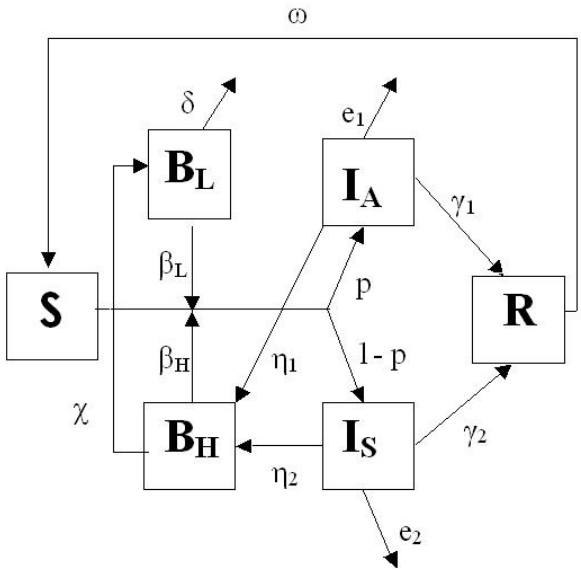


Figure 3.1: Diagram of our model following the movement of susceptible individuals (S) to either an asymptomatic infected class (I_A) or a symptomatic infected class (I_S) through interaction with a sufficiently large concentration of vibrios, both in a non-HI state (B_L) and an HI state (B_H). Individuals in the recovered class (R) lose immunity over time and move back to the susceptible class.

Table 3.1: Summary of model notation and units

Notation	**Description**	**Units**
$S(t)$	Number of susceptible humans on day t	humans
$I_A(t)$	Number of asymptomatic infected humans on day t	humans
$I_S(t)$	Number of symptomatic infected humans on day t	humans
$R(t)$	Number of recovered or immune humans on day t	humans
$B_L(t)$	Concentration of non-HI vibrio population in drinking water on day t	cells/ml
$B_H(t)$	Concentration of HI vibrio population in drinking water on day t	cells/ml
p	Probability of infected individual being asymptomatic	none
β_L	Rate of drinking non-HI vibrios in untreated water	day^{-1}
β_H	Rate of drinking HI vibrios in untreated water	day^{-1}
κ_L	Half saturation constant of non-HI vibrios	cells/ml
κ_H	Half saturation constant of HI vibrios	cells/ml
e_1	Cholera-related death rate for asymptomatic infecteds	day^{-1}
e_2	Cholera-related death rate for symptomatic infecteds	day^{-1}
γ_1	Cholera recovery rate for asymptomatic infecteds	day^{-1}
γ_2	Cholera recovery rate for symptomatic infecteds	day^{-1}
ω	Rate of waning cholera immunity	day^{-1}
η_1	Rate of contribution to HI vibrios in environment by asymptomatic infecteds	$\frac{\text{cells}}{\text{ml-day-human}}$
η_2	Rate of contribution to HI vibrios in environment by symptomatic infecteds	$\frac{\text{cells}}{\text{ml-day-human}}$
χ	Transition rate of vibrios from HI to non-HI state	day^{-1}
δ	Death rate of vibrios in the environment	day^{-1}

3.2.3 The Basic Reproduction Number, $\mathcal{R}_0$

The basic reproduction number, $\mathcal{R}_0$, is the expected number of secondary cases produced in a completely susceptible population by a typical infected individual. If $\mathcal{R}_0 < 1$, then an average infected individual produces less than one new infection over the course of its infectious period and the infection cannot grow. Conversely, if $\mathcal{R}_0 > 1$, then an average infected individual produces more than one new infection and the disease can invade the population. For cholera, a disease characterized by hidden infections, determining the appropriate value of $\mathcal{R}_0$ is challenging, but essential for developing methods of control.

To determine the basic reproduction number for our compartmental model, we use methods set forth by van den Driessche and Watmough [27]. For only the calculations done in this section, consider the system to be ordered such that $x = (I_A, I_S, B_H, B_L, R, S)$. Let N denote the total number of human individuals in a disease-free population. The disease-free equilibrium (DFE) is $x_0 = (0, 0, 0, 0, 0, N)$. New infections occur in compartments 1 and 2 only. Compartment 3 and 4 are classified as infectious compartments, but vibrios in these compartments are not considered to be new infections. For $1 \leq i \leq 6$, let $\mathbb{F}_i(x)$ be the rate of new infections in compartment i, $V_i^+(x)$ be the rate of transfer of humans or vibrios into compartment i by all other means, and $V_-^+(x)$ be the rate of transfer of humans or vibrios out of compartment i. Setting $\mathbb{V}_i(x) = V_i^-(x) - V_i^+(x)$, we have the following two vectors.

$$\mathbb{F} = \begin{pmatrix} p\left[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\right] S \\ (1-p)\left[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\right] S \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbb{V} = \begin{pmatrix} (e_1+\gamma_1)I_A \\ (e_2+\gamma_2)I_S \\ \chi B_H - \eta_1 I_A - \eta_2 I_S \\ \delta B_L - \chi B_H \\ \omega R - \gamma_1 I_A - \gamma_2 I_S \\ \left(\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\right] S - \omega R \end{pmatrix}$$

Compartments 1 through 4 are considered to be the infectious compartments. As done in [27], we define $\mathbf{F} = \left[\frac{\partial \mathbb{F}_i}{\partial x_j}(x_0)\right]$ and $\mathbf{V} = \left[\frac{\partial \mathbb{V}_i}{\partial x_j}(x_0)\right]$ for $1 \leq i, j \leq 4$.

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & pN\frac{\beta_H}{\kappa_H} & pN\frac{\beta_L}{\kappa_L} \\ 0 & 0 & (1-p)N\frac{\beta_H}{\kappa_H} & (1-p)N\frac{\beta_L}{\kappa_L} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} e_1+\gamma_1 & 0 & 0 & 0 \\ 0 & e_2+\gamma_2 & 0 & 0 \\ -\eta_1 & -\eta_2 & \chi & 0 \\ 0 & 0 & -\chi & \delta \end{pmatrix}$$

We then proceed to calculate $\mathbf{V}^{-1}$ and find

$$\mathbf{V}^{-1} = \begin{pmatrix} \frac{1}{e_1+\gamma_1} & 0 & 0 & 0 \\ 0 & \frac{1}{e_2+\gamma_2} & 0 & 0 \\ \frac{\eta_1}{(e_1+\gamma_1)\chi} & \frac{\eta_2}{(e_2+\gamma_2)\chi} & \frac{1}{\chi} & 0 \\ \frac{\eta_1}{(e_1+\gamma_1)\delta} & \frac{\eta_2}{(e_2+\gamma_2)\delta} & \frac{1}{\delta} & \frac{1}{\delta} \end{pmatrix}.$$

The product $\mathbf{FV}^{-1}$ is

$$\mathbf{FV}^{-1} =$$
$$\begin{pmatrix} pN\frac{\eta_1}{(e_1+\gamma_1)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & pN\frac{\eta_2}{(e_2+\gamma_2)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & pN\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & pN\frac{\beta_L}{\kappa_L\delta} \\ (1-p)N\frac{\eta_1}{(e_1+\gamma_1)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & (1-p)N\frac{\eta_2}{(e_2+\gamma_2)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & (1-p)N\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right) & (1-p)N\frac{\beta_L}{\kappa_L\delta} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Van den Driessche and Watmough call $\mathbf{FV}^{-1}$ the next generation matrix. To interpret the entries of $\mathbf{FV}^{-1}$, consider the fate of an infectious human or vibrio introduced into compartment k of a disease-free population. The (j,k) entry of $\mathbf{V}^{-1}$ is the average length of time this human or vibrio spends in compartment j during its lifetime, assuming the population remains near the DFE $x_0 = (0,0,0,0,0,N)$ and barring reinfection. The (i,j) entry of $\mathbf{F}$ is the rate at which an infectious human or vibrio in compartment j produces new infections in compartment i. Hence, the (i,k) entry of the product FV^{-1} is the expected number of new infections in compartment i produced by the infectious human or vibrio originally introduced into compartment k.

The basic reproduction number is defined to be

$$\mathcal{R}_0 = \rho(\mathbf{FV}^{-1}) \tag{3.8}$$

where $\rho(A)$ denotes the spectral radius of a matrix A. Van den Driessche and Watmough show that R_0 defined by (3.8) is a threshold parameter for the stability of the disease-free equilibrium. The following theorem summarizes this concept.

Theorem 4. *Consider the disease transmission model given by (3.1)-(3.6). If x_0 is a disease-free equilibrium of the model, then x_0 is locally asymptotically stable if $R_0 < 1$, but unstable if $R_0 > 1$, where R_0 is defined by (3.8).*

See [27] for the proof. After calculating the eigenvalues of $\mathbf{FV}^{-1}$ above, we find the reproduction number for our cholera model to be

$$\mathcal{R}_0 = \rho(\mathbf{F}\mathbf{V}^{-1}) = pN\frac{\eta_1}{(e_1+\gamma_1)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right)+(1-p)N\frac{\eta_2}{(e_2+\gamma_2)}\left(\frac{\beta_L}{\delta\kappa_L}+\frac{\beta_H}{\chi\kappa_H}\right). \tag{3.9}$$

Unlike previous estimates of $\mathcal{R}_0$ for cholera [37, 42, 63], the expression above is a measure of all secondary infections, asymptomatic or symptomatic, generated by an average infected individual in a population of N susceptibles. Thus, we expect values of $\mathcal{R}_0$ based on our derivation to exceed those estimated from models or data which do not account for asymptomatic cholera infections.

The expression in (3.9) is a weighted average of the basic reproductive number generated from each of the asymptomatic and symptomatic infected classes. Seen within each component of the sum, the expressions $\frac{\beta_H}{\kappa_H}$ and $\frac{\beta_L}{\kappa_L}$ are the number of new cases generated in terms of the HI and non-HI vibrios, respectively, per unit time as measured by the concentrations. The terms $\frac{1}{\chi}$ and $\frac{1}{\delta}$ are the expected times vibrios remain in the HI and non-HI states before decaying or dying. To ultimately reduce a populations $\mathcal{R}_0$, our result in (3.9) echos the importance of accelerating recovery from cholera (γ_1, γ_2) and reducing shedding rates (η_1, η_2) in both infected classes as well as reducing the populations rates of vibrio ingestion (β_H, β_L).

3.2.4 Parameter Sensitivity Analysis

Cholera-related parameters values can be difficult to determine and can vary considerably between populations [37, 42]. A sensitivity analysis is conducted to determine the impact of several parameters in our model on various output measures. Three outputs are chosen as measures of outbreak severity. Over the total time period of the model, outbreak severity is measured by (1) the maximum and (2) the total size of the symptomatic infected class and (3) the total size of the combined infected classes. For the sensitivity analysis, we employ a Latin Hypercube Sampling (LHS) scheme and calculate partial rank correlation coefficients (PRCC) and p-values corresponding to each parameter and output measure. For detailed methodology, see [11] or [43]. The p-value is used to determine if a PRCC is significantly

Table 3.2: **Parameters used in sensitivity analysis.** Lower and upper bounds for sampled ranges are displayed. All values are in units day^{-1}.

	Description	**Lower**	**Upper**	**Ref.**
p	Proportion of infections that are asymptomatic	0.70	0.99	[42]
r	Multiplicative factor determining β_H, where $\beta_H = r\beta_L$	0.01	1.0	[20, 37]
ω	Rate of waning immunity	0.0005	0.005	[42]
e_2	Cholera-related death rate of symptomatic infected	0.0005	0.02	[42]
γ_1	Cholera recovery rate for asymptomatic infected	0.1	0.2	[38]
γ_2	Cholera recovery rate for symptomatic infected	0.02	0.14	[42]
η_1	Vibrio shedding rate for asymptomatic infected	0.05	1.0	[20, 50]
s	Multiplicative factor determining η_2, where $\eta_2 = s\eta_1$	1	100	[20, 50]

different from zero. A statistically significant PRCC value near -1 or 1 suggests that changes in the input parameter influence change in the outcome measure. Specifically, a negative (positive) PRCC indicates that an increase in the input parameter decreases (increases) the output measure.

Each parameter of interest is sampled using a uniform distribution over a range of biologically realistic values. See Table 3.2 for a summary of parameter descriptions and sampled ranges. Using equations (3.1)-(3.6), n model simulations over a period of 100 days are performed by randomly pairing sampled values for all LHS parameters. The outcome measures are checked to be monotone with respect to each input parameter. We apply the MATLAB 'regress' function to each column-pair of ranked input parameters and outcome measures and then apply the MATLAB 'corr' function to the linear regression residuals, which calculates the PRCCs and corresponding p-values. See Table 3.3 for the PRCC and p-value results. An output is assumed sensitive to an input if the corresponding PRCC is less than -0.50 or greater than 0.50, and the corresponding p-value is less than 0.05.

Results of the sensitivity analysis suggest that all outcome measures are sensitive to changes in p, r, η_1, and s. Since the parameter p measures the proportion of infected who are asymptomatic, the significance in determining the maximum and total number of symptomatic infecteds is not surprising. The effect of p on the third outcome measure suggests a more subtle dynamic. As the proportion of infected who are asymptomatic increases, the

Table 3.3: **Sensitivity analysis results** For each parameter, the PRCC corresponding to each of the three measures of outbreak severity is displayed. Statistical significance of associated p-value is indicated with astericks (* p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001).

Parameter	**Maximum I_S**	**Total I_S**	**Total ($I_A + I_S$)**
p	-0.95 ***	-0.97 ***	-0.59 ***
r	0.68 ***	0.41 ***	-0.53 ***
ω	0.11 *	0.35 ***	0.59 ***
e_2	-0.22 ***	-0.18 ***	-0.11 *
γ_1	-0.13 *	-0.08	0.10
γ_2	-0.76 ***	-0.28 ***	-0.42 ***
η_1	0.67 ***	0.59 ***	0.70 ***
s	0.62 ***	0.58 ***	0.69 ***

total size of the combined infected populations over the entire outbreak decreases. The decrease in this measure of outbreak severity is due to the decreased concentration of vibrios in drinking water, a result of fewer symptomatic infecteds. The parameter r measures the ingestion rate of hyperinfectious vibrios relative to the ingestion rate of less-infectious vibrios ($\beta_H = r\beta_L$). This parameter most influences the maximum size of symptomatic population during the outbreak. The importance of parameters η_1 and s to all outcome measures is not surprising as the parameters measure the contribution of asymptomatic individuals to the environmental bacteria concentration and increased contribution of symptomatic individuals, respectively.

In addition, the maximum size of the symptomatic infected population is significantly affected by the rate of recovery of symptomatic infecteds, γ_2, and variation in the total number of infected individuals is influenced by the rate of waning immunity, ω. We further conclude that, with ranges as suggested in the literature, the parameters e_2 and γ_1 account for little variation in the outcome measures and exact estimates or control measures affecting these parameters have little contribution to the predicted outcome of a cholera outbreak.

3.3 Optimal Control Problem with Multiple Controls

3.3.1 Introduction of Controls and Objective Functional

Several cholera treatment and prevention options do exist. Mild and moderate cases can be successfully treated with oral rehydration salts alone. This cheap and effective treatment can be implemented in homes without professional assistance, and thus we do not consider treatment of asymptomatic infections as a control with trade offs. We extend our model described by system (3.1) - (3.6) to include the effects of following three cholera controls. The controls are expressed as Lebesgue measurable functions of time. The notation for each control as well as the notation for each set of admissable controls is defined.

- **Rehydration Therapy combined with Antibiotic Treatment.** Classified as symptomatic infected in our model, individuals with severe infections require rehydration with intravenous fluids and antibiotics. Rehydration therapy can considerably reduce the chance of death due to cholera and, when combined with antibiotics, the time spent in the infected class. Thus, those receiving this combined treatment are assumed to have an increased rate of recovery ($\gamma_3 > \gamma_2$) and a decreased rate of death due to cholera ($e_3 < e_2$). We denote the proportion of symptomatic infecteds who receive rehydration and antibiotic treatment as $u(t)$. The set of admissible controls representing this treatment is

$$U = \{0 \leq u(t) \leq u_{max} < 1 | u \textit{ Lebesgue measurable}\}.$$

- **Vaccination.** The rate at which individuals gain immunity through vaccination is denoted by $v(t)$. Because asymptomatic infecteds may not be aware of their infection, we assume susceptibles and asymptomatic infecteds are indistinguishable with respect to vaccination. Vaccinated susceptibles move to the immune class. Vaccinating asymptomatic infected individuals has no effect, but still implies a cost. Natural immunity and immunity gained from vaccination are assumed to wane at the same

rate, ω. The set of admissible vaccination controls is

$$V = \{0 \leq v(t) \leq v_{max} | v \textit{ Lebesgue measurable}\}.$$

- **Sanitation.** We let $m(t)$ denote the level of sanitation efforts, scaled so that $0 \leq m(t) \leq m_{max} \leq 1$. Sanitation efforts include the sterilization of cooking, washing, and drinking water by means of boiling or use of cloth filters. It also includes treatment of general sewage water before it enters water supplies. The effect of sanitation efforts is modeled as a reduction in vibrio ingestion rates. The level of sanitation decreases the transmission rate of the hyperinfectious vibrios (β_H) and the transmission rate of the less infectious vibrios (β_L) as reflected by the terms $(1-m(t))\beta_L$ and $(1-m(t))\beta_H$ respectively. The set of admissible sanitation controls is

$$M = \{0 \leq m(t) \leq m_{max} \leq 1 | m \textit{ Lebesgue measurable}\}.$$

For simplicity, we will denote a control triple by $\bar{u} = (u, v, m)$ and define the set of admissible controls as $\bar{U} = U \times V \times M$. Presented in (3.10)-(3.15) is the system of six state differential equations extended to include the effects of the three controls. Controls **u**, **v**, and **m** are functions of time and emphasized in bold. Table 3.4 displays additional parameter notation, descriptions and units for the model with control. All other parameters are the same as described in Table 3.1.

The system of state equations with controls is

$$\frac{dS}{dt} = -(1 - \mathbf{m(t)})\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big]S(t) + \omega R(t) - \mathbf{v(t)}S(t), \quad (3.10)$$

$$\frac{dI_A}{dt} = p(1 - \mathbf{m(t)})\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big]S(t) - (e_1 + \gamma_1)I_A(t), \quad (3.11)$$

Table 3.4: Additional notation for model with control

Notation	Description	Units
$u(t)$	Proportion of symptomatic infecteds receiving rehydration and antibiotics on day t	none
u_{max}	Upper bound for proportion of symptomatic infecteds receiving rehydration and antibiotics	none
$v(t)$	Rate at which susceptibles and asymptomatic infecteds are vaccinated on day t	day^{-1}
v_{max}	Upper bound for vaccination rate	day^{-1}
$m(t)$	Sanitation level on day t	none
m_{max}	Upper bound for sanitation level	none
e_3	Cholera-related death rate for symptomatic infecteds receiving rehydration and antibiotics	day^{-1}
γ_3	Cholera recovery rate for symptomatic infecteds with rehydration and antibiotics	day^{-1}

$$\frac{dI_S}{dt} = (1-p)(1-\mathbf{m(t)})\big[\beta_L \frac{B_L(t)}{\kappa_L + B_L(t)} + \beta_H \frac{B_H(t)}{\kappa_H + B_H(t)}\big] S(t) - (e_2+\gamma_2)(1-\mathbf{u(t)})I_S(t) - (e_3+\gamma_3)\mathbf{u(t)}I_S(t), \tag{3.12}$$

$$\frac{dR}{dt} = \gamma_1 I_A(t) + \gamma_2(1-\mathbf{u(t)})I_S(t) + \gamma_3 \mathbf{u(t)} I_S(t) - \omega R(t) + \mathbf{v(t)}S(t), \tag{3.13}$$

$$\frac{dB_H}{dt} = \eta_1 I_A(t) + \eta_2 I_S(t) - \chi B_H(t), \tag{3.14}$$

$$\frac{dB_L}{dt} = \chi B_H(t) - \delta B_L(t), \tag{3.15}$$

with initial conditions

$$S(0) = S_0, \quad I_A(0) = I_{A0}, \quad I_S(0) = I_{S0}, \quad R(0) = R_0, \quad B_L(0) = B_{L0}, \quad B_H(0) = B_{H0}. \tag{3.16}$$

A short-term analysis for optimal intervention strategies is practical because the dynamics and parameter values of the model will change over a period of time. One may want

to reassess the situation at a later time and compute the optimal strategy given the current scenario. Because we investigate intervention strategies over a period of 100 days in this chapter, we exclude natural birth and death rates. The model could easily be modified to include a recruitment term to signify an influx of individuals that might be seen in a refugee camp and/or to include population growth dynamics.

A successful mitigation scheme is one which minimizes both the number of deaths due to disease and the cost of the controls until the final time, T. In computing the cost of vaccinating, we assume that asymptomatic infecteds are indistinguishable from susceptibles with respect to vaccination, thus incurring costs.

A control scheme $\bar{u}$ is considered optimal if it minimizes the objective functional defined as

$$J(\bar{u}) = \int_0^T \big[A\left(e_1 I_A(t) + e_2(1-u(t))I_S(t) + e_3 u(t) I_S(t)\right) + B_1 m(t) + C_1 m(t)^2 + B_2 u(t) I_S(t) + C_2 u(t)^2 + B_3 v(t)(S(t) + I_A(t)) + C_3 v^2(t)\big] dt \tag{3.17}$$

where $A, B_1, B_2, B_3, C_1, C_2, C_3$ are positive balancing coefficients which transform the integrand into units of dollars. The first sum, multiplied by A, is the cost of death due to cholera and the remaining expressions are costs for implementation of the three controls. Quadratic expressions of the controls are included to indicate non-linear costs potentially arising at high intervention levels.

The optimal control problem is stated as

$$\min_{\bar{u} \in \bar{U}} J(\bar{u})$$

subject to the state system (3.10)-(3.15) and initial conditions (3.16).

3.3.2 Existence and Optimality System

The existence of the optimal control triple can be obtained using a result by Fleming and Rishel ([34], Theorem 4.1, pp 68-69). We first show that the set of all feasible solutions to the control problem is non-empty.

Theorem 5. *Given $\bar{u} \in \bar{U}$ there exists a bounded solution $(S, I_A, I_S, R, B_H, B_L)$ to the initial value problem defined in (3.10)-(3.16).*

Proof: The solutions to (3.10)-(3.15) are bounded below by the trivial solution. An upper bound can be proven using the fact that a supersolution $(\bar{S}, \bar{I_A}, \bar{I_S}, \bar{R}, \bar{B_H}, \bar{B_L})$ satisfying

$$\begin{aligned}
\frac{d\bar{S}}{dt} &= \omega\bar{R}(t), \\
\frac{d\bar{I}_A}{dt} &= (\beta_L + \beta_H)\bar{S}(t), \\
\frac{d\bar{I}_S}{dt} &= (\beta_L + \beta_H)\bar{S}(t), \\
\frac{d\bar{R}}{dt} &= \gamma_1\bar{I}_A(t) + (\gamma_2 + \gamma_3)\bar{I}_S(t), \\
\frac{d\bar{B}_H}{dt} &= \eta_1\bar{I}_A(t) + \eta_2\bar{I}_S(t), \\
\frac{d\bar{B}_L}{dt} &= \chi\bar{B}_H(t),
\end{aligned}$$

is bounded on a finite time interval. Standard results in [54] give existence of a solution to (3.10)-(3.15), a nonlinear system with measurable coefficients. □

Theorem 6. *There exist an optimal triple $\bar{u}^* = (u^*, v^*, m^*) \in \bar{U}$ with corresponding states $(S^*, I_A^*, I_S^*, R^*, B_H^*, B_L^*)$ that minimizes the objective functional $J(\bar{u})$ defined by (3.17).*

Proof: Verifying the conditions needed for the result in [34], we note that the right-hand sides of the differential equations (3.10)-(3.15) are linear in each of the controls and can be written as $\bar{f}(t, \bar{x}, \bar{u}) = \bar{\alpha}(t, \bar{x}) + \bar{\beta}(t, \bar{x})\bar{u}$ where $\bar{x} = (S, I_A, I_S, R, B_H, B_L)$. The boundedness of solutions gives $|\bar{f}(t, \bar{x}, \bar{u})| \leq \bar{C}(1 + |\bar{x}| + |\bar{u}|)$ for $0 \leq t \leq T$. Furthermore, $\bar{U}$ is a closed, convex set. The integrand of our objective functional,

$$L(t,\bar{x},\cdot) = \big[A\left(e_1 I_A(t) + e_2(1-u(t))I_S(t) + e_3 u(t) I_S(t)\right) + B_1 m(t) + C_1 m(t)^2$$
$$+ B_2 u(t) I_S(t) + C_2 u(t)^2 + B_3 v(t)(S(t) + I_A(t)) + C_3 v^2(t)\big],$$

is convex on $\bar{U}$. We must verify the last condition by showing there exists δ_1, δ_2, $\beta > 0$ such that

$$L(t,\bar{x},\bar{u}) \geq \delta_1 |\bar{u}|^\beta - \delta_2.$$

However, this holds for any $\delta_2 > 0$ if we assume $\delta_1 = \min\{C_1, C_2, C_3\}$ and $\beta = 2$. Using the existence result in [34], we conclude there exists an optimal triple $\bar{u}^* \in \bar{U}$. □

We use Pontryagin's Maximum Principle [66] to characterize the controls. Forming the Hamiltonian, we have

$$\begin{aligned}
H = {} & A(e_1 I_A + e_2(1-u)I_S + e_3 u I_S) \\
& + B_1 m + C_1 m^2 + B_2 u I_S + C_2 u^2 + B_3 v(S + I_A) + C_3 v^2 \\
& + \lambda_S \left(-(1-m)\big[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\big] S + \omega R - vS \right) \\
& + \lambda_{I_A} \left(p(1-m)\big[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\big] S - (e_1 + \gamma_1) I_A \right) \\
& + \lambda_{I_S} \big((1-p)(1-m)\big[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\big] S \\
& \qquad - (e_2 + \gamma_2)(1-u) I_S - (e_3 + \gamma_3) u I_S \big) \\
& + \lambda_R \left(\gamma_1 I_A + \gamma_2 (1-u) I_S + \gamma_3 u I_S - \omega R + vS \right) \\
& + \lambda_{B_H} \left(\eta_1 I_A + \eta_2 I_S - \chi B_H \right) + \lambda_{B_L} \left(\chi B_H - \delta_L B_L \right)
\end{aligned}$$

where λ_S, λ_{I_A}, λ_{I_S}, λ_R, λ_{B_H}, λ_{B_L} are the adjoint variables associated with their respective states.

Theorem 7. *Given an optimal triple $\bar{u}^* = (u^*, v^*, m^*) \in \bar{U}$ and corresponding states $(S^*, I_A^*, I_S^*, R^*, B_H^*, B_L^*)$, there exist adjoint functions satisfying*

$$\frac{d\lambda_S}{dt} = v(\lambda_S - \lambda_R - B_3)$$
$$+ (1-m)\big[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\big](\lambda_s - p\lambda_{I_A} - (1-p)\lambda_{I_S}), \tag{3.18}$$
$$\frac{d\lambda_{I_A}}{dt} = -Ae_1 - B_3 v + e_1\lambda_{I_A} + \gamma_1(\lambda_{I_A} - \lambda_R) - \eta_1\lambda_{B_H}, \tag{3.19}$$
$$\frac{d\lambda_{I_S}}{dt} = -A(e_2(1-u) + e_3 u) - B_2 u + \lambda_{I_S}(e_2(1-u) + e_3 u)$$
$$+ (\lambda_{I_S} - \lambda_R)(\gamma_2(1-u) + \gamma_3 u) - \eta_2\lambda_{B_H}, \tag{3.20}$$
$$\frac{d\lambda_R}{dt} = \omega(\lambda_R - \lambda_S), \tag{3.21}$$
$$\frac{d\lambda_{B_H}}{dt} = (1-m)\beta_H \frac{\kappa_H}{(\kappa_H + B_H)^2} S(\lambda_S - p\lambda_{I_A} - (1-p)\lambda_{I_S}) + \chi(\lambda_{B_H} - \lambda_{B_L}), \tag{3.22}$$
$$\frac{d\lambda_{B_L}}{dt} = (1-m)\beta_L \frac{\kappa_L}{(\kappa_H + B_L)^2} S(\lambda_S - p\lambda_{I_A} - (1-p)\lambda_{I_S}) + \lambda_{B_L}\delta, \tag{3.23}$$

with transversality conditions

$$\lambda_S(T) = 0, \quad \lambda_{I_A}(T) = 0, \quad \lambda_{I_S}(T) = 0, \quad \lambda_R(T) = 0, \quad \lambda_{B_H}(T) = 0, \quad \lambda_{B_L}(T) = 0.$$

Furthermore, the optimal controls are characterized by

$$u^*(t) =$$
$$\max\left(0, \min\left(\frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}, u_{\max}\right)\right),$$

$$v^*(t) = \max\left(0, \min\left(\frac{(\lambda_R - \lambda_S)S + B_3(I_A + S)}{-2C_3}, v_{\max}\right)\right),$$

and

$$m^*(t) = \max\left(0, \min\left(\frac{B_1 + (\lambda_S - p\lambda_{I_A} - (1-p)\lambda_{I_S}\big[\beta_L \frac{B_L}{\kappa_L + B_L} + \beta_H \frac{B_H}{\kappa_H + B_H}\big]S}{-2C_1}, m_{\max}\right)\right).$$

Proof: The differential equations for the adjoints are standard results from Pontryagin's Maximum Principle [66]. The right-hand sides of the differential equations can be easily computed by

$$\begin{aligned}
\frac{d\lambda_S}{dt} &= \frac{-\partial H}{\partial S},\\
\frac{d\lambda_{I_A}}{dt} &= \frac{-\partial H}{\partial I_A},\\
\frac{d\lambda_{I_S}}{dt} &= \frac{-\partial H}{\partial I_S},\\
\frac{d\lambda_R}{dt} &= \frac{-\partial H}{\partial R},\\
\frac{d\lambda_{B_H}}{dt} &= \frac{-\partial H}{\partial B_H},\\
\frac{d\lambda_{B_L}}{dt} &= \frac{-\partial H}{\partial B_L}.
\end{aligned}$$

The final time conditions are due to the transversality conditions. Because there is no salvage term in the objective functional, the final time conditions are zero.

In characterizing each control, we consider three cases concerning the control bounds. We show this in detail for the characterization of u^*.

1. On the set $\{t|0 < u^*(t) < u_{\max}\}$, we have

$$\begin{aligned}
0 &= \left.\frac{\partial H}{\partial u}\right|_{u^*}\\
&= (-Ae_2 + Ae_3 + B_2)I_S + 2C_2u^* + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2).
\end{aligned}$$

Solving the above for u^* yields

$$u^*(t) = \frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}.$$

2. On the set $\{t|u^*(t) = 0\}$, we have

$$0 = \left.\frac{\partial H}{\partial u}\right|_{u^*}$$
$$= (-Ae_2 + Ae_3 + B_2)I_S + 2C_2u^* + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2).$$

Since $-2C_2 < 0$, we have

$$0 \geq \frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}$$

and thus

$$u^* = \max\left(0, \frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}\right)$$

holds on this set.

3. On the set $\{t|u^*(t) = u_{\max}\}$, we have

$$0 \geq \left.\frac{\partial H}{\partial u}\right|_{u^*}$$
$$= (-Ae_2 + Ae_3 + B_2)I_S + 2C_2u_{\max} + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)$$

or equivalently,

$$-2C_2u_{\max} \geq (-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2).$$

Dividing both sides by the negative quantity $-2C_2$ we have

$$u_{\max} \leq \frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}$$

and thus

$$u^* = \min\left(\frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}, u_{\max}\right)$$

holds on this set.

After examining these three cases, we characterize the optimal rehydration control as

$$u^*(t) = \max\left(0, \min\left(\frac{(-Ae_2 + Ae_3 + B_2)I_S + I_S\lambda_{I_S}(e_2 - e_3 + \gamma_2 - \gamma_3) + \lambda_R I_S(\gamma_3 - \gamma_2)}{-2C_2}, u_{\max}\right)\right).$$

Similar three-step arguments hold for characterizing the optimal vaccination schedule, v^*, and the optimal sanitation schedule, m^*.

We also note that $\frac{\partial^2 H}{\partial u^2} \mid_{u^*} = 2C_2 > 0$, $\frac{\partial^2 H}{\partial v^2} \mid_{v^*} = 2C_3 > 0$, and $\frac{\partial^2 H}{\partial m^2} \mid_{m^*} = 2C_1 > 0$, indicating that the optimal controls minimize the Hamiltonian. □

3.4 Population-Specific Strategies of Cholera Control

In this section, we use our model to project cholera outbreaks for two endemic populations with diverse parameter values. In all numerical simulations, we use an iterative scheme employing a Runge Kutta method of the fourth order. See Chapter 1 and Appendix A for details. Parameter values for the two populations are presented in Section 3.4.1. Outbreak dynamics in the absence of control are displayed and compared in Section 3.4.2. For both populations, optimal strategies of multiple controls are computed and displayed in Section 3.4.3. We also investigate optimal strategies of one control and present results in Section 3.4.4. Lastly, we conduct a sensitivity analysis of our model with constant control parameters in Section 3.4.5.

3.4.1 Parameter Values

For population-specific parameters, we refer to King *et al.* [42] and the log-likelihood estimates for the two-path model. Displayed in Table B.1, estimates are based on monthly cholera mortality data from Bengal districts in the early 1900's. The effects of vaccination, antibiotics, and modern sanitation are not evident in the values. For our two populations, we choose the Calcutta district and the Bogra district primarily because of their diverse values of p, the proportion of infections which are asymptomatic. In Bogra, it is estimated that 76% of infections are asymptomatic. In contrast, Calcutta estimates indicate nearly 99% of infections are asymptomatic. The two districts also differ considerably in their estimated rates of waning immunity, cholera-related death, and cholera recovery.

For parameter values concerning the vibrios, we use estimates consistent with Hartley *et al.* [37] and Codeço [20]. The half saturation constant (or infectious dose ID_{50}) for non-HI vibrios is estimated to be $K_L = 10^6$ cells/ml. Laboratory experiments demonstrate that when inoculated into the intestines of mice, freshly shed *Vibrio cholerae* greatly outcompete bacteria grown in vitro, by as much as 700-fold [55]. Thus, for HI vibrios, it is estimated that $K_H = \frac{10^6}{700}$ cells/ml [37]. The HI advantage is temporary. Average time spent in the HI state is estimated to be $\frac{1}{\chi} \approx 5$ hours. Average lifespan of the non-HI vibrios is estimated to be $\frac{1}{\delta} \approx 30$ days [41].

All individuals with cholera actively shed vibrios [20, 50]. As found in a study by Levine *et al.*, symptomatic infecteds shed as much as 10^3 more vibrios than asymptomatic infecteds during their infectious period [50]. Based on the findings in this study and the results of a sensitivity analysis for our model, we let $\eta_1 = 0.5 \frac{\text{cells}}{\text{ml-day-human}}$ and $\eta_2 = 50 \frac{\text{cells}}{\text{ml-day-human}}$ as the vibrio shedding rates for asymptomatic and symptomatic infecteds respectively.

Ingestion of HI vibrios occurs at the rate β_H while ingestion of non-HI vibrios occurs at the rate β_L. Little is known about exact values of β_L and β_H, but it is clear that they play an important role in controlling cholera. Consistent with Hartley *et al.* [37], we assume $\beta_L = 0.215$ per day in both populations and allow β_H to vary. Figure 3.2 illustrates how varying the values of β_H can impact the explosiveness of the outbreak. Because the HI

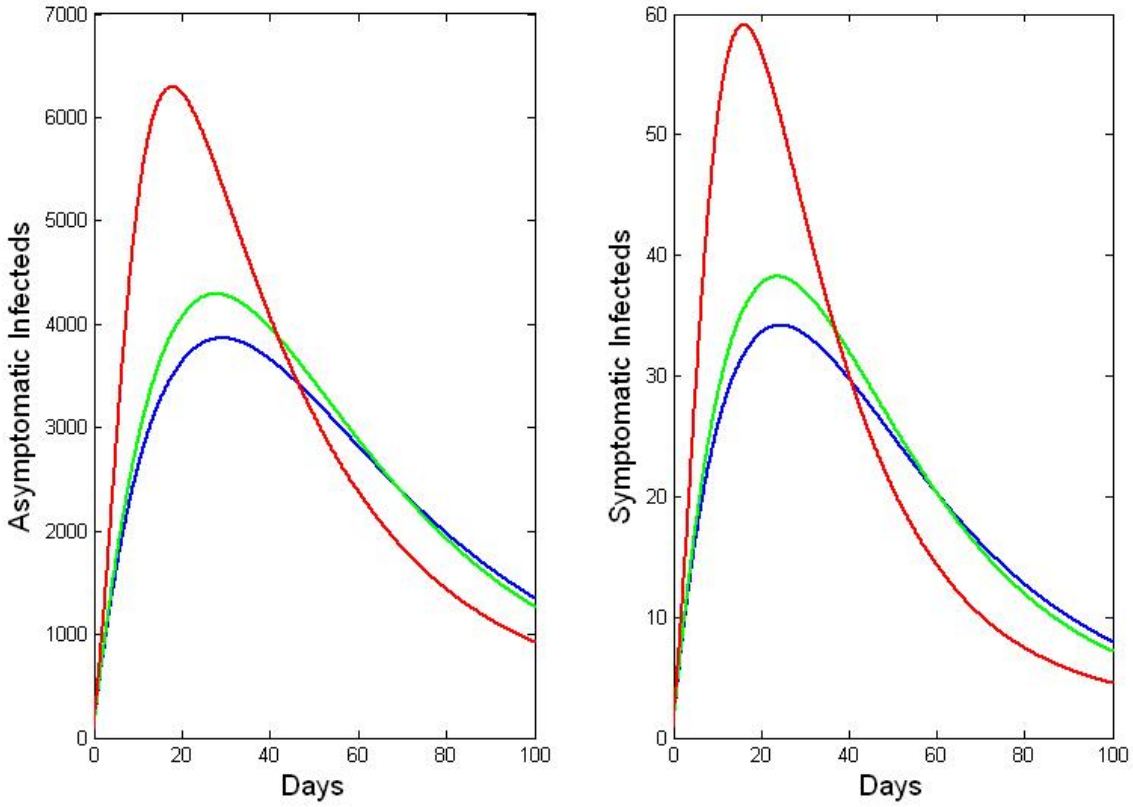


Figure 3.2: Simulation of three cholera outbreaks using system (3.1)-(3.6). All parameter values remain the same between simulations with the exception of β_H, the ingestion rate of HI vibrios. Red curves indicate an explosive outbreak resulting from $\beta_H = \beta_L$. Such a relationship may be representative of a a highly populated city or refugee camp in which human-to-human transmission is frequent. The green curves represent the outbreak for the decreased value $\beta_H = \frac{\beta_L}{5}$. Lastly, the blue curves represent the least explosive outbreak occurring when $\beta_H = \frac{\beta_L}{10}$ and may be representative of a rural community in which human-to-human transmission plays a lesser role in disease spread. Here, the environment-to-human route is dominant.

vibrio state is short-lived, the ingestion rate of HI vibrios has been associated with human-to-human transmission of cholera [37]. Areas more densely populated are likely to have higher rates β_H than areas less densely populated.

The district of Calcutta (currently renamed Kolkata) is a small district containing the city, also named Calcutta, and its surrounding area. The city of Calcutta served as the the capital of West Bengal during the early 1900's. During this time, its population exceeded 1 million people [70]. Presently, it is the third largest Indian city. In contrast, the district of Bogra is much larger in land area and less populated. The 1901 census revealed the total population of the Bogra district was less than 900,000 and no town within the district exceeded 10,000 inhabitants [12]. Because of the difference in the two districts' population densities, we assume different amounts of human-to-human transmission and choose the values of β_H accordingly. We let $\beta_H = \beta_L$ for the Calcutta population and $\beta_H = \frac{\beta_L}{10}$ for the Bogra population.

Values of parameters associated with the controls are chosen to reflect the appropriate consequences [91]. That is, a symptomatic infected receiving rehydration treatment experiences a 90% decrease in the rate of death due to disease, i.e. $e_3 = 0.1e_2$. These individuals also experience an increased recovery rate such that the average time spent in the I_S class is halved, i.e. $\gamma_3 = 2\gamma_2$.

Displayed in Table 3.5 are the parameter values used in all numerical simulation for the Bogra and Calcutta populations.

3.4.2 Cholera Outbreaks in the Absence of Intervention

Equations (3.1)-(3.6) with population-specific parameter values in Table 3.5 are used to simulate cholera outbreak dynamics in the absence of intervention. At the initial time, the two populations each consist of 100 infected individuals and 10,000 susceptibles individuals. The initial ratio of asymptomatic to symptomatic infecteds is chosen to reflect the population-specific value of p. The initial HI and non-HI vibrio concentrations are $\frac{1}{10}\kappa_H$ cells/ml and $\frac{1}{10}\kappa_L$ cells/ml respectively. Initial positive concentration of vibrios are indicative of endemic areas such as the Calcutta and Bogra districts.

Table 3.5: Population-specific parameter values

Parameter		**Bogra**	**Calcutta**	**Ref.**
Initial number of susceptibles	S_0	10,000	10,000	
Initial number of asymptomatic infecteds	I_{A0}	76	99	
Initial number of symptomatic infecteds	I_{S0}	24	1	
Initial number of recovered individuals	R_0	0	0	
Initial concentration of HI vibrios	B_{H0}	$\frac{\kappa_H}{10}$	$\frac{\kappa_H}{10}$	
Initial concentration of Non-HI vibrios	B_{L0}	$\frac{\kappa_L}{10}$	$\frac{\kappa_L}{10}$	
Proportion of infections being asymptomatic	p	0.76	0.9895	[42]
Ingestion rate of non-HI vibrios	β_L	0.215 day^{-1}	0.215 day^{-1}	[20]
Ingestion rate of non-HI vibrios	β_H	0.0215 day^{-1}	0.215 day^{-1}	
Non-HI saturation constant	κ_L	10^6 cells/ml	10^6 cells/ml	[37]
HI saturation constant	κ_H	$\frac{10^6}{700}$ cells/ml	$\frac{10^6}{700}$ cells/ml	[37]
Cholera-related death rate (asymptomatic)	e_1	0 day^{-1}	0 day^{-1}	[42]
Cholera-related death rate (symptomatic)	e_2	0.00065753 day^{-1}	.0128 day^{-1}	[42]
Cholera-related death rate (symptomatic with treatment)	e_3	0.000065753 day^{-1}	.00128 day^{-1}	
Cholera recovery rate (asymptomatic)	γ_1	0.15 day^{-1}	0.15 day^{-1}	[38]
Cholera recovery rate (symptomatic)	γ_2	0.0592 day^{-1}	.0310 day^{-1}	[42]
Cholera recovery rate (symptomatic with treatment)	γ_3	0.1184 day^{-1}	.0620 day^{-1}	
Rate of waning immunity	ω	0.0011 day^{-1}	.00192 day^{-1}	[42]
Vibrio shedding rate (asymptomatic)	η_1	0.5 $\frac{\text{cells}}{\text{ml-day-human}}$	0.5 $\frac{\text{cells}}{\text{ml-day-human}}$	[50]
Vibrio shedding rate (symptomatic)	η_2	50 $\frac{\text{cells}}{\text{ml-day-human}}$	50 $\frac{\text{cells}}{\text{ml-day-human}}$	[50]
Vibrio transition rate	χ	5 day^{-1}	5 day^{-1}	[37]
Vibrio death rate	δ	$\frac{1}{30}$ day^{-1}	$\frac{1}{30}$ day^{-1}	[37]

Displayed in Figure 3.3 are the susceptible and recovered human classes during an outbreak. Figure 3.4 displays the two infected classes and Figure 3.5 displays the vibrio concentrations. Cholera sweeps through both populations within weeks of introduction. Our model projects an approximate 2:1 ratio of asymptomatic to symptomatic infections at the peak of the outbreak for the Bogra population, while the Calcutta population represents one with fewer severe infections and an approximate 50:1 ratio.

At their maximum values on days 18 and 23 respectively, the Bogra asymptomatic infected class contains 1677 individuals and the symptomatic infected class contains 986 individuals. The HI and non-HI classes reach their maximum concentrations on days 23 and 42 with 10,011 cells/ml and 894,600 cells/ml respectively. For the Calcutta population, the class of asymptomatic infecteds reaches its maximum on day 14 with 2428 individuals. The maximum number of symptomatic infecteds is 52 and attained on day 21. The HI and non-HI vibrio concentrations attain maximums on days 17 and 28 with concentrations of 740 cells/ml and 93,181 cells/ml respectively. Despite the different dynamics, the estimates indicate the two populations are affected similarly in terms of cholera-related deaths. During the 100-day period, the model projects a total of 28 deaths for the Bogra population and 31 deaths for the Calcutta population.

3.4.3 Optimal Strategies of Multiple Controls

The cost-effective balance of multiple controls can differ between two populations. To illustrate, cost coefficients are fixed within objective functional (3.17) and the optimal schedule of the three controls is simulated for the Bogra population as well as the Calcutta population.

We bound $\mathbf{u}(t)$ above so that we effectively treat no more than 70% of symptomatic infections at any time. Daily vaccination rate is bounded above so that no more than 5% of the combined susceptible and aymptomatic infected classes are vaccinated per day. Sanitation efforts are modeled as a reduction in the population's bacterial ingestion rates and are bounded above so that a 40% reduction is maximal.

The cost of cholera vaccine was estimated to range between US \$0.50 - \$1.00 per dose in 1995 [60]. More recently, trials of a cholera vaccine manufactured in Vietnam at a cost

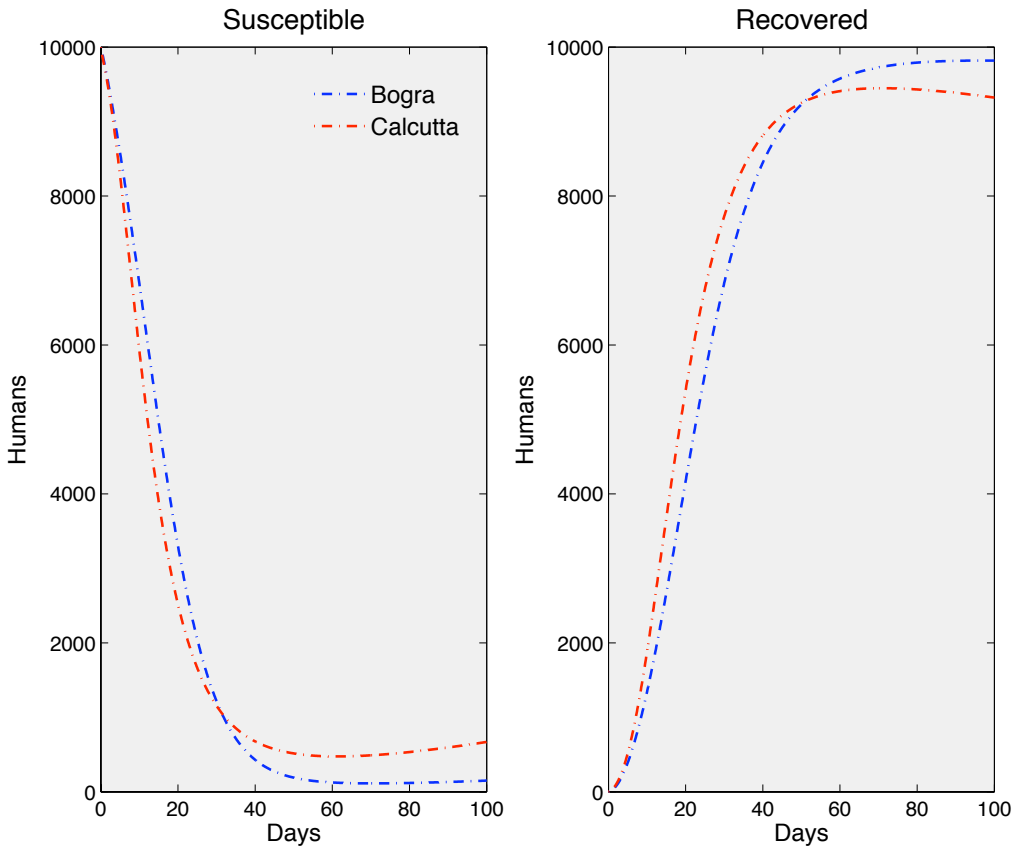


Figure 3.3: Susceptible and recovered human classes during cholera outbreak using system (3.1)-(3.6). Curves reflect outbreak dynamics in the absence of intervention.

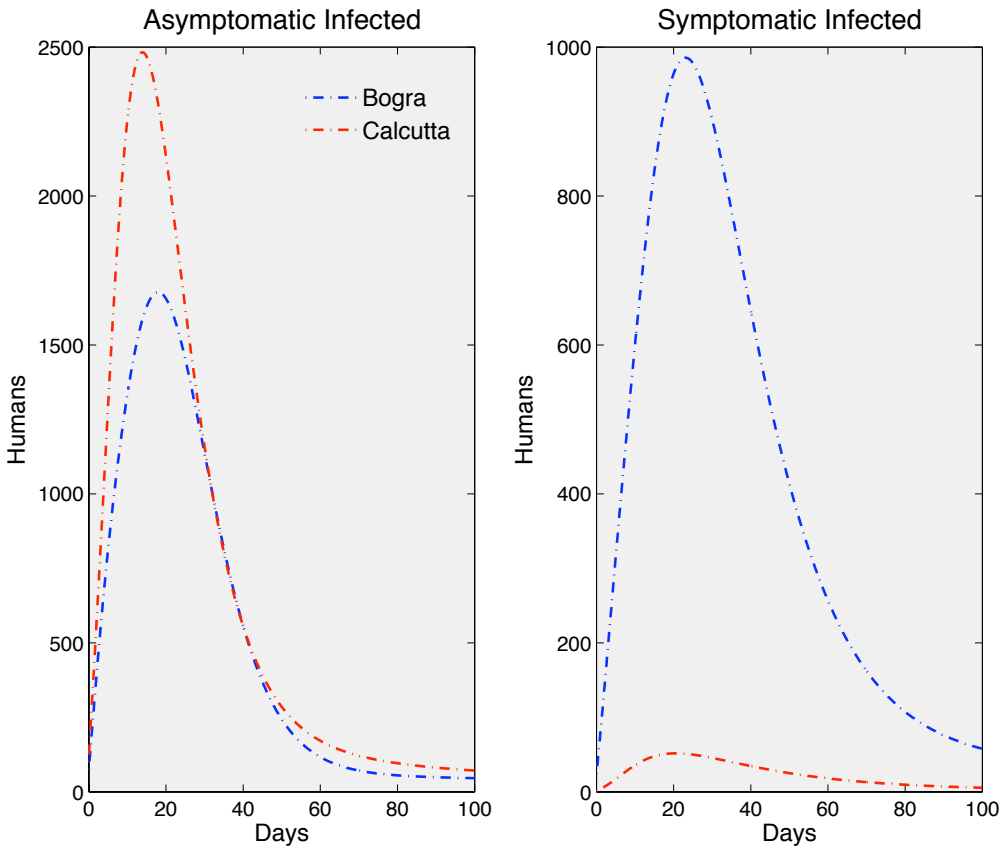


Figure 3.4: Asymptomatic and symptomatic infected human classes during cholera outbreak using system (3.1)-(3.6).

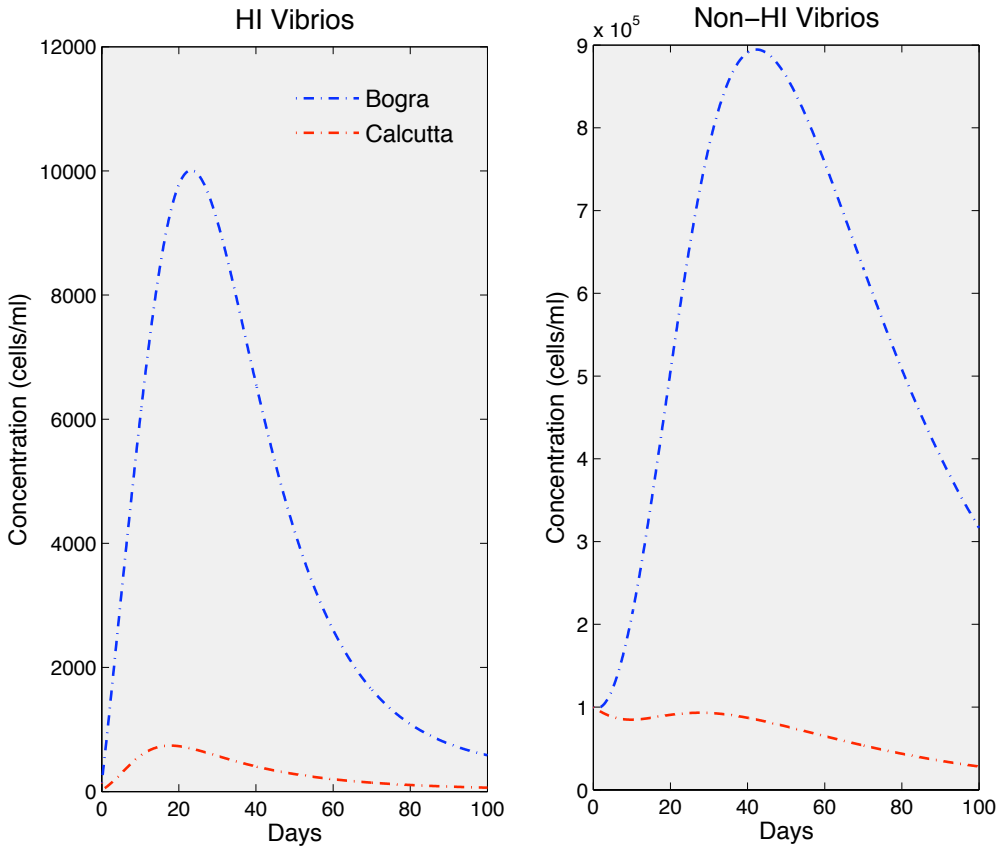


Figure 3.5: Hyperinfectious and less-infectious vibrio concentrations during cholera outbreak using system (3.1)-(3.6).

of only about US $0.20 per dose have produced encouraging results. Both require multiple doses for an effective vaccination. These estimates are for the vaccine itself and do not include any cost for facilities usage or professional administration of the drug. We fix the cost coefficient for the linear component within the vaccine term as $B_3 = \$6.00$. The cost coefficient for the linear rehydration and antibiotic component is fixed as $B_2 = \$2.00$, implying a cost of $2.00 per day for one symptomatic infected to receive professional cholera treatment. Thus, it is assumed that after 3 days of rehydration and antibiotic treatment, one symptomatic infected costs more to treat than it would cost to have vaccinated him or her prior to infection. Estimating the cost of sanitation efforts is not as straightforward. Sanitation effort may be in the form of increasing public awareness and education or providing families with saris and demonstrations of filtering. It is clear that the cost per unit of effort in reducing transmission in the total population is much greater than the per person cost of rehydration and antibiotic treatment or vaccination. The linear coefficient of sanitation cost is fixed as $B_1 = \$100.00$. The quadratic cost coefficients for each of the three controls are assumed equal, i.e., $C_1 = C_2 = C_3$. In the cost representations of each control within the objective functional, the linear terms dominate while quadratic terms add small additional costs associated with higher levels of treatment. The cost of human death, A, is fixed large enough to ensure minimizing death due to disease is a primary goal. The fixed cost coefficients and control bounds used in all simulations with control are displayed in Table 3.6.

Given the fixed costs in Table 3.6 and population-specific parameters in Table 3.5, the optimal balance of the three controls for the Bogra population (blue) and Calcutta population (red) are displayed in Figure 3.6. The corresponding human population dynamics are displayed as solid lines in Figure 3.7 and Figure 3.8. Solid lines in Figure 3.9 illustrate the corresponding vibrio concentrations. In all figures, dotted lines recall dynamics in absence of control.

Optimal control results provide clearly different strategies for relative application of sanitation efforts and treatment of symptomatic infecteds for the two populations. For the Bogra population, the optimal balance of rehydration and antibiotic treatment starts at the introduction of disease with two individuals receiving treatment, then increases according

Table 3.6: Fixed cost coefficients and control bounds

Parameter	**Notation**	**Value**
Cost coefficient (Infecteds)	A	3000 dollars per human
Cost coefficient (Sanitation)	B_1	100.00 dollars per unit effort
Cost coefficient (Rehydration and Antibiotics)	B_2	2.00 dollars per human
Cost coefficient (Vaccine)	B_3	6.00 dollars per human/day
Quadratic cost coefficient (Sanitation)	C_1	10 dollars per effort^2
Quadratic cost coefficient (Rehydration and Antibiotics)	C_2	10 dollars
Quadratic cost coefficient (Vaccine)	C_3	10 dollars per 1/day^2
Maximum rate of vaccination	v_{max}	0.05 day^{-1}
Maximum proportion of symptomatic infecteds receiving rehydration and antibiotic treatment	u_{max}	0.7
Maximum sanitation effort	m_{max}	0.4 unit effort

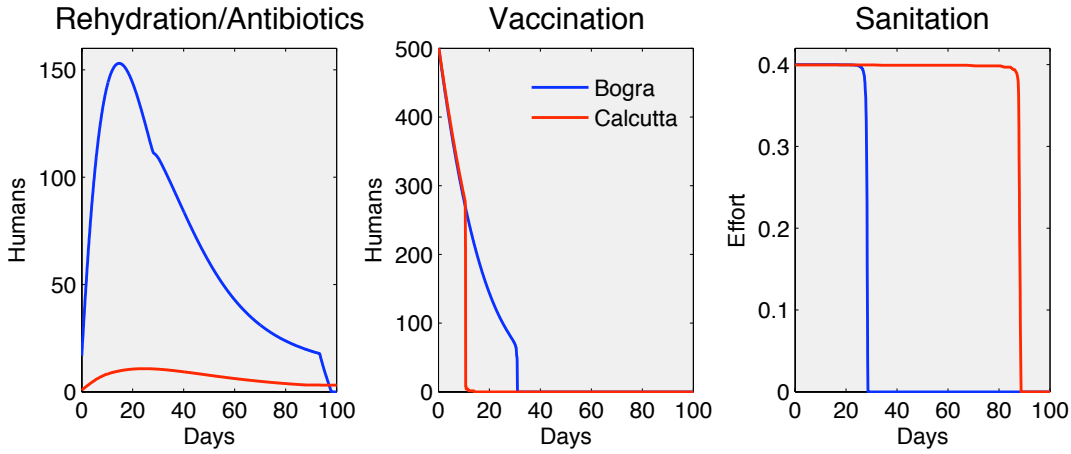


Figure 3.6: After fixing cost coefficients, the optimal balance of the three controls for the two diverse populations, Bogra (Blue) and Calcutta (Red) are displayed. Results indicate clearly different strategies for the relative application of sanitation efforts and treatment of symptomatic infecteds, but similar high levels of initial vaccination.

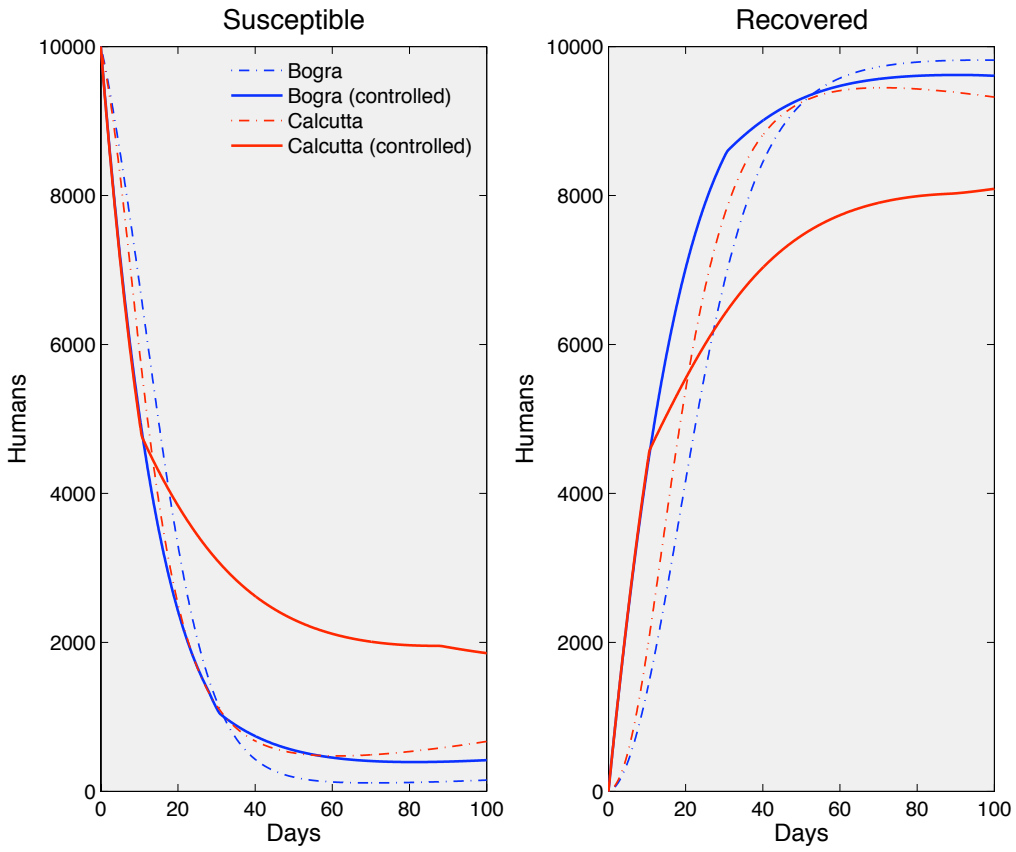


Figure 3.7: Solid lines indicate susceptible and recovered human classes during cholera outbreak using system (3.10)-(3.15) and given the population-specific optimal strategy of multiple controls in Figure 3.6. Dotted lines indicate population dynamics in absence of control.

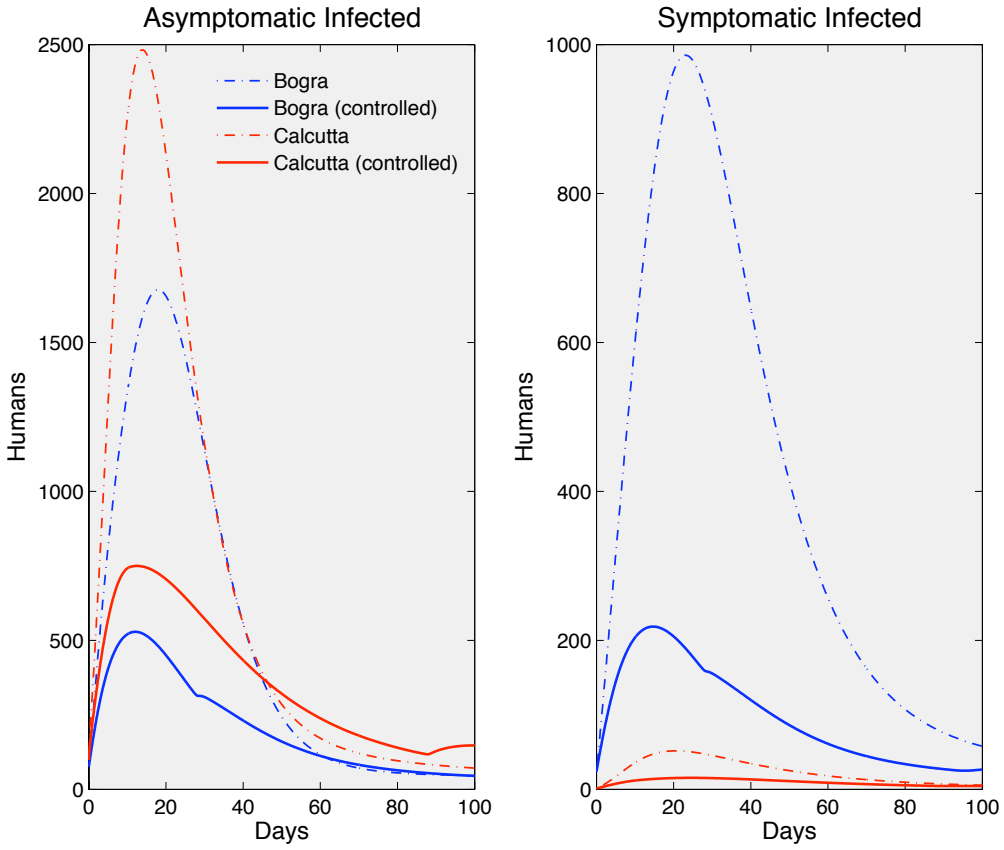


Figure 3.8: Solid lines indicate asymptomatic and symptomatic infected human classes during cholera outbreak using system (3.10)-(3.15) and population-specific optimal strategy of multiple controls in Figure 3.6. Dotted lines indicate population dynamics in absence of control.

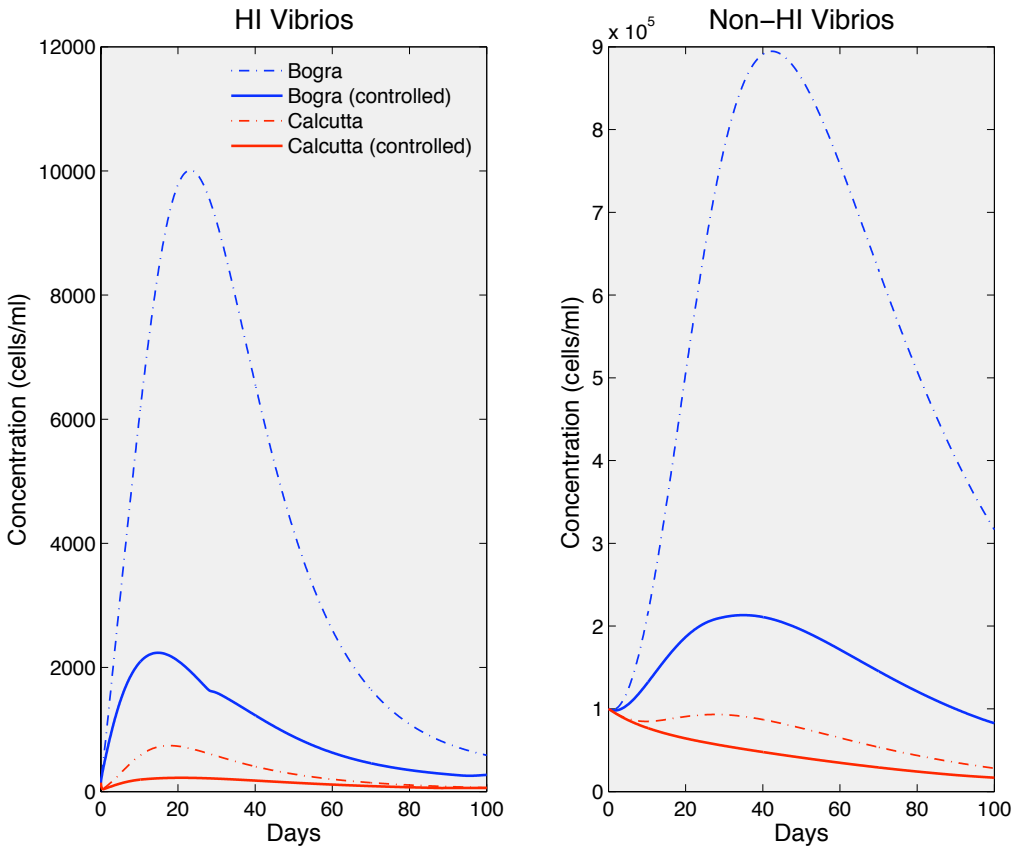


Figure 3.9: Solid lines indicate hyperinfectious and less-infectious vibrio concentrations using system (3.10)-(3.15) and population-specific optimal strategy of multiple controls in Figure 3.6. Dotted lines indicate population dynamics in absence of control.

to the increase in symptomatic infecteds. At its maximum quantity on day 14, a total of 150 symptomatic infecteds receive rehydration and antibiotic treatment. Afterwards, a decline in the number treated occurs and this particular control is ended on day 93. Consequentially, fewer funds are allocated for sanitation efforts in the optimal scheme; however, temporary maximum sanitation treatment is advantageous at the onset of infection until day 28. In combination with other controls, high levels of sanitation are most beneficial at the beginning of an outbreak to decrease the explosiveness of infection, providing more time to effectively implement the other treatments.

For the Calcutta population, far fewer rehydration and antibiotic supplies are necessary and instead, results suggest funding maximum sanitation for nearly the entire duration of infection. The optimal balance of rehydration and antibiotic treatment begins with one symptomatic infected at the start of an outbreak and increases until day 19, at which time 11 symptomatic infecteds receive treatment. A decline in the number of individuals receiving treatment occurs after day 11. After day 85, the number of symptomatic infecteds receiving treatment remains steadily at four for the remaining time. Sanitation efforts are applied at the maximum level initially and continue at this level for 85 days. No gradual decline in effort is visible and the optimal schedule suggests terminating sanitation efforts on day 85.

The optimal control strategies show that vaccinating at maximum rates initially is optimal in preventing deaths regardless of the populations ratio of asymptomatic to symptomatic infections. Bogra and Calcutta results both indicate that vaccination is crucial to apply during the first few weeks of disease detection. Vaccination of 500 individuals per day is initially optimal for both populations, but followed by a decreasing daily rate. For the Bogra population, the decline is such that on day 31, 150 individuals are vaccinated per day. The vaccination regime continues until termination on day 37 at which time 100 individuals per day are vaccinated. For the Calcutta population, a sharper decline is indicated due to the decreasing size of the susceptible class. The decline is such that on day 11 the optimal vaccination rate is 275 individuals per day. After day 11, no further vaccination is optimal.

Our model indicates similar and significant reductions in both populations' death toll and infectious classes given the optimal intervention. For the Bogra population, cholera

related deaths are reduced reduced by 89% from 28 to 3 deaths over the 100-day period. With the optimal intervention, the Bogra asymptomatic infected class attains its maximum on day 12 with 528 individuals and symptomatic class on day 15 with 218 individuals. These numbers indicate a 68% and 77% reduction in the maximum number of asymptomatic and symptomatic infecteds, respectively.

The optimal intervention yields 5 cholera-related deaths over the 100-day period for the Calcutta population, an 84% reduction from the projected 31 deaths in the absence of intervention. With the intervention, the Calcutta outbreak is less severe with a maximum number of 750 asymptomatic infecteds and 15 symptomatic infecteds on days 13 and 25 respectively. These values indicate a 66% and 60% reduction in maximum number of symptomatic and asymptomatic infecteds, respectively.

Both populations, with their respective optimal strategy, experience similar, significant reductions in vibrio concentrations because of the reduced number of infected individuals. By the final time, an increased supply of susceptible individuals is visible in both populations due to the controls. The susceptible population is heightened more for Calcutta because of less movement from the susceptible class to the recovered class via vaccination. This is also reflected in the size of each population's recovered classes. By the final time, a decreased supply of recovered individuals exists in each population, more noticeably so in Calcutta.

Although the balancing coefficients in the objective functional are not meant to be accurate estimates of all costs associated with control implementation, we use the value of the objective functional to assess the relative costs of the two optimal strategies. For the optimal allocation of controls in the Bogra population, the objective functional in (3.17) is estimated at \$64,645.00 (\$2,585.80 per life saved). For the optimal application of controls in the Calcutta population, the objective functional in (3.17) is estimated at \$43,748.00 (\$1,682.60 per life saved). The increased cost of the Bogra strategy is largely due to the increased treatment of symptomatic infecteds.

3.4.4 Optimal Strategies of Single Control

To assess the benefit gained from strategies comprised of multiple controls, we simulate optimal applications of one control alone and the corresponding population dynamics. Assuming only one of the controls is feasible, we can reformulate our optimal control problem by setting the remaining two controls to identically zero in system (3.10)-(3.15) and objective functional (3.17). Fixing cost coefficients and control bounds as in Table 3.6, we simulate the optimal schedule of each intervention method. Here, we display results for the Calcutta population.

Figures 3.10 (a) - (c) display the optimal schedule of each control when applied in absence of the other two. In Figure 3.11, the asymptomatic and symptomatic populations corresponding to each of the optimal stratgies of one control are displayed. For comparison, we also display the corresponding population dynamics for the optimal strategy of all three controls.

Recall that in the absence of all controls, our model projects 31 cholera-related deaths for the Calcutta population over a 100-day period. Comparing the optimal strategies for each single control, the strategy comprised of only rehydration and antibiotic treatment is most effective in reducing the death toll. With the optimal treatment in Figure 3.10 (a), our model projects 10 cholera-related deaths. However, this optimal intervention does the least to reduce the total number infected during the outbreak. On the other hand, the initial high rate of vaccination in Figure 3.10 (b) effectively decreases the size of both infected classes while also modestly reducing the death toll. The optimal vaccination schedule reduces cholera-related deaths to 20 over the 100-day period. In the absence of the other two controls, sanitation does little to reduce the death due to cholera. Our model projects 28 deaths with the optimal schedule in Figure 3.10 (c). However, sanitation does dampen the explosiveness of the outbreak.

Although not displayed here, optimal applications of one control were also simulated for the Bogra population. For both populations, the optimal strategy balancing the three controls is considerably more effective in reducing both cholera-related death and total infections than any of the optimal strategies of one intervention method.

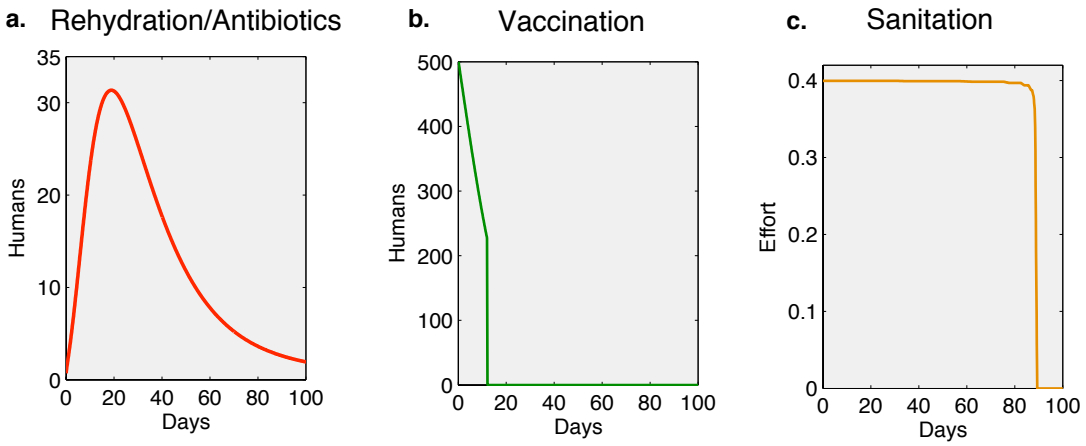


Figure 3.10: Optimal applications of one intervention method for Calcutta population. (a) In absence of sanitation and vaccination, the optimal number of symptomatic infecteds receiving rehydration and antibiotics treatment during the outbreak increases. (b) In absence of rehydration/antibiotic treatment and sanitation, initial high rates of vaccination remain optimal for a short time. (c) In absence of rehydration/antibiotic treatment and vaccination, maximum levels of sanitation throughout the outbreak period remain optimal.

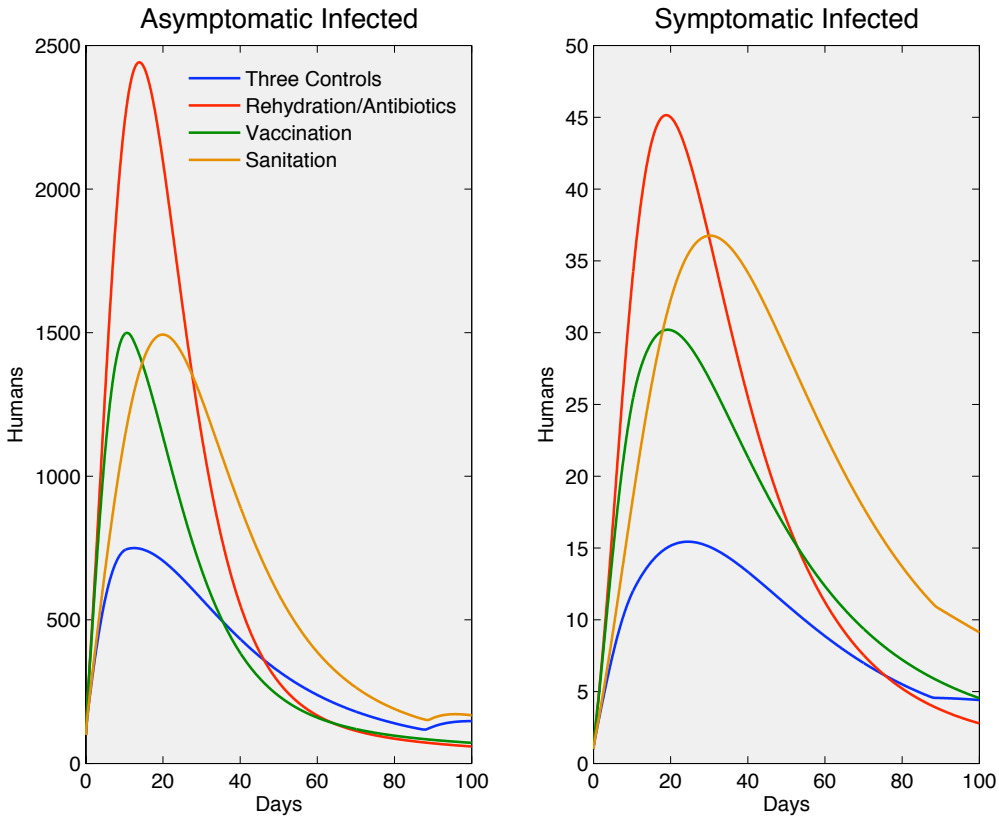


Figure 3.11: Dynamics of two infected classes for optimal strategies of one control and multiple controls. The optimal strategy balancing the the three controls is more effective in reducing both cholera-related death and total infections than any of the optimal strategies of one control.

3.4.5 Sensitivity Analysis of Control-Related Parameters

Lastly, a sensitivity analysis for each population is conducted using equations (3.10)-(3.15). For the analysis, we assume control functions are constant functions of time ($u(t) = u$, $v(t) = v$, $m(t) = m$) and compare the ability of parameters u, v, m, γ_3 to affect change in the three outcomes measuring outbreak severity. Over the total time period of the model, outbreak severity is again measured by (1) the maximum and (2) the total size of the symptomatic infected class and (3) the total size of the combined infected classes. The three most influential parameters from the sensitivity analysis in Section 3.2.4 are also included in the analysis. The remaining parameters are fixed according to the population-specific values given in Table 3.5. A summary of the parameters used in the LHS and the ranges over which they were sampled is displayed in Table 3.7.

Results of the sensitivity analysis are displayed in Table 3.8 for the Bogra population and in Table 3.9 for the Calcutta population. The results indicate that each of the three controls are key in reducing the severity of an outbreak, but population-specific parameters influence which controls are more important. The results suggest that, in both populations, vaccination is the most effective control in decreasing the number of individuals who become infected during an outbreak. Reducing the maximum number of symptomatic infecteds during an outbreak can be addressed with all three controls. Results for both districts attest that rehydration and antibiotic treatment can have the greatest impact of the three controls in reducing the maximum number of symptomatic infected individuals during an outbreak. For Calcutta, sanitation effort can strongly influence the total number of individuals who become infected during an outbreak, while in Bogra the effect of sanitation is marginal.

The population-specific results of the sensitivity analysis reiterate the key conclusion of our optimal control simulations. Strategies of multiple controls can have dramatic effects in reducing the severity of a cholera outbreak. However, the optimal balance of multiple intervention methods can vary between populations.

Table 3.7: Parameters used in sensitivity analysis with controls. Upper and lower bounds for sampled ranges are displayed. All values are in units day^{-1}.

	Description	**Lower**	**Upper**
r	Multiplicative factor determining β_H, where $\beta_H = r\beta_L$	0.01	1.0
η_1	Vibrio shedding rate for asymptomatic infected	0.05	1.0
s	Multiplicative factor determining η_2, where $\eta_2 = s\eta_1$	1	100
m	Constant level of sanitation effort	0.001	0.40
u	Constant proportion of symptomatic infecteds receiving rehydration and antibiotic treatment	0.001	0.70
v	Constant rate of vaccination	0.001	0.05
γ_3	Recovery rate for symptomatic infected with treatment	0.05	0.3

Table 3.8: Sensitivity analysis results for Bogra. For each parameter, the PRCC corresponding to each of the three measures of outbreak severity is displayed. Statistical significance of associated p-value is indicated with astericks (* p-value < 0.05, ** p-value < 0.01, ***p-value < 0.001).

Parameter	**Maximum I_S**	**Total I_S**	**Total ($I_A + I_S$)**
r	0.85 ***	0.79 ***	0.79 ***
η_1	0.73 ***	0.74 ***	0.74 ***
s	0.74 ***	0.75 ***	0.75 ***
m	-0.52 ***	-0.46	-0.46 ***
u	-0.54 ***	-0.20 ***	-0.19 ***
v	-0.51 ***	-0.83 ***	-0.83 ***
γ_3	-0.61 ***	-0.21***	-0.21 ***

Table 3.9: Sensitivity analysis results for Calcutta. For each parameter, the PRCC corresponding to each of the three measures of outbreak severity is displayed. Statistical significance of associated p-value is indicated with astericks (* p-value < 0.05, ** p-value < 0.01, ***p-value < 0.001).

Parameter	**Maximum** I_S	**Total** I_S	**Total** $(I_A + I_S)$
r	0.71 ***	0.81 ***	0.81 ***
η_1	0.57 ***	0.80 ***	0.80 ***
s	0.51 ***	0.73 ***	0.73 ***
m	-0.43 ***	-0.63 ***	-0.64 ***
u	-0.78 ***	-0.36 ***	-0.37 ***
v	-0.27 ***	-0.85 ***	-0.86 ***
γ_3	-0.74 ***	-0.32 ***	-0.32 ***

3.5 Concluding Remarks and New Contributions

This work presents a new mathematical model for the spread of cholera that incorporates both mild and severe human infections, both hyperinfectious and less-infectious bacteria, and waning immunity. A new interpretation of the basic reproductive number arises from our model and provides a measure of all secondary infections (mild and severe).

Our work also presents the first application of optimal control theory to a cholera model. The model formulation, sensitivity analyses, optimal control problem, and associated solution provide a new methodology to evaluate trade-offs in multiple intervention schemes and assess which are most effective. An important result of this analysis is that a cost-effective balance of prevention and treatment methods can successfully control a cholera outbreak, but that the balance of treatment is specific to the population. A one-size-fits-all approach is inappropriate. Strategies of multiple controls are shown to be more effective in reducing the death toll and severity of an outbreak than any strategy consisting of one control alone.

Optimal control theory in our cholera model is a starting point for more elaborate models which can include spatial and/or age dependence and can provide insight in determining which age groups or locales should be given priority for various treatments. Our results broaden the application of optimal control theory to disease dynamics and provide a framework for designing cost effective treatment strategies for diseases other than cholera.

Acknowledgements

This work was done in collaboration with Elsa Schaefer, Holly Gaff, Renee Fister, and Suzanne Lenhart. We acknowledge helpful discussions with L. J. Gross, D. Hartley, J. G. Morris, and P. van den Driessche. An article summarizing the results has been submitted for publication.

Chapter 4

Optimal Control of a Spatiotemporal Epidemic Model and Application to Rabies

4.1 Background

In the past, diffusive spatiotemporal models have been used to study the spread of epidemics among a population of susceptible (S), infected (I), and recovered (R) individuals. In 1981, Webb [88] analyzed the following coupled system of nonlinear PDE's with prescribed initial values and no-flux boundary values:

$$S_t(x,t) = S_{xx}(x,t) - aS(x,t)I(x,t), \qquad t \geq 0,\ -L \leq x \leq L, \tag{4.1}$$

$$I_t(x,t) = I_{xx}(x,t) + aS(x,t)I(x,t) - \lambda I(x,t), \tag{4.2}$$

$$R_t(x,t) = R_{xx}(x,t) + \lambda I(x,t), \tag{4.3}$$

$$S_x(\pm L,t) = I_x(\pm L,t) = R_x(\pm L,t) = 0, \qquad t \geq 0, \tag{4.4}$$

$$S(x,0) = S_0(x), \quad I(x,0) = I_0(x), \quad R(x,0) = R_0(x) \qquad -L \leq x \leq L. \tag{4.5}$$

The infection rate a and removal rate λ are given positive constants and initial populations $S_0(x), I_0(x), R_0(x)$ are assumed to be nonnegative and continuous on $[-L, L]$. Webb showed that for system (4.1)-(4.5) there exists a unique classical solution defined for all $t \geq 0$ and as $t \to \infty$, $S(x,t)$ converges uniformly in x to S_∞ and $I(x,t)$ converges uniformly in x to $I_\infty(x)$, where S_∞ is a positive constant function on $[-L, L]$ and $I_\infty = 0$ on $[-L, L]$. Thus, the infection dies out as $t \to \infty$, but not for lack of susceptible individuals, some of whom never contract the disease. In 1987, Fitzgibbon and Morgan [33] extended this model to bounded domains of arbitrary dimension and proved asymptotic results similar to that of Webb.

In this chapter of the dissertation we seek weak solutions [31] to epidemic models with reaction-diffusion equations that include transport effects and no-flux boundary conditions. In 2002, Bendahmane, *et al.* [9] proved the existence of at least one weak solution for a nonlinear reaction-diffusion system of partial differential equations with L^1 data and no-flux boundary conditions. We formulate an S-I-R model similar to that of Bendahmane, *et al.* with L^∞ data and include the effects of vaccinating as a control variable. After designing an objective functional, we use the principles of optimal control theory to characterize the optimal strategy of vaccination. We illustrate results with numerical approximations to the optimality system.

For the numerical results, parameter values are chosen to model the spread of rabies and raccoons. It was Murray *et al.* who first used partial differential equations to study the spatial spread of rabies within fox populations [40, 57, 58]. Using the model from [58], Evans and Pritchard [32] applied control of initial conditions in culling and quarantine to drive the population to a desired profile. Other extensions of Murray's work incorporate environmental and habitat heterogeneity [69, 78, 79]. Our goal is to investigate optimal vaccination regimes in both homogeneous and heterogeneous spatial domains and address how long distance raccoon translocation may affect optimal strategies.

4.2 Model Formulation and Optimal Control Problem

We formulate a spatiotemporal S-I-R model as follows. Let S, I, R represent either the density or size of the susceptible, infected, and immune classes of a population at a given location and time. Let Ω be an open and bounded subset of $\mathbf{R}^n$. Set $Q = \Omega \times (0,T)$ for some fixed time $T > 0$. Given a control $v(x,t)$, the corresponding state variables $(S,I,R) = (S,I,R)(v)$ satisfy the system

$$\begin{aligned} L_1 S &= b(S+R) - \mu_1 S - \beta SI - avS, && \text{for } (x,t) \in Q, \\ L_2 I &= \beta SI - \mu_2 I, \\ L_3 R &= -\mu_1 R + avS \end{aligned} \tag{4.6}$$

where the operators L_k, $k = 1,2,3$ are defined as

$$L_k u \equiv \frac{\partial u}{\partial t} - \sum_{i,j=1}^{n} (a_{ij}^k u_{x_i})_{x_j} + \sum_{i=1}^{n} (b_i^k u)_{x_i}.$$

Initial conditions and no-flux boundary conditions are given by

$$S(x,0) = S_0(x), \quad I(x,0) = I_0(x), \quad R(x,0) = R_0(x) \text{ for } x \in \Omega, \tag{4.7}$$

$$\frac{\partial S}{\partial \nu} = 0, \quad \frac{\partial I}{\partial \nu} = 0, \quad \frac{\partial R}{\partial \nu} = 0 \text{ on } \partial\Omega \times (0,T). \tag{4.8}$$

We assume only the S and R populations give birth. The parameter b is the birth rate, μ_1 is a natural death rate, and μ_2 is an increased death rate due to disease. The horizontal incidence term βIS represents the infection rate of susceptibles. We assume the simple mass action law applies with β as the mass action coefficient. The control function v represents the vaccination rate of the susceptible population. The rate at which susceptibles are

effectively vaccinated is assumed to be proportional to the total susceptible population size. The proportionality constant is taken to be av for some $a > 0$. The boundary conditions imply that the population does not diffuse across the boundary. Denote the outward normal vector on $\partial\Omega$ by η. Then ν is the conormal outward vector with components $\nu_i = \sum_{j=1}^{n} a_{ij}\eta_j$.

We define the class of admissible controls as

$$U = \{v \in L^\infty(Q) \mid v : Q \to [0, v_{\max}]\}$$

for some positive constant $v_{\max}$. The best strategy for controlling an epidemic outbreak during a time period of length T may be one which minimizes the number of individuals who become infected and the cost of vaccination during the time period. It may also be of importance to sustain a sizable susceptible population. Therefore, we seek to minimize the objective functional

$$J(v) = \int_Q (AI - BS + C(v))dxdt$$

where A, B are constant weights and $C(\cdot)$ is a lower semi-continuous convex function representing the cost of vaccination. For example, one may take $C(v) = c_1 v^2$ or $C(v) = c_2 v + c_3 v^2$ with positive constants c_1, c_2, and c_3.

Throughout this chapter we will make the following assumptions:

1. $a, v_{\max} \in \mathbf{R}_+$,
2. $S_0(x), I_0(x), R_0(x) \in L^\infty(\Omega)$ and $S_0(x) > 0, I_0(x) > 0, R_0(x) \geq 0$,
3. $b, \mu_1, \mu_2 \in L^\infty(Q)$ and $b \geq 0, \mu_1 \geq 0, \mu_2 \geq 0$,
4. $a_{ij}^k, b_i^k \in C^1(\bar{Q})$ and $a_{ij}^k = a_{ji}^k$ for $k = 1, 2, 3$,
5. $\sum_{i,j=1}^{n} a_{ij}^k \xi_i \xi_j \geq \theta \sum_{i=1}^{n} \xi_i^2, k = 1, 2, 3$, where $\theta > 0$.

Define the space $V = L^2(0, T; H^1(\Omega))$ and the bilinear forms

$$B^k[t,u,\phi] = \int_\Omega \sum_{i,j=1}^{n} a_{ij}^k(x,t) u_{x_i} \phi_{x_j} dxdt + \int_\Omega \sum_{i=1}^{n} (b_i^k(x,t)u)_{x_i} \phi dxdt$$

for $u, \phi \in V$ and $k = 1, 2, 3$.

Definition 8. For the system (4.6)-(4.8), a weak solution $(S, I, R) \in V^3$ with $S_t, I_t, R_t \in L^2(0,T; H^1(\Omega)^*)$ must be nonnegative and satisfy (4.7),(4.8) and the following variational formulation of (4.6)

$$\begin{aligned}
\int_0^T \langle S_t, \phi_1 \rangle dt + \int_0^T B^1[t,S,\phi_1]dt &= \int_Q b(S+R)\phi_1 dxdt - \int_Q \beta SI\phi_1 dxdt - \int_Q (\mu_1 - av)S\phi_1 dxdt \\
\int_0^T \langle I_t, \phi_2 \rangle dt + \int_0^T B^2[t,I,\phi_2]dt &= \int_Q \beta SI\phi_2 dxdt - \int_Q \mu_2 I\phi_2 dxdt \\
\int_0^T \langle R_t, \phi_3 \rangle dt + \int_0^T B^3[t,R,\phi_3]dt &= \int_Q avS\phi_3 dxdt - \int_Q \mu_1 R\phi_3 dxdt
\end{aligned} \tag{4.9}$$

for all $\phi_1, \phi_2, \phi_3 \in V$ where $\langle\ ,\ \rangle$ inner product is the duality between $H^1(\Omega)^*$ and $H^1(\Omega)$.

Remark 9. Due to results in Evans [31], $S, I, R \in C([0,T]; L^2(\Omega))$. Thus, initial conditions $S_0(x), I_0(x), R_0(x) \in L^\infty(\Omega)$ are both biologically and mathematically appropriate.

Remark 10. To make sense of terms such as $\int_Q \beta SI\phi dxdt$ in (4.9), we seek $L^\infty(Q)$ bounds on the state variables in Theorem 11.

Throughout the following sections of results, similar estimation techniques are used within the proofs. Because of their frequency, we display them here and reference them as necessary. Four useful inequalities and equalities follow.

1. For $u \in V$ and $s \in (0,T]$, we have

$$\begin{aligned}
\int_{\Omega\times(0,s)} u_t u dxdt &= \frac{1}{2}\int_{\Omega\times(0,s)} (u^2)_t dxdt \\
&= \frac{1}{2}\int_\Omega u^2(x,s) - u^2(x,0)dx.
\end{aligned} \tag{4.10}$$

2. For $u, v \in V$ Cauchy's Inequality with $\epsilon > 0$ states

$$\int_Q uvdxdt \leq \epsilon \int_Q u^2 dxdt + \frac{1}{4\epsilon} \int_Q v^2 dxdt. \tag{4.11}$$

3. For $k = 1, 2, 3$ and $u \in V$, an application of Cauchy's Inequality with ϵ gives

$$|\int_Q \sum_{i=1}^{n} (b_i^k u)_{x_i} u| dxdt \leq \frac{\theta}{2} \int_Q |\bigtriangledown u|^2 dxdt + C \int_Q |u|^2 dxdt \tag{4.12}$$

where C depends on θ and the bounds on the coefficients b_i^k and $(b_i^k)_{x_i}$.

4. For $u, \bar{u}, v, \bar{v} \in V$, we can write

$$vu - \bar{v}\bar{u} = v(u - \bar{u}) + \bar{u}(v - \bar{v}). \tag{4.13}$$

4.3 Existence and Optimality System

4.3.1 Existence of State Solution and Optimal Control

We start by proving existence of a solution to the state system given a control $v \in U$. Then we will show that an optimal control does exist.

Theorem 11. *For T sufficiently small, given $v(x,t) \in U$, there exists a unique nonnegative solution $(S, I, R) \in V^3$ satisfying (4.7), (4.8), and (4.9). Furthermore, $0 \leq S(x,t) \leq C$, $0 \leq I(x,t) \leq C$, and $0 \leq R(x,t) \leq C$ a.e. $(x,t) \in Q$ for some constant C.*

Proof: We use Banach's fixed point theorem which states that if given a Banach space X and a nonlinear mapping $A : X \to X$ such that $\|A(u) - A(\hat{u})\| \leq \gamma \|u - \hat{u}\|$ for all u,$\hat{u} \in X$ and some $\gamma < 1$ then A has a unique fixed point.

For a solution $(S, I, R) \in V^3$, results in Evans [31] indicate $S \in C([0,T]; L^2(\Omega))$, $I \in C([0,T]; L^2(\Omega))$, and $R \in C([0,T]; L^2(\Omega))$. In addition, we desire $L^\infty(Q)$ bounds on S, I, and R. Because of this, we apply Banach's fixed point theorem in the Banach space X^3 where

$$X = C([0,T];L^2(\Omega)) \bigcap \{u \in L^\infty(Q) | 0 \le u \le M \text{ a.e.} (x,t) \in Q\}$$

for fixed

$$M \ge 2 \max\{\|S_0\|_{L^\infty(\Omega)}, \|I_0\|_{L^\infty(\Omega)}, \|R_0\|_{L^\infty(\Omega)}\}.$$

We use the norm

$$\|u_1, u_2, u_3\|_{X^3} = (\|u_1\|_X + \|u_2\|_X + \|u_3\|_X)$$

where

$$\|u\|_X = \sup_{0 \le t \le T} \|u(t)\|_{L^2(\Omega)}.$$

Assume $v \in U$. We transform our state equations by a change of functions. Let $w = e^{-\lambda t}S$, $y = e^{-\lambda t}I$, and $z = e^{-\lambda t}R$. The equations satisfied by these functions are

$$\begin{aligned} \hat{L_1} w &= b(w+z) - \mu_1 w - e^{\lambda t}\beta w y - avw, \\ \hat{L_2} y &= e^{\lambda t}\beta w y - \mu_2 y, \\ \hat{L_3} z &= -\mu_1 z + avw \qquad\qquad \text{a.e. } (x,t) \in Q \end{aligned} \tag{4.14}$$

where

$$\hat{L_k} u \equiv \frac{\partial u}{\partial t} - \sum_{i,j=1}^{n} (a_{ij}^k u_{x_i})_{x_j} + \sum_{i=1}^{n} b_i^k u_{x_i} + \left(\sum_{i=1}^{n} (b_i^k)_{x_i} + \lambda \right) u$$

for $k = 1, 2, 3$. The system has initial conditions

$$w(x,0) = S_0(x), \qquad y(x,0) = I_0(x), \qquad z(x,0) = R_0(x) \qquad \text{for x} \in \Omega \tag{4.15}$$

and boundary conditions

$$\frac{\partial w}{\partial \nu}(x,t) = 0, \qquad \frac{\partial y}{\partial \nu}(x,t) = 0, \qquad \frac{\partial z}{\partial \nu}(x,t) = 0 \qquad \text{for all } x \in \partial\Omega, t \in (0,T). \tag{4.16}$$

Let $(h_1, h_2, h_3) \in X^3$. Consider the system of linear parabolic PDEs

$$\begin{aligned} \hat{L_1} w &= b(w + h_3) - \mu_1 w - e^{\lambda t}\beta w h_2 - avw, \\ \hat{L_2} y &= e^{\lambda t}\beta h_1 y - \mu_2 y, \\ \hat{L_3} z &= -\mu_1 z + avh_1 \qquad\qquad \text{a.e. } (x,t) \in Q \end{aligned} \tag{4.17}$$

with initial conditions

$$w(x,0) = S_0(x), \qquad y(x,0) = I_0(x), \qquad z(x,0) = R_0(x) \qquad \text{for x} \in \Omega \tag{4.18}$$

and boundary conditions

$$\frac{\partial w}{\partial \nu}(x,t) = 0, \qquad \frac{\partial y}{\partial \nu}(x,t) = 0, \qquad \frac{\partial z}{\partial \nu}(x,t) = 0 \qquad \text{for all } x \in \partial\Omega,\ t \in (0,T). \tag{4.19}$$

Given $(h_1, h_2, h_3) \in X^3$, the system of linear equations (4.17)-(4.19) has a unique weak solution solution $(w, y, z) \in V^3$ [31]. We will show that the solution $(w, y, z) \in X^3$ as well. By results in Evans [31], we know $w \in C([0,T]; L^2(\Omega))$, $y \in C([0,T]; L^2(\Omega))$, and $z \in C([0,T]; L^2(\Omega))$. We must show that $0 \leq w(x,t) \leq M$, $0 \leq y(x,t) \leq M$, and $0 \leq z(x,t) \leq M$ a.e. $(x,t) \in Q$ to conclude $(w, y, z) \in X^3$. First, we show the lower bounds

hold.

Rearranging (4.17) we find

$$\begin{aligned} \hat{L_1} w + (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)w &= bh_3 \geq 0, \\ \hat{L_2} y + (-e^{\lambda t}\beta h_1 + \mu_2)y &= 0, \\ \hat{L_3} z + \mu_1 z &= avh_1 \geq 0 \qquad \text{a.e. } (x,t) \in Q \end{aligned} \tag{4.20}$$

since $0 \leq b, 0 \leq av, 0 \leq h_1, 0 \leq h_3$ a.e. $(x,t) \in Q$. Noting that $(b_i^k)_{x_i}, b, \mu_1, \mu_2, \beta, av \in L^\infty(Q)$, we can first choose λ sufficiently large and then choose T sufficiently small so that

$$\begin{aligned} \lambda + \sum_{i=1}^{n} (b_i^1)_{x_i} - b + \mu_1 + e^{\lambda t}\beta h_2 + av &\geq 0, \\ \lambda + \sum_{i=1}^{n} (b_i^2)_{x_i} - e^{\lambda t}\beta h_1 + \mu_2 &\geq 0, \\ \lambda + \sum_{i=1}^{n} (b_i^3)_{x_i} + \mu_1 &\geq 0 \qquad \text{a.e. } (x,t) \in Q. \end{aligned} \tag{4.21}$$

Thus, the extension of the Parabolic Maximum Principle to weak solutions [46] gives us

$$\begin{aligned} &0 \leq w(x,t), \\ &0 \leq y(x,t), \\ &0 \leq z(x,t) \qquad \text{a.e. } (x,t) \in Q. \end{aligned}$$

Second, we show the upper bound holds. That is, $w(x,t) \leq M$, $y(x,t) \leq M$, and $z(x,t) \leq M$ a.e. $(x,t) \in Q$. There exists a positive constant $C = \hat{C}M$ such that $|bh_3| \leq C$, $|avh_1| \leq C$ a.e. $(x,t) \in Q$ and $\hat{C}$ depends only on the $L^\infty(Q)$ bounds on b and av. Therefore, an upper bound on the right hand sides of (4.20) is such that

$$\begin{aligned}
&\hat{L_1}w + (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)w \leq C, \\
&\hat{L_2}y + (-e^{\lambda t}\beta h_1 + \mu_2)y \leq C, \\
&\hat{L_3}z + (\mu_1)z \leq C \qquad\qquad \text{a.e. } (x,t) \in Q.
\end{aligned} \tag{4.22}$$

Consider the functions $W(x,t) = w(x,t) - Ct, Y(x,t) = y(x,t) - Ct$ and $Z(x,t) = z(x,t) - Ct$. We have

$$\begin{aligned}
\hat{L_1}W + (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)W &= \hat{L_1}w + (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)w \\
&\quad - \hat{L_1}(Ct) - (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)Ct \\
&= \hat{L_1}w + (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)w - C \\
&\quad - \left(\lambda + \sum_{i=1}^{n}(b_i^1)_{x_i}\right)Ct - (-b + \mu_1 + e^{\lambda t}\beta h_2 + av)Ct \\
&\leq -\left(\lambda + \sum_{i=1}^{n}(b_i^1)_{x_i} - b + \mu_1 + e^{\lambda t}\beta h_2 + av\right)Ct \\
&\leq 0
\end{aligned}$$

because of inequality (4.22) and the first choice of λ large and second choice of T sufficiently small so that (4.21) holds. Similarly, we also have

$$\begin{aligned}
\hat{L_2}Y + (-e^{\lambda t}\beta h_1 + \mu_2)Y &\leq 0 \\
\hat{L_3}Z + \mu_1 Z &\leq 0
\end{aligned}$$

along with initial conditions

$$W(x,0) = S_0(x), \qquad Y(x,0) = I_0(x), \qquad Z(x,0) = R_0(x) \qquad \text{for } x \in \Omega$$

and boundary conditions

$$\frac{\partial W}{\partial \nu} = 0, \qquad \frac{\partial Y}{\partial \nu} = 0, \qquad \frac{\partial Z}{\partial \nu} = 0 \qquad \text{for } t \in (0,T), x \in \partial\Omega.$$

Again recalling the extension of the Maximum Principle to weak solutions [46], we conclude $W(x,t) \leq \|S_0(x)\|_{L^\infty(\Omega)}, Y(x,t) \leq \|I_0(x)\|_{L^\infty(\Omega)}, Z(x,t) \leq \|R_0(x)\|_{L^\infty(\Omega)}$ a.e. $(x,t) \in Q$. Equivalently, we have

$$\begin{aligned} w(x,t) &\leq \|S_0(x)\|_{L^\infty(\Omega)} + Ct, \\ y(x,t) &\leq \|I_0(x)\|_{L^\infty(\Omega)} + Ct, \\ z(x,t) &\leq \|R_0(x)\|_{L^\infty(\Omega)} + Ct \qquad \text{a.e. } (x,t) \in Q. \end{aligned}$$

For $T \leq \frac{M}{2C}$, we have

$$\begin{aligned} w(x,t) &\leq \|S_0(x)\|_{L^\infty(\Omega)} + \frac{M}{2} \\ &\leq \max\{\|S_0(x)\|_{L^\infty(\Omega)}, \|I_0(x)\|_{L^\infty(\Omega)}, \|R_0(x)\|_{L^\infty(\Omega)}\} + \frac{M}{2} \\ &\leq \frac{M}{2} + \frac{M}{2} \\ &= M \qquad\qquad \text{a.e. } (x,t) \in Q \end{aligned}$$

and similarly

$$y(x,t) \leq M \text{ and } z(x,t) \leq M \text{ a.e. } (x,t) \in Q.$$

Thus we have shown for $(h_1, h_2, h_3) \in X^3$, the solution $(w, y, z) \in V^3$ to (4.17)-(4.19) is also in X^3.

Define the map $A : X^3 \to X^3$ such that $A(h_1, h_2, h_3) = (w, y, z)$. We now show A is a strict contraction. Choose $(h_1, h_2, h_3), (\bar{h_1}, \bar{h_2}, \bar{h_3}) \in X^3$. Define $(w, y, z) = A(h_1, h_2, h_3)$

and $(\bar{w}, \bar{y}, \bar{z}) = A(\bar{h_1}, \bar{h_2}, \bar{h_3})$. Consider the differences $(w - \bar{w})$, $(y - \bar{y})$, and $(z - \bar{z})$. The PDEs satisfied by these differences are

$$\frac{\partial(w-\bar{w})}{\partial t} + \lambda(w-\bar{w}) - \sum_{i,j=1}^{n}(a_{ij}^1(w-\bar{w})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^1(w-\bar{w}))_{x_i} = b(w-\bar{w}) + b(h_3 - \bar{h_3})$$
$$- \mu_1(w-\bar{w}) - e^{\lambda t}\beta(wh_2 - \bar{w}\bar{h}) - av(w-\bar{w}),$$
$$\frac{\partial(y-\bar{y})}{\partial t} + \lambda(y-\bar{y}) - \sum_{i,j=1}^{n}(a_{ij}^2(y-\bar{y})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^2(y-\bar{y}))_{x_i} = e^{\lambda t}\beta(h_1 y - \bar{h_1}\bar{y}) - \mu_2(y-\bar{y}),$$
$$\frac{\partial(z-\bar{z})}{\partial t} + \lambda(z-\bar{z}) - \sum_{i,j=1}^{n}(a_{ij}^3(z-\bar{z})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^3(z-\bar{z}))_{x_i} = -\mu_1(z-\bar{z}) + av(h_1 - \bar{h_1})$$

a.e. $(x,t) \in Q$ with initial conditions

$$(w-\bar{w})(x,0) = 0, \qquad (y-\bar{y})(x,0) = 0, \qquad (z-\bar{z})(x,0) = 0 \qquad \text{for x} \in \Omega$$

and boundary conditions

$$\frac{\partial(w-\bar{w})}{\partial \nu} = 0, \qquad \frac{\partial(y-\bar{y})}{\partial \nu} = 0, \qquad \frac{\partial(z-\bar{z})}{\partial \nu} = 0 \qquad \text{for all } t \in (0,T),\ x \in \partial\Omega.$$

We multiply the PDEs satisfied by $(w - \bar{w})$, $(y - \bar{y})$, and $(z - \bar{z})$ by test functions $(w - \bar{w})$, $(y - \bar{y})$, and $(z - \bar{z})$ respectively and integrate over the domain $Q_s = \Omega \times (0, s)$ for arbitrary $s \in (0, T)$. After integrating by parts on the a_{ij} terms, we have

$$\int_{Q_s}(w-\bar{w})_t(w-\bar{w})dxdt + \lambda\int_{Q_s}(w-\bar{w})^2dxdt + \int_{Q_s}\sum_{i,j=1}^{n}a_{ij}^1(w-\bar{w})_{x_i}(w-\bar{w})_{x_j}dxdt$$
$$+ \int_{Q_s}\sum_{i=1}^{n}(b_i^1(w-\bar{w}))_{x_i}(w-\bar{w})dxdt = \int_{Q_s}b(w-\bar{w})^2dxdt + \int_{Q_s}b(h_3-\bar{h_3})(w-\bar{w})dxdt$$
$$- \int_{Q_s}\mu_1(w-\bar{w})^2dxdt - \int_{Q_s}e^{\lambda t}\beta(wh_2 - \bar{w}\bar{h_2})(w-\bar{w})dxdt - \int_{Q_s}av(w-\bar{w})^2dxdt,$$

$$\int_{Q_s}(y-\bar{y})_t(y-\bar{y})dxdt+\lambda\int_{Q_s}(y-\bar{y})^2dxdt+\int_{Q_s}\sum_{i,j=1}^{n}a_{ij}^2(y-\bar{y})_{x_i}(y-\bar{y})_{x_j}dxdt$$
$$+\int_{Q_s}\sum_{i=1}^{n}(b_i^2(y-\bar{y}))_{x_i}(y-\bar{y})dxdt=\int_{Q_s}e^{\lambda t}\beta(h_1y-\bar{h_1}\bar{y})(y-\bar{y})dxdt-\int_{Q_s}\mu_2(y-\bar{y})^2dxdt,$$
$$\int_{Q_s}(z-\bar{z})_t(z-\bar{z})dxdt+\lambda\int_{Q_s}(z-\bar{z})^2dxdt+\int_{Q_s}\sum_{i,j=1}^{n}a_{ij}^3(z-\bar{z})_{x_i}(z-\bar{z})_{x_j}dxdt$$
$$+\int_{Q_s}\sum_{i=1}^{n}(b_i^3(z-\bar{z}))_{x_i}(z-\bar{z})dxdt=\int_{Q_s}\mu_1(z-\bar{z})^2dxdt+\int_{Q_s}av(h_1-\bar{h_1})(z-\bar{z})dxdt.$$

We use (4.10)-(4.12) and the ellipticity condition in assumption (4) to estimate several terms on the left hand side of the above equations. Then summing our inequalities, we obtain

$$\begin{aligned}
&\frac{1}{2}\int_{\Omega}(w-\bar{w})^2(x,s)+(y-\bar{y})^2(x,s)+(z-\bar{z})^2(x,s)dx+\lambda\int_{Q_s}(w-\bar{w})^2+(y-\bar{y})^2+(z-\bar{z})^2dxdt\\
&\quad+\theta\int_{Q_s}|\bigtriangledown(w-\bar{w})|^2+|\bigtriangledown(y-\bar{y})|^2+|\bigtriangledown(z-\bar{z})|^2dxdt\leq\frac{\theta}{2}\int_{Q_s}|\bigtriangledown(w-\bar{w})|^2+|\bigtriangledown(y-\bar{y})|^2\\
&\quad+|\bigtriangledown(z-\bar{z})|^2dxdt+C_0\int_{Q_s}(w-\bar{w})^2+(y-\bar{y})^2+(z-\bar{z})^2dxdt+\int_{Q_s}b(w-\bar{w})^2dxdt\\
&\quad+\int_{Q_s}b(h_3-\bar{h_3})(w-\bar{w})-\int_{Q_s}\mu_1(w-\bar{w})^2dxdt+\int_{Q_s}e^{\lambda t}\beta(wh_2-\bar{w}\bar{h_2})(w-\bar{w})dxdt \qquad (4.23)\\
&\quad-\int_{Q_s}av(w-\bar{w})^2dxdt+\int_{Q_s}e^{\lambda t}\beta(h_1y-\bar{h_1}\bar{y})(y-\bar{y})dxdt-\int_{Q_s}\mu_2(y-\bar{y})^2dxdt\\
&\quad+\int_{Q_s}\mu_1(z-\bar{z})^2dxdt+\int_{Q_s}av(h_1-\bar{h_1})(z-\bar{z})dxdt
\end{aligned}$$

where C_0 is a constant which depends on θ and the bounds on the coefficients b_i^k and $(b_i^k)_{x_i}$.

Next we use (4.11), (4.13), and the fact that $e^{\lambda t}\leq e^{2\lambda t}$ to estimate terms on the right hand side of the above inequality. An example of such estimation is

$$\begin{aligned}
\int_{Q_s}e^{\lambda t}\beta(h_1y-\bar{h_1}\bar{y})(y-\bar{y})dxdt&=\int_{Q_s}e^{\lambda t}\beta\left(y(h_1-\bar{h_1})+\bar{h_1}(y-\bar{y})\right)(y-\bar{y})dxdt\\
&=\int_{Q_s}e^{\lambda t}\beta y(h_1-\bar{h_1})(y-\bar{y})dxdt+\int_{Q_s}e^{\lambda t}\beta\bar{h_1}(y-\bar{y})^2dxdt
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{2}\int_{Q_s} e^{2\lambda t}y^2\beta^2(y-\bar{y})^2dxdt + \frac{1}{2}\int_{Q_s}(h_1-\bar{h_1})^2dxdt \\
&\qquad + \int_{Q_s} e^{\lambda t}\beta\bar{h_1}(y-\bar{y})^2dxdt \\
&\leq \frac{1}{2}\int_{Q_s} e^{2\lambda t}y^2\beta^2(y-\bar{y})^2dxdt + \frac{1}{2}\int_{Q_s}(h_1-\bar{h_1})^2dxdt \\
&\qquad + \int_{Q_s} e^{2\lambda t}\beta\bar{h_1}(y-\bar{y})^2dxdt \\
&\leq D_0(M+M^2)e^{2\lambda T}\int_{Q_s}(y-\bar{y})^2dxdt + \frac{1}{2}\int_{Q_s}(h_1-\bar{h_1})^2dxdt
\end{aligned}
$$

where D_0 is a positive constant that depends only on the $L^\infty(Q)$ bounds on β. After similar estimates on terms in (4.23), we have

$$
\begin{aligned}
&\frac{1}{2}\int_\Omega (w-\bar{w})^2(x,s)+(y-\bar{y})^2(x,s)+(z-\bar{z})^2(x,s)dx + \lambda\int_{Q_s}(w-\bar{w})^2+(y-\bar{y})^2+(z-\bar{z})^2dxdt \\
&\quad + \theta\int_{Q_s} |\bigtriangledown(w-\bar{w})|^2+|\bigtriangledown(y-\bar{y})|^2+|\bigtriangledown(z-\bar{z})|^2dxdt \leq \frac{\theta}{2}\int_{Q_s}|\bigtriangledown(w-\bar{w})|^2+|\bigtriangledown(y-\bar{y})|^2 \\
&\quad + |\bigtriangledown(z-\bar{z})|^2dxdt + C_0\int_{Q_s}(w-\bar{w})^2+(y-\bar{y})^2+(z-\bar{z})^2dxdt + \int_{Q_s} b(w-\bar{w})^2dxdt \\
&\quad + \frac{1}{2}\int_{Q_s}(h_3-\bar{h_3})^2 + \frac{1}{2}\int_{Q_s} b^2(w-\bar{w})^2dxdt - \int_{Q_s}\mu_1(w-\bar{w})^2dxdt \\
&\quad + \frac{1}{2}\int_{Q_s}(h_2-\bar{h_2})^2dxdt + \frac{1}{2}\int_{Q_s} e^{2\lambda t}\beta^2w^2(w-\bar{w})^2dxdt + \int_{Q_s} e^{2\lambda t}\beta\bar{h_2}(w-\bar{w})^2dxdt \\
&\quad - \int_{Q_s} av(w-\bar{w})^2dxdt + \frac{1}{2}\int_{Q_s}(h_1-\bar{h_1})^2dxdt + \frac{1}{2}\int_{Q_s} e^{2\lambda t}\beta^2y^2(y-\bar{y})^2dxdt \\
&\quad + \int_{Q_s} e^{2\lambda t}\beta\bar{h_1}(y-\bar{y})^2dxdt - \int_{Q_s}\mu_2(y-\bar{y})^2dxdt + \int_{Q_s}\mu_1(z-\bar{z})^2dxdt \\
&\quad + \frac{1}{2}\int_{Q_s}(h_1-\bar{h_1})^2dxdt + \frac{1}{2}\int_{Q_s}(av)^2(z-\bar{z})^2dxdt.
\end{aligned}
$$

Noting the $L^\infty(Q)$ bounds on the coefficients, the product av and the variables $\bar{w}, \bar{y}, \bar{z}, h_1, h_2, h_3$, we can bound several terms on the right hand side above by constants and combine similar squared terms to get

$$\frac{1}{2}\int_\Omega (w-\bar{w})^2(x,s)+(y-\bar{y})^2(x,s)+(z-\bar{z})^2(x,s)dx+(\lambda-C_1-C_2(M+M^2)e^{2\lambda T})\int_{Q_s}(w-\bar{w})^2$$
$$+(y-\bar{y})^2+(z-\bar{z})^2dxdt+\frac{\theta}{2}\int_{Q_s}|\bigtriangledown(w-\bar{w})|^2+|\bigtriangledown(y-\bar{y})|^2+|\bigtriangledown(z-\bar{z})|^2dxdt$$
$$\leq\int_{Q_s}(h_1-\bar{h_1})^2+(h_2-\bar{h_2})^2+(h_3-\bar{h_3})^2dxdt$$

where C_1 and C_2 are constants depending on the $L^\infty(Q)$ bounds of the coefficients and av only. Moreover, for λ first chosen large and then T chosen sufficiently small, we have $\lambda-C_1-C_2(M+M^2)e^{2\lambda T}>0$ and thus

$$\int_\Omega (w-\bar{w})^2(s,x)+(y-\bar{y})^2(s,x)+(z-\bar{z})^2(s,x)dx$$
$$\leq 2\int_{Q_s}(h_1-\bar{h_1})^2+(h_2-\bar{h_2})^2+(h_3-\bar{h_3})^2dxdt. \tag{4.24}$$

Estimating the right hand side of (4.24), we have

$$2\int_{Q_s}(h_1-\bar{h_1})^2+(h_2-\bar{h_2})^2+(h_3-\bar{h_3})^2dxdt$$
$$\leq 2T\left(\sup_{0\leq t\leq T}\int_\Omega(h_1-\bar{h_1})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_2-\bar{h_2})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_3-\bar{h_3})^2dx\right). \tag{4.25}$$

Thus, (4.24) together with (4.25) yields

$$\int_\Omega(w-\bar{w})^2(s,x)dx\leq 2T\Big(\sup_{0\leq t\leq T}\int_\Omega(h_1-\bar{h_1})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_2-\bar{h_2})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_3-\bar{h_3})^2dx\Big),$$
$$\int_\Omega(y-\bar{y})^2(s,x)dx\leq 2T\Big(\sup_{0\leq t\leq T}\int_\Omega(h_1-\bar{h_1})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_2-\bar{h_2})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_3-\bar{h_3})^2dx\Big),$$
$$\int_\Omega(z-\bar{z})^2(s,x)dx\leq 2T\Big(\sup_{0\leq t\leq T}\int_\Omega(h_1-\bar{h_1})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_2-\bar{h_2})^2dx+\sup_{0\leq t\leq T}\int_\Omega(h_3-\bar{h_3})^2dx\Big).$$

We take the square root of both sides in each inequality. Then, we bound the right hand side above by recognizing

$$
\begin{aligned}
&\left(2T\Big(\sup_{0\le t\le T}\int_\Omega (h_1-\bar{h_1})^2dx+\sup_{0\le t\le T}\int_\Omega (h_2-\bar{h_2})^2dx+\sup_{0\le t\le T}\int_\Omega (h_3-\bar{h_3})^2dx\Big)\right)^{1/2}\\
&\quad\le (2T)^{1/2}\Big(\left(\sup_{0\le t\le T}\int_\Omega (h_1-\bar{h_1})^2dx\right)^{1/2}+\left(\sup_{0\le t\le T}\int_\Omega (h_2-\bar{h_2})^2dx\right)^{1/2}\\
&\qquad+\left(\sup_{0\le t\le T}\int_\Omega (h_3-\bar{h_3})^2dx\right)^{1/2}\Big)\\
&\quad= (2T)^{1/2}\Big(\sup_{0\le t\le T}\left(\int_\Omega (h_1-\bar{h_1})^2dx\right)^{1/2}+\sup_{0\le t\le T}\left(\int_\Omega (h_2-\bar{h_2})^2dx\right)^{1/2}\\
&\qquad+\sup_{0\le t\le T}\left(\int_\Omega (h_3-\bar{h_3})^2dx\right)^{1/2}\Big)\\
&\quad= (2T)^{1/2}\left(\|h_1-\bar{h_1}\|_X+\|h_2-\bar{h_2}\|_X+\|h_3-\bar{h_3}\|_X\right).
\end{aligned}
$$

This allows us to write

$$
\begin{aligned}
\left(\int_\Omega (w-\bar{w})^2(s,x)dx\right)^{1/2}&\le (2T)^{1/2}\left(\|h_1-\bar{h_1}\|_X+\|h_2-\bar{h_2}\|_X+\|h_3-\bar{h_3}\|_X\right),\\
\left(\int_\Omega (y-\bar{y})^2(s,x)dx\right)^{1/2}&\le (2T)^{1/2}\left(\|h_1-\bar{h_1}\|_X+\|h_2-\bar{h_2}\|_X+\|h_3-\bar{h_3}\|_X\right),\\
\left(\int_\Omega (z-\bar{z})^2(s,x)dx\right)^{1/2}&\le (2T)^{1/2}\left(\|h_1-\bar{h_1}\|_X+\|h_2-\bar{h_2}\|_X+\|h_3-\bar{h_3}\|_X\right).
\end{aligned}
$$

Taking the supremum over $0\le s\le T$ on both sides, we obtain

$$
\|w-\bar{w}\|_X+\|y-\bar{y}\|_X+\|z-\bar{z}\|_X\le 3\,(2T)^{1/2}\left(\|h_1-\bar{h_1}\|_X+\|h_2-\bar{h_2}\|_X+\|h_3-\bar{h_3}\|_X\right).
$$

We must choose T sufficiently small because of previous assumptions, but we can also ensure that T is small enough so that $3\,(2T)^{1/2}<1$. In doing so, we obtain the desired inequality

$$\|w-\bar{w}\|_X + \|y-\bar{y}\|_X + \|z-\bar{z}\|_X = \|A(h_1,h_2,h_3) - A(\bar{h_1},\bar{h_2},\bar{h_3})\|_{X^3}$$
$$\leq \gamma\left(\|h_1-\bar{h_1}\|_X + \|h_2-\bar{h_2}\|_X + \|h_3-\bar{h_3}\|_X\right)$$

for some $\gamma < 1$. Thus, the mapping A is a strict contraction. By Banach's fixed point theorem, there exists a unique solution $(w,y,z) \in V^3 \bigcap X^3$ to the PDE system (4.14)-(4.16).

It follows that for sufficiently small T, there exists an unique nonnegative solution $(S,I,R) \in V^3$ satisfying (4.7), (4.8), and (4.9). Furthermore, $0 \leq S(x,t) \leq e^{\lambda T}M$, $0 \leq I(x,t) \leq e^{\lambda T}M$, and $0 \leq R(x,t) \leq e^{\lambda T}M$ a.e. $(x,t) \in Q$. □

By use of a time interval stacking methods, an extension of this existence result can be done to overcome the restriction of T being sufficiently small. First, select $T_1 > 0$ such that the conditions in the proof are met. Apply Banach's fixed point theorem to find a solution existing on the interval $[0,T_1]$. Then repeat the argument to extend the solution to $[T_1,T_2]$ for some $T_2 > 0$. This process can be repeated numerous times. It should be noted that uniqueness of the optimal control is proven in Theorem 15 for only sufficiently small T.

Theorem 12. *There exists an optimal control $v^* \in U$ that minimizes the functional $J(v)$.*

Proof: The control variable v is uniformly bounded in Q and by Theorem 11 the state variables S, I, and R are also uniformly bounded in Q. Therefore

$$\inf_{v\in U} J(v) = \inf_{v\in U} \int_Q (AI - BS + C(v))dxdt > -\infty$$

and there exists a minimizing sequence $v^n \in U$ such that

$$\lim_{n\to\infty} J(v^n) = \inf_{v\in U} J(v).$$

Because of Theorem 11, we can define

$$(S^n, I^n, R^n) = (S,I,R)(v^n) \text{ for each n.}$$

We proceed to obtain uniform bounds on $\|S^n\|_V$, $\|I^n\|_V$, and $\|R^n\|_V$ for all n. Multiply the PDEs satisfied by S^n, I^n, and R^n by the test functions S^n, I^n, and R^n respectively and integrate over the domain $Q_s = \Omega \times (0, s)$ for arbitrary $s \in (0, T]$. Integration by parts is used on the a_{ij} terms and the integral equations are summed. For any n and $s \in (0, T]$, we have

$$\int_{Q_s} (S_t^n S^n + I_t^n I^n + R_t^n R^n) dxdt + \int_0^s B^1[\cdot, S^n, S^n] + B^2[\cdot, I^n, I^n] + B^3[\cdot, R^n, R^n] dt \tag{4.26}$$

$$\begin{aligned}
&= \int_{Q_s} b(S^n)^2 + \int_{Q_s} bR^n S^n - \int_{Q_s} \mu_1 (S^n)^2 dxdt - \int_{Q_s} \beta I^n (S^n)^2 dxdt \\
&\quad - \int_{Q_s} av^n (S^n)^2 dxdt + \int_{Q_s} \beta S^n (I^n)^2 dxdt - \int_{Q_s} \mu_2 (I^n)^2 dxdt \\
&\quad - \int_{Q_s} \mu_1 (R^n)^2 dxdt + \int_{Q_s} av^n S^n R^n dxdt.
\end{aligned}$$

Using (4.10)-(4.11) to estimate terms in (4.26), we have

$$\begin{aligned}
&\frac{1}{2} \int_\Omega (S^n)^2(x,s) + (I^n)^2(x,s) + (R^n)^2(x,s) dx + \theta \int_{Q_s} \left(| \bigtriangledown S^n|^2 + | \bigtriangledown I^n|^2 + | \bigtriangledown R^n|^2 \right) dxdt \\
&\quad \leq \frac{1}{2} \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx + \frac{\theta}{2} \int_{Q_s} | \bigtriangledown S^n|^2 + | \bigtriangledown I^n|^2 + | \bigtriangledown R^n|^2 dxdt \qquad (4.27) \\
&\quad + C \int_{Q_s} |S^n|^2 + |I^n|^2 + |R^n|^2 dxdt + \int_{Q_s} b(S^n)^2 dxdt + \frac{1}{2} \int_{Q_s} b^2 (R^n)^2 dxdt \\
&\quad + \frac{1}{2} \int_{Q_s} (S^n)^2 dxdt - \int_{Q_s} \mu_1 (S^n)^2 dxdt - \int_{Q_s} \beta I^n (S^n)^2 dxdt \\
&\quad - \int_{Q_s} av^n (S^n)^2 dxdt + \int_{Q_s} \beta S^n (I^n)^2 dxdt - \int_{Q_s} \mu_2 (I^n)^2 dxdt - \int_{Q_s} \mu_1 (R^n)^2 dxdt \\
&\quad + \frac{1}{2} \int_{Q_s} (av^n)^2 (S^n)^2 dxdt + \frac{1}{2} \int_{Q_s} (R^n)^2 dxdt
\end{aligned}$$

where C is a new positive constant that depends on θ and the bounds on the b_i^k and $(b_i^k)_{x_i}$ coefficients. Noting that our state variables, control variable, and coefficients are uniformly bounded in Q, we collect similar squared terms to transform (4.27) into

$$\int_\Omega (S^n)^2(x,s) + (I^n)^2(x,s) + (R^n)^2(x,s)dx + \theta \int_{Q_s} (|\bigtriangledown S^n|^2 + |\bigtriangledown I^n|^2 + |\bigtriangledown R^n|^2)dxdt \tag{4.28}$$

$$\leq \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx + \hat{C} \int_{Q_s} |S^n|^2 + |I^n|^2 + |R^n|^2 dxdt$$

where $\hat{C}$ is a positive constant which depends on θ and the $L^\infty(Q)$ bounds on the coefficients, control variable, and state variables. In particular, we have

$$\int_\Omega (S^n)^2(x,s) + (I^n)^2(x,s) + (R^n)^2(x,s)dx \leq \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx$$
$$+ \hat{C} \int_{Q_s} |S^n|^2 + |I^n|^2 + |R^n|^2 dxdt.$$

An application of Gronwall's Inequality to the above line yields

$$\int_\Omega (S^n)^2(x,s) + (I^n)^2(x,s) + (R^n)^2(x,s)dx \leq \left(1 + \hat{C}se^{\hat{C}s}\right) \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx$$

implying

$$\int_Q |S^n|^2 + |I^n|^2 + |R^n|^2 dxdt \leq T\left(1 + \hat{C}Te^{\hat{C}T}\right) \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx. \tag{4.29}$$

Thus, (4.29) together with (4.28) gives

$$\sup_{s\in(0,T)} \left(\int_\Omega (S^n)^2(x,s) + (I^n)^2(x,s) + (R^n)^2(x,s)dx \right) + \theta \int_Q (|\bigtriangledown S^n|^2 + |\bigtriangledown I^n|^2 + |\bigtriangledown R^n|^2)dxdt$$
$$\leq C_1 \int_\Omega (S_0)^2 + (I_0)^2 + (R_0)^2 dx$$

where C_1 depends on T, θ, and $L^\infty(Q)$ bounds on the states, control and coefficients. From

this, we can conclude that

$$\|S^n\|_V, \|I^n\|_V, \|R^n\|_V \text{ are uniformly bounded independent of } n. \tag{4.30}$$

From the PDE system and (4.30), we have

$$\|S_t^n\|, \|I_t^n\|, \|R_t^n\| \text{ are uniformly bounded in } L^2(0,T;H^1(\Omega)^*) \text{ independent of } n. \tag{4.31}$$

In light of (4.30) and (4.31) and because $v^n \in L^\infty(Q)$, there exists subsequences S^n, I^n, R^n and v^n such that

$$\begin{aligned}
&S^n \rightharpoonup S^* \text{ weakly in } V, \qquad S_t^n \rightharpoonup S_t^* \text{ weakly in } L^2(0,T;H^1(\Omega)^*)\\
&I^n \rightharpoonup I^* \text{ weakly in } V, \qquad I_t^n \rightharpoonup I_t^* \text{ weakly in } L^2(0,T;H^1(\Omega)^*)\\
&R^n \rightharpoonup R^* \text{ weakly in } V, \qquad R_t^n \rightharpoonup R_t^* \text{ weakly in } L^2(0,T;H^1(\Omega)^*)\\
&v^n \rightharpoonup v^* \text{ weakly in } L^2(Q).
\end{aligned}$$

We must show that S^*, I^*, R^* are the states associated with the optimal control v^*. The variational formulation of system (4.6) satisfied by S^n, I^n, R^n and test functions ϕ_1, ϕ_2, $\phi_3 \in V$ is

$$\begin{aligned}
\int_0^T \langle S_t^n, \phi_1\rangle dt + \int_0^T B^1[t, S^n, \phi_1]dt &= \int_Q b(S^n + R^n)\phi_1 dxdt - \int_Q \beta S^n I^n \phi_1 dxdt\\
&\quad - \int_Q (\mu_1 + av^n) S^n \phi_1 dxdt\\
\int_0^T \langle I_t^n, \phi_2\rangle dt + \int_0^T B^2[t, I^n, \phi_2]dt &= \int_Q \beta S^n I^n \phi_2 dxdt - \int_Q \mu_2 I^n \phi_2 dxdt\\
\int_0^T \langle R_t^n, \phi_3\rangle dt + \int_0^T B^3[t, R^n, \phi_3]dt &= \int_Q av^n S^n \phi_3 dxdt - \int_Q \mu_1 R^n \phi_3 dxdt.
\end{aligned} \tag{4.32}$$

To pass the limit in the above system, we need stronger convergence results. Using the

compactness result in Corollary 4 of Theorem 5 in [76], we have

$$S^n \to S^* \text{ strongly in } L^2(Q) \tag{4.33}$$

$$I^n \to I^* \text{ strongly in } L^2(Q) \tag{4.34}$$

$$R^n \to R^* \text{ strongly in } L^2(Q). \tag{4.35}$$

This gives sufficient convergence results for passing the limit in (4.32) as $n \to \infty$. We conclude $(S^*, I^*, R^*) = (S, I, R)(v^*)$. To finish the proof, recall $C(\cdot)$ is a lower semi-continuous convex function in the objective functional. Because every lower semi-continuous convex function of a real vector space remains lower semi-continuous when supplied with the weak topology [30], we have

$$\int_Q C(v^*)dxdt \leq \liminf_{n\to\infty} \int_Q C(v^n)dxdt.$$

This property gives us

$$\begin{aligned} J(v^*) &= \int_Q AI^* - BS^* + C(v^*)dxdt \\ &\leq \liminf_{n\to\infty} \int_Q AI^n - BS^n + C(v^n)dxdt \end{aligned}$$

and because we chose v^n to be a minimizing sequence, we have

$$\begin{aligned} &= \lim_{n\to\infty} \int_Q AI^n - BS^n + C(v^n)dxdt \\ &= \inf_{v\in U} J(v). \end{aligned}$$

Thus, v^* is the optimal control which minimizes the objective functional, i.e.

$$J(v^*) = \min_{v\in U} J(v).$$

□

4.3.2 Optimality System and Uniqueness

To characterize the optimal control, we will take Gateaux derivative of J with respect to v in the direction h, i.e. $\lim_{\epsilon\to 0}\frac{J(v+\epsilon h)-J(v)}{\epsilon}$. Because the objective functional J contains the state variables, we also differentiate the maps $v \to S = S(v)$, $v \to I = I(v)$, and $v \to R = R(v)$. These derivatives are called the sensitivities.

Theorem 13. *(Sensitivities) The mapping $v \in U \to (S, I, R) \in V^3$ is differentiable in the sense that there exists Ψ_1, Ψ_2, $\Psi_3 \in V$ such that*

$$\begin{aligned}\frac{S(v+\epsilon h)-S(v)}{\epsilon} &\rightharpoonup \Psi_1\\ \frac{I(v+\epsilon h)-I(v)}{\epsilon} &\rightharpoonup \Psi_2\\ \frac{R(v+\epsilon h)-R(v)}{\epsilon} &\rightharpoonup \Psi_3\end{aligned}$$

in V as $\epsilon \to 0$ for any $v \in U$ and $h \in L^\infty(Q)$ such that $v + \epsilon h \in U$ for ϵ small. Also Ψ_1, Ψ_2, Ψ_3 satisfy

$$\begin{aligned}L_1\Psi_1 &= b(\Psi_1+\Psi_3)-\mu_1\Psi_1-\beta\Psi_1 I-\beta\Psi_2 S-av\Psi_1-ahS, && \text{for } (x,t)\in Q\\ L_2\Psi_2 &= \beta\Psi_1 I+\beta\Psi_2 S-\mu_2\Psi_2, && (4.36)\\ L_3\Psi_3 &= -\mu_1\Psi_3+av\Psi_1+ahS,\end{aligned}$$

$$\Psi_1(x,0)=\Psi_2(x,0)=\Psi_3(x,0)=0 \qquad \text{for } x\in\Omega \tag{4.37}$$

$$\frac{\partial\Psi_1}{\partial\nu}=\frac{\partial\Psi_2}{\partial\nu}=\frac{\partial\Psi_3}{\partial\nu}=0 \qquad \text{on } \partial\Omega\times(0,T). \tag{4.38}$$

Proof: Choose $h \in L^\infty(Q)$, $v \in U$ such that $(v + \epsilon h) \in U$ for ϵ small. Define $(S^\epsilon, I^\epsilon, R^\epsilon) = (S, I, R)(v + \epsilon h)$. The equations satisfied by the quotients $(\frac{S^\epsilon-S}{\epsilon}), (\frac{I^\epsilon-I}{\epsilon})$, and $(\frac{R^\epsilon-R}{\epsilon})$ are

$$\frac{\partial(\frac{S^\epsilon - S}{\epsilon})}{\partial t} - \sum_{i,j=1}^{n}(a_{ij}^1(\frac{S^\epsilon - S}{\epsilon})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^1(\frac{S^\epsilon - S}{\epsilon}))_{x_i} = b((\frac{S^\epsilon - S}{\epsilon}) + (\frac{R^\epsilon - R}{\epsilon})) - \mu_1(\frac{S^\epsilon - S}{\epsilon})$$
$$- \beta(\frac{S^\epsilon I^\epsilon - SI}{\epsilon}) - \Big(\frac{a(v+\epsilon h)S^\epsilon - avS}{\epsilon}\Big), \tag{4.39}$$
$$\frac{\partial(\frac{I^\epsilon - I}{\epsilon})}{\partial t} - \sum_{i,j=1}^{n}(a_{ij}^2(\frac{I^\epsilon - I}{\epsilon})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^2(\frac{I^\epsilon - I}{\epsilon}))_{x_i} = \beta(\frac{S^\epsilon I^\epsilon - SI}{\epsilon}) - \mu_2(\frac{I^\epsilon - I}{\epsilon}),$$
$$\frac{\partial(\frac{S^\epsilon - S}{\epsilon})}{\partial t} - \sum_{i,j=1}^{n}(a_{ij}^3(\frac{R^\epsilon - R}{\epsilon})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^3(\frac{R^\epsilon - R}{\epsilon}))_{x_i} = -\mu_1(\frac{R^\epsilon - R}{\epsilon}) + \Big(\frac{a(v+\epsilon h)S^\epsilon - avS}{\epsilon}\Big).$$

The initial and boundary conditions satisfied by the quotients are

$$(\frac{S^\epsilon - S}{\epsilon})(x,0) = 0, \quad (\frac{I^\epsilon - I}{\epsilon})(x,0) = 0, \quad (\frac{R^\epsilon - R}{\epsilon})(x,0) = 0 \text{ for } x \in \Omega,$$

$$\frac{\partial(\frac{S^\epsilon - S}{\epsilon})}{\partial \nu} = 0, \quad \frac{\partial(\frac{I^\epsilon - I}{\epsilon})}{\partial \nu} = 0, \quad \frac{\partial(\frac{R^\epsilon - R}{\epsilon})}{\partial \nu} = 0 \text{ on } \partial\Omega \times (0,T).$$

In system (4.39), we use (4.13) to rewrite terms like $\frac{a(v+\epsilon h)S^\epsilon - avS}{\epsilon}$ as

$$\frac{a(v+\epsilon h)S^\epsilon - avS}{\epsilon} = ahS^\epsilon + \frac{S^\epsilon - S}{\epsilon}av.$$

As in the proof of Theorem 12, we multiple each equation by the appropriate test function, integrate over $Q_s = \Omega \times (0,s)$ for arbitrary $s \in (0,T]$, and sum the equations. Noting the state variables and control are bounded in $L^\infty(Q)$ independent of ϵ and using the estimation techniques as before, we find that for any $s \in (0,T]$

$$\int_\Omega (\frac{S^\epsilon - S}{\epsilon})^2(x,s) + (\frac{I^\epsilon - I}{\epsilon})^2(x,s) + (\frac{R^\epsilon - R}{\epsilon})^2(x,s)dx$$
$$+ \theta \int_{Q_s} |\bigtriangledown(\frac{S^\epsilon - S}{\epsilon})|^2 + |\bigtriangledown(\frac{I^\epsilon - I}{\epsilon})|^2 + |\bigtriangledown(\frac{R^\epsilon - R}{\epsilon})|^2 dxdt$$
$$\leq \hat{C}\int_{Q_s} (\frac{S^\epsilon - S}{\epsilon})^2 + (\frac{I^\epsilon - I}{\epsilon})^2 + (\frac{R^\epsilon - R}{\epsilon})^2 dxdt + \bar{C}\int_{Q_s}(ah)^2 dxdt$$

where $\hat{C}$ depends on θ and the $L^\infty(Q)$ bounds on the states, control, and coefficients. After an application of Gronwall's inequality and simplification, we have

$$\sup_{s\in(0,T]}\left(\int_\Omega (\frac{S^\epsilon - S}{\epsilon})^2(x,s) + (\frac{I^\epsilon - I}{\epsilon})^2(x,s) + (\frac{R^\epsilon - R}{\epsilon})^2(x,s)dx\right)$$
$$+\theta\int_Q |\bigtriangledown(\frac{S^\epsilon - S}{\epsilon})|^2 + |\bigtriangledown(\frac{I^\epsilon - I}{\epsilon})|^2 + |\bigtriangledown(\frac{R^\epsilon - R}{\epsilon})|^2 dxdt$$
$$\leq \bar{C}\int_Q (ah)^2 dxdt. \tag{4.40}$$

Thus, the right hand side of (4.40) is bounded independent of ϵ. We can conclude that $\|(\frac{S^\epsilon - S}{\epsilon})\|_V, \|(\frac{I^\epsilon - I}{\epsilon})\|_V, \|(\frac{R^\epsilon - R}{\epsilon})\|_V$ are uniformly bounded for ϵ small. These bounds justify the existence of Ψ_1, Ψ_2 and $\Psi_3 \in V$ and the convergence of the S, I, R quotients,

$$(\frac{S^\epsilon - S}{\epsilon}) \rightharpoonup \Psi_1, \quad (\frac{I^\epsilon - I}{\epsilon}) \rightharpoonup \Psi_2, \quad (\frac{R^\epsilon - R}{\epsilon}) \rightharpoonup \Psi_3 \ \text{ in } V.$$

As in the previous proof, we can also conclude

$$(\frac{S^\epsilon - S}{\epsilon})_t \rightharpoonup (\Psi_1)_t, \quad (\frac{I^\epsilon - I}{\epsilon})_t \rightharpoonup (\Psi_2)_t, \quad (\frac{R^\epsilon - R}{\epsilon})_t \rightharpoonup (\Psi_3)_t \ \text{ in } L^2(0,T;H^1(\Omega)^*)$$

and

$$(\frac{S^\epsilon - S}{\epsilon}) \to \Psi_1, \quad (\frac{I^\epsilon - I}{\epsilon}) \to \Psi_2, \quad (\frac{R^\epsilon - R}{\epsilon}) \to \Psi_3 \ \text{ in } L^2(Q).$$

These convergences justify that Ψ_1, Ψ_2, Ψ_3 solve system (4.36) with initial conditions (4.37) and boundary conditions (4.38). □

The sensitivities Ψ_1, Ψ_2, and Ψ_3 solve a linearized version of the state PDEs. We can rewrite (4.36) in terms of the linear operator $\mathbf{L}$ such that

$$\mathbf{L}\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\end{pmatrix} = \begin{pmatrix}-ahS\\ 0\\ ahS\end{pmatrix},$$

where

$$\mathbf{L}\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\end{pmatrix} = \begin{pmatrix}L_1\Psi_1\\ L_2\Psi_2\\ L_3\Psi_3\end{pmatrix} + M\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\end{pmatrix}$$

and

$$M = \begin{pmatrix}-b+\mu_1+\beta I & \beta S + av & -b\\ -\beta I & -\beta S+\mu_2 & 0\\ -av & 0 & \mu_1\end{pmatrix}.$$

We now find the adjoint of the operator $\mathbf{L}$. For $(p_1, p_2, p_3) \in V^3$ we must have

$$\int_Q (p_1,p_2,p_3)\Big(\mathbf{L}\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\end{pmatrix}\Big)dxdt = \int_Q (\Psi_1,\Psi_2,\Psi_3)\Big(\mathbf{L}^*\begin{pmatrix}p_1\\ p_2\\ p_3\end{pmatrix}\Big)dxdt$$

in the appropriate weak sense.

We refer to p_1, p_2, and p_3 as adjoint variables and define the adjoint PDEs as

$$\mathbf{L}^*\begin{pmatrix}p_1\\ p_2\\ p_3\end{pmatrix} = \begin{pmatrix}-B\\ A\\ 0\end{pmatrix} \tag{4.41}$$

where

$$\mathbf{L}^*\begin{pmatrix}p_1\\ p_2\\ p_3\end{pmatrix} = \begin{pmatrix}L_1^*p_1\\ L_2^*p_2\\ L_3^*p_3\end{pmatrix} + M^T\begin{pmatrix}p_1\\ p_2\\ p_3\end{pmatrix}$$

and the adjoint operators $L_k^*, k = 1, 2, 3$ are defined as

$$L_k^*p_k = -\frac{\partial p_k}{\partial t} - \sum_{i,j=1}^{n}(a_{ij}^k(x,t)(p_k)_{x_i})_{x_j} - \sum_{i=1}^{n} b_i^k(x,t)(p_k)_{x_i}.$$

The right hand side of equation (4.41) is the derivative of the integrand in $J(v)$ with respect to the S,I, and R state variables respectively. We must also attach appropriate boundary

conditions to the adjoint PDEs (4.41) to form the adjoint system.

Theorem 14. *Given an optimal control v^* and corresponding state solution (S^*, I^*, R^*), there exists a weak solution $(p_1, p_2, p_3) \in V^3$ satisfying the adjoint system (4.41) and boundary conditions*

$$p_k(x, T) = 0 \qquad \text{for } x \in \Omega, \tag{4.42}$$

$$\frac{\partial p_k}{\partial \nu} + (\sum_{i=1}^{n} b_i^k \cdot \eta_i) p_k = 0 \qquad \text{on } \partial\Omega \times (0, T) \tag{4.43}$$

for $k = 1, 2, 3$.

Furthermore, in the case where $C(v) = cv^2$ for some $c > 0$ the optimal control is characterized by

$$v^* = min\Big(\Big(\frac{(p_1 - p_3)aS}{2c}\Big)^+, v_{\max}\Big). \tag{4.44}$$

Proof: Suppose v^* is an optimal control and $(S^*, I^*, R^*) = (S, I, R)(v^*)$ are the corresponding state variables. The adjoint system becomes

$$\begin{aligned} L_1^* p_1 &= -B + bp_1 - \mu_1 p_1 - \beta I^* p_1 + \beta I^* p_2 + a v^* p_3, \\ L_2^* p_2 &= A - \beta S^* p_1 - a v^* p_1 + \beta S^* p_2 - \mu_2 p_2, \\ L_3^* p_3 &= b p_1 - \mu_1 p_3, \end{aligned} \tag{4.45}$$

$$p_1(x, T) = 0, \quad p_2(x, T) = 0, \quad p_3(x, T) = 0 \text{ for } x \in \Omega, \tag{4.46}$$

$$\frac{\partial p_1}{\partial \nu} + (\sum_{i=1}^{n} b_i^1 \cdot \eta_i) p_1 = 0, \ \frac{\partial p_2}{\partial \nu} + (\sum_{i=1}^{n} b_i^2 \cdot \eta_i) p_2 = 0, \ \frac{\partial p_3}{\partial \nu} + (\sum_{i=1}^{n} b_i^3 \cdot \eta_i) p_3 = 0, \text{ on } \partial\Omega \times (0, T). \tag{4.47}$$

Because the system is linear in the adjoint variables, there exists (p_1, p_2, p_3) satisfying

(4.45)-(4.47). To characterize the control, we differentiate the map from the control to the objective functional. Consider $v^\epsilon \equiv v^* + \epsilon h \in U$ and corresponding state solution $(S^\epsilon, I^\epsilon, R^\epsilon) = (S, I, R)(v^\epsilon)$. Since the minimum of the objective functional is attained at v^*, we have

$$\begin{aligned} 0 &\leq \lim_{\epsilon \to 0^+} \frac{J(v^* + \epsilon h) - J(v^*)}{\epsilon} \\ &= \lim_{\epsilon \to 0^+} \int_Q \Big(A\Big(\frac{I^\epsilon - I^*}{\epsilon}\Big) - B\Big(\frac{S^\epsilon - S^*}{\epsilon}\Big) + \frac{c(v^\epsilon)^2 - c(v^*)^2}{\epsilon} \Big) dxdt \end{aligned}$$

which in light of (4.13) can be written as

$$\begin{aligned} &= \lim_{\epsilon \to 0^+} \int_Q \Big(A\Big(\frac{I^\epsilon - I^*}{\epsilon}\Big) - B\Big(\frac{S^\epsilon - S^*}{\epsilon}\Big) + c\Big(\frac{v^\epsilon - v^*}{\epsilon}\Big)(v^\epsilon + v^*) \Big) dxdt \\ &= \int_Q (A\Psi_2 - B\Psi_1 + 2chv^*) dxdt \\ &= \int_Q \Big((\Psi_1, \Psi_2, \Psi_3) \begin{pmatrix} -B \\ A \\ 0 \end{pmatrix} + 2chv^* \Big) dxdt \\ &= \int_Q \Big((\Psi_1, \Psi_2, \Psi_3) \mathbf{L}^* \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + 2chv^* \Big) dxdt \end{aligned}$$

and in the appropriate weak sense, we have

$$= \int_Q \Big((p_1, p_2, p_3) \mathbf{L} \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{pmatrix} + 2chv^* \Big) dxdt$$

$$= \int_Q \Big((p_1, p_2, p_3) \begin{pmatrix} -ahS^* \\ 0 \\ ahS^* \end{pmatrix} + 2chv^* \Big) dxdt$$
$$= \int_Q (-hp_1 aS^* + hp_3 aS^* + 2chv^*) dxdt$$
$$= \int_Q h(2cv^* - p_1 aS^* + p_3 aS^*) dxdt.$$

The last line in the above inequalities simplifies to

$$0 \leq \int_Q h(2cv^* - ap_1 S^* + ap_3 S^*) dxdt$$

To characterize v^*, we consider three cases:

- On the set $\{(x,t) \in Q | v^*(x,t) = 0\}$, we choose nonnegative h with support on this set. Therefore, $0 \leq 2cv^* + (p_3 - p_1)aS^*$ or $0 = v^* \geq \frac{(p_1 - p_3)aS^*}{2c}$.
- On the set $\{(x,t) \in Q | 0 < v^*(x,t) < v_{\max}\}$, we can choose h with arbitrary sign and support on this set. Therefore $0 = 2cv^* + (p_3 - p_1)aS^*$ or $v^* = \frac{(p_1 - p_3)aS^*}{2c}$.
- On the set $\{(x,t) \in Q | v^*(x,t) = v_{\max}\}$, we choose nonpositive h. Therefore, $0 \geq 2cv^* + (p_3 - p_1)aS^*$ or $v_{\max} = v^* \leq \frac{(p_1 - p_3)aS^*}{2c}$.

Considering these three cases, the appropriate characterization of v^* is

$$v^* = min\Big(\Big(\frac{(p_1 - p_3)aS^*}{2c} \Big)^+, v_{\max} \Big).$$

□

The optimality system consists of the state system (4.6)-(4.8), the adjoint system (4.41)-(4.43), and the characterization of the control (4.44). For the case where $C(v) = cv^2$ for some $c > 0$, we have the following uniqueness result.

Theorem 15. *When T is sufficiently small, the solution of the optimality system is unique.*

Proof: Suppose (S,I,R),(p_1,p_2,p_3) and $(\bar{S},\bar{I},\bar{R})$, $(\bar{p_1},\bar{p_2},\bar{p_3})$ are two solutions of (4.6)-(4.8),(4.41)-(4.43). The associated control characterizations are

$$v = min\left(\left(\frac{(p_1-p_3)aS}{2c}\right)^+, v_{\max}\right), \qquad \bar{v} = min\left(\left(\frac{(\bar{p_1}-\bar{p_3})a\bar{S}}{2c}\right)^+, v_{\max}\right).$$

Through a change of functions, let

$$w = e^{-\lambda t}S, \quad y = e^{-\lambda t}I, \quad z = e^{-\lambda t}R, \quad q_1 = e^{\lambda t}p_1, \quad q_2 = e^{\lambda t}p_2, \quad q_3 = e^{\lambda t}p_3,$$
$$\bar{w} = e^{-\lambda t}\bar{S}, \quad \bar{y} = e^{-\lambda t}\bar{I}, \quad \bar{z} = e^{-\lambda t}\bar{R}, \quad \bar{q_1} = e^{\lambda t}\bar{p_1}, \quad \bar{q_2} = e^{\lambda t}\bar{p_2}, \quad \bar{q_3} = e^{\lambda t}\bar{p_3}.$$

We consider the differences $w-\bar{w}$, $y-\bar{y}$, $z-\bar{z}$, $q_1-\bar{q_1}$, $q_2-\bar{q_2}$, and $q_3-\bar{q_3}$. The PDEs satisfied by the differences are

$$\begin{aligned}
&\frac{\partial(w-\bar{w})}{\partial t} + \lambda(w-\bar{w}) - \sum_{i,j=1}^{n}(a_{ij}^1(w-\bar{w})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^1(w-\bar{w}))_{x_i} = b((w-\bar{w})+(z-\bar{z})) - \mu_1(w-\bar{w})\\
&\quad - e^{\lambda t}\beta(wy-\bar{w}\bar{y}) - (avw - a\bar{v}\bar{w}),\\
&\frac{\partial(y-\bar{y})}{\partial t} + \lambda(y-\bar{y}) - \sum_{i,j=1}^{n}(a_{ij}^2(y-\bar{y})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^2(y-\bar{y}))_{x_i} = e^{\lambda t}\beta(wy-\bar{w}\bar{y}) - \mu_2(y-\bar{y}),\\
&\frac{\partial(z-\bar{z})}{\partial t} + \lambda(z-\bar{z}) - \sum_{i,j=1}^{n}(a_{ij}^3(z-\bar{z})_{x_i})_{x_j} + \sum_{i=1}^{n}(b_i^3(z-\bar{z}))_{x_i} = \mu_1(z-\bar{z}) + (avz - a\bar{v}\bar{z}),\\
&-\frac{\partial(q_1-\bar{q_1})}{\partial t} + \lambda(q_1-\bar{q_1}) - \sum_{i,j=1}^{n}(a_{ij}^1(q_1-\bar{q_1})_{x_i})_{x_j} - \sum_{i=1}^{n}b_i^1(q_1-\bar{q_1})_{x_i} = b(q_1-\bar{q_1}) - \mu_1(q_1-\bar{q_1})\\
&\quad - e^{\lambda t}\beta(yq_1-\bar{y}\bar{q_1}) + e^{\lambda t}\beta(yq_2-\bar{y}\bar{q_2}) + (avq_3 - a\bar{v}\bar{q_3}),\\
&-\frac{\partial(q_2-\bar{q_2})}{\partial t} + \lambda(q_2-\bar{q_2}) - \sum_{i,j=1}^{n}(a_{ij}^2(q_2-\bar{q_2})_{x_i})_{x_j} - \sum_{i=1}^{n}b_i^2(q_2-\bar{q_2})_{x_i} = -e^{\lambda t}\beta(wq_1-\bar{w}\bar{q_1})\\
&\quad - (avq_1 - a\bar{v}\bar{q_1}) + e^{\lambda t}\beta(wq_2-\bar{w}\bar{q_2}) - \mu_2(q_2-\bar{q_2}),\\
&-\frac{\partial(q_3-\bar{q_3})}{\partial t} + \lambda(q_3-\bar{q_3}) - \sum_{i,j=1}^{n}(a_{ij}^3(q_3-\bar{q_3})_{x_i})_{x_j} - \sum_{i=1}^{n}b_i^3(q_3-\bar{q_3})_{x_i} = b(q_1-\bar{q_1}) - \mu_1(q_3-\bar{q_3})
\end{aligned}$$

with initial and final time conditions

$$(w-\bar{w})(0,x) = (y-\bar{y})(0,x) = (z-\bar{z})(0,x) = 0,$$

$$(q_1-\bar{q_1})(T,x) = (q_2-\bar{q_2})(T,x) = (q_3-\bar{q_3})(T,x) = 0 \qquad \text{for all x} \in \Omega,$$

and boundary conditions

$$\frac{\partial(q_1-\bar{q_1})}{\partial\nu} + (\sum_{i=1}^{n} b_i^1\eta_i)(q_1-\bar{q_1}) = 0,$$
$$\frac{\partial(q_2-\bar{q_2})}{\partial\nu} + (\sum_{i=1}^{n} b_i^2\eta_i)(q_2-\bar{q_2}) = 0,$$
$$\frac{\partial(q_3-\bar{q_3})}{\partial\nu} + (\sum_{i=1}^{n} b_i^3\eta_i)(q_3-\bar{q_3}) = 0,$$
$$\frac{\partial(w-\bar{w})}{\partial\nu} = \frac{\partial(y-\bar{y})}{\partial\nu} = \frac{\partial(z-\bar{z})}{\partial\nu} = 0 \qquad \text{for all } t\in[0,T], x\in\partial\Omega.$$

Multiplying each PDE by the appropriate test function in V and integrating over the domain Q, we obtain

$$\frac{1}{2}\int_\Omega (w-\bar{w})^2(T,x)dx + \lambda\int_Q (w-\bar{w})^2 dxdt + \int_Q \sum_{i,j=1}^{n} a_{ij}^1 (w-\bar{w})_{x_i}(w-\bar{w})_{x_j} dxdt$$
$$+ \int_Q \sum_{i=1}^{n} (b_i^1(w-\bar{w}))_{x_i}(w-\bar{w})dxdt = \int_Q b(w-\bar{w})^2 dxdt + \int_Q b(z-\bar{z})(w-\bar{w})dxdt$$
$$- \int_Q \mu_1(w-\bar{w})^2 dxdt - \int_Q e^{\lambda t}\beta(wy-\bar{w}\bar{y})(w-\bar{w})dxdt - \int_Q (avw-a\bar{v}\bar{w})(w-\bar{w})dxdt,$$

$$\frac{1}{2}\int_\Omega (y-\bar{y})^2(T,x)dx + \lambda\int_Q (y-\bar{y})^2 dxdt + \int_Q \sum_{i,j=1}^{n} a_{ij}^2 (y-\bar{y})_{x_i}(y-\bar{y})_{x_j} dxdt$$
$$+ \int_Q \sum_{i=1}^{n} (b_i^2(y-\bar{y}))_{x_i}(y-\bar{y})dxdt = \int_Q e^{\lambda t}\beta(wy-\bar{w}\bar{y})(y-\bar{y})dxdt - \int_Q \mu_2(y-\bar{y})^2 dxdt,$$

$$\frac{1}{2}\int_\Omega (z-\bar{z})^2(T,x)dx+\lambda\int_Q (z-\bar{z})^2dxdt+\int_Q\sum_{i,j=1}^n a_{ij}^3(z-\bar{z})_{x_i}(z-\bar{z})_{x_j}$$
$$+\int_Q\sum_{i=1}^n (b_i^3(z-\bar{z}))_{x_i}(z-\bar{z})dxdt=\int_Q \mu_1(z-\bar{z})^2dxdt+\int_Q (avz-a\bar{v}\bar{z})(z-\bar{z})dxdt,$$
$$\frac{1}{2}\int_\Omega (q_1-\bar{q_1})^2(0,x)dx+\lambda\int_Q (q_1-\bar{q_1})^2dxdt+\int_Q\sum_{i,j=1}^n a_{ij}^1(q_1-\bar{q_1})_{x_i}(q_1-\bar{q_1})_{x_j}dxdt$$
$$+\int_Q\sum_{i=1}^n b_i^1(q_1-\bar{q_1})(q_1-\bar{q_1})_{x_i}dxdt=\int_Q b(q_1-\bar{q_1})^2dxdt-\int_Q \mu_1(q_1-\bar{q_1})^2dxdt$$
$$-\int_Q e^{\lambda t}\beta(yq_1-\bar{y}\bar{q_1})(q_1-\bar{q_1})dxdt+\int e^{\lambda t}\beta(yq_2-\bar{y}\bar{q_2})(q_1-\bar{q_1})dxdt$$
$$+\int_Q (avq_3-a\bar{v}\bar{q_3})(q_1-\bar{q_1})dxdt,$$
$$\frac{1}{2}\int_\Omega (q_2-\bar{q_2})^2(0,x)dx+\lambda\int_Q (q_2-\bar{q_2})^2dxdt+\int_Q\sum_{i,j=1}^n a_{ij}^1(q_2-\bar{q_2})_{x_i}(q_2-\bar{q_2})_{x_j}dxdt$$
$$+\int_Q\sum_{i=1}^n b_i^2(q_2-\bar{q_2})(q_2-\bar{q_2})_{x_i}dxdt=-\int_Q e^{\lambda t}\beta(wq_1-\bar{w}\bar{q_1})(q_2-\bar{q_2})dxdt$$
$$-\int_Q (avq_1-a\bar{v}\bar{q_1})(q_2-\bar{q_2})dxdt+\int_Q e^{\lambda t}\beta(wq_2-\bar{w}\bar{q_2})(q_2-\bar{q_2})dxdt-\int_Q \mu_2(q_2-\bar{q_2})^2dxdt,$$
$$\frac{1}{2}\int_\Omega (q_3-\bar{q_3})^2(0,x)dx+\lambda\int_Q (q_3-\bar{q_3})^2dxdt+\int_Q\sum_{i,j=1}^n a_{ij}^3(q_3-\bar{q_3})_{x_i}(q_3-\bar{q_3})_{x_j}dxdt$$
$$+\int_Q\sum_{i=1}^n b_i^3(q_3-\bar{q_3})(q_3-\bar{q_3})_{x_i}dxdt=\int_Q b(q_1-\bar{q_1})(q_3-\bar{q_3})dxdt-\int_Q \mu_1(q_3-\bar{q_3})^2dxdt.$$

Here we will use estimation techniques such as

$$|av-a\bar{v}|=|a\left(min\left(\left(\frac{(q_1-q_3)aw}{2c}\right)^+,v_{\max}\right)\right)-a\left(min\left(\left(\frac{(\bar{q_1}-\bar{q_3})a\bar{w}}{2c}\right)^+,v_{\max}\right)\right)|$$
$$\leq |a\left(\frac{a(q_1-q_3)w-a(\bar{q_1}-\bar{q_3})\bar{w}}{2c}\right)|$$
$$=\frac{a^2}{2c}|q_1(w-\bar{w})+\bar{w}(q_1-\bar{q_1})-q_3(w-\bar{w})-\bar{w}(q_3-\bar{q_3})|.$$

This estimate, along with techniques in (4.10)-(4.13) can be combined to estimate terms like

$$
\begin{aligned}
|\int_Q (avz - a\bar{v}\bar{z})(z-\bar{z})dxdt| &\leq \int_Q |(avz - a\bar{v}\bar{z})(z-\bar{z})|dxdt \\
&= \int_Q |z(av - a\bar{v})(z-\bar{z}) + a\bar{v}(z-\bar{z})^2|dxdt \\
&\leq \int_Q \frac{a^2|z|}{2c}\Big(|q_1||w-\bar{w}| + |\bar{w}||q_1 - \bar{q_1}| + |q_3||w-\bar{w}| \\
&\qquad + |\bar{w}||q_3 - \bar{q_3}|\Big)|z-\bar{z}| + a|z-\bar{z}|^2 dxdt \\
&\leq C\int_Q |w-\bar{w}|^2 + |z-\bar{z}|^2 + |q_1 - \bar{q_1}|^2 + |q_3 - \bar{q_3}|^2 dxdt
\end{aligned}
$$

for some constant C which depends on the $L^\infty(Q)$ bounds on the states and adjoints.

After summing the six integral equations and estimating several terms as in the example above, we obtain

$$
\begin{aligned}
&\frac{1}{2}\int_\Omega (w-\bar{w})^2(T,x) + (y-\bar{y})^2(T,x) + (z-\bar{z})^2(T,x) + (q_1-\bar{q_1})^2(0,x) + (q_2-\bar{q_2})^2(0,x) \\
&\quad + (q_3-\bar{q_3})^2(0,x)dx + (\lambda - C_1 - C_2 e^{\lambda T})\int_Q |w-\bar{w}|^2 + |y-\bar{y}|^2 + |z-\bar{z}|^2 + |q_1-\bar{q_1}|^2 \\
&\quad + |q_2-\bar{q_2}|^2 + |q_3-\bar{q_3}|^2 dxdt + \frac{\theta}{2}\int_Q |\bigtriangledown(w-\bar{w})|^2 + |\bigtriangledown(y-\bar{y})|^2 + |\bigtriangledown(z-\bar{z})|^2 \\
&\quad + |\bigtriangledown(q_1-\bar{q_1})|^2 + |\bigtriangledown(q_2-\bar{q_2})|^2 + |\bigtriangledown(q_3-\bar{q_3})|^2 dxdt \leq 0.
\end{aligned}
$$

If we choose λ sufficiently large and T sufficiently small so that $\lambda - C_1 - C_2 e^{\lambda T} > 0$, then we must have $w = \bar{w}$, $y = \bar{y}$, $z = \bar{z}$, $q_1 = \bar{q_1}$, $q_2 = \bar{q_2}$, and $q_3 = \bar{q_3}$. Therefore, $S = \bar{S}$, $I = \bar{I}$,$R = \bar{R}$,$p_1 = \bar{p_1}$,$p_2 = \bar{p_2}$, and $p_3 = \bar{p_3}$. Finally, because the optimal control is characterized in terms of the states and the adjoints, we have $v = \bar{v}$. □

The solutions to the optimality system are unique, implying an unique characterization of the control.

4.4 Application to Rabies and Raccoons

In this section, we provide numerical illustrations of our optimal spatiotemporal vaccination strategies with parameter values chosen specifically for rabies and raccoons. Rabies, a viral neuroinvasive disease, is transmitted through the bite of a rabid animal. In an effort to eliminate rabies among raccoons, two types of oral rabies vaccine baits are distributed over designated areas [77]. One type of bait is made of fish meal, a binder, and fish oil in the form of a hard brown pellet about 1-1/4" square in size. Inside is a vaccine chamber, which is filled with the rabies vaccine. These baits are placed by hand. The other type of bait contains the same ingredients and is distributed by plane. The air-dropped baits are flat, clear packets about 3/4" by 2" in size with an oily coating. After eating a bait, a healthy raccoon will develop antibodies in two to three weeks that will protect it if exposed to an infected raccoon.

Spatial heterogeneity and long-distance translocation (LDT) play important roles in the spatiotemporal dynamics and management of rabies among raccoons. In 2002, Smith *et al.* [78] incorporate spatial heterogeneity into a network model based on the adjacency of Connecticut townships. In their model, the rate of disease spread into a township depends on the fraction of adjacent townships that are experiencing rabies among raccoons. Smith *et al.* utilize a spatiotemporal data set from the state of Connecticut in determining appropriate rates of spread. Spanning over 50 months, the extensive data set indicates the month in which the first case of raccoon rabies was reported within 168 Connecticut townships. Five models were investigated, including one with a simple homogeneous rate of spread and four others with rates of spread correlating to human density and/or separation of townships by a river. All models incorporated LDT of raccoons by humans. The model accounting for major rivers and constant LDT dispersal rates was shown to best fit the irregular pattern of disease spread across the state. The authors found that large rivers act as semi-permeable barriers, leading to a 7-fold reduction in the local rates of rabies propagation. Human population density was weakly associated with the rate of local spread and rate of translocation. In 2005, Smith *et al.* [79] use a similar network model to investigate the role of LDT events in shaping the rabies epidemic in Connecticut. While LDT events are

common, only a few were shown to become new epidemic foci. In this work, the authors revisit the idea of spatial heterogeneity and show the spread of rabies is associated with forest cover. Analysis and comparison with the Connecticut data revealed that the spread of rabies slowed substantially between townships that were lightly forested. In the best model, rabies did not cross rivers separating heavily forested townships.

We confine our attention to the spread of rabies over a period of 52 weeks on a rectangular domain, $\Omega \subseteq R^2$. We choose dimensions 30 km $\times$ 20 km, an area of suitable size for investigating rabies spread over a period of 52 weeks. In agreement with the Connecticut data set [78], infection starting in one corner of our domain can spread, possibly irregularly, to the opposite corner within a 52-week period.

For this application, we simplify our original model to include only diffusive movement in the x and y directions. Given a control $v(x, y, t)$ representing the density of vaccine baits at location $(x, y) \in \Omega$ on week t, the corresponding susceptible (S), infected (I), and immune (R) raccoon population densities satisfy

$$\frac{\partial S}{\partial t}(x,y,t) = a_{11}(x,y)S_{xx}(x,y,t) + a_{22}(x,y)S_{yy}(x,y,t) + b(t)(S(x,y,t) + R(x,y,t)) \\ - \mu_1 S(x,y,t) - \beta S(x,y,t)I(x,y,t) - av(x,y,t)S(x,y,t), \tag{4.48}$$

$$\frac{\partial I}{\partial t}(x,y,t) = a_{11}(x,y)I_{xx}(x,y,t) + a_{22}I_{yy}(x,y,t) + \beta S(x,y,t)I(x,y,t) - \mu_2 I(x,y,t), \tag{4.49}$$

$$\frac{\partial R}{\partial t}(x,y,t) = a_{11}(x,y)R_{xx}(x,y,t) + a_{22}(x,y)R_{yy}(x,y,t) - \mu_1 R(x,y,t) + av(x,y,t)S(x,y,t) \tag{4.50}$$

for all $(x, y, t) \in \Omega \times [0, T]$. Initial conditions are specified by

$$S(x,y,0) = S_0(x,y), \quad I(x,y,0) = I_0(x,y), \quad R(x,y,0) = R_0(x,y) \text{ for } (x,y) \in \Omega \tag{4.51}$$

and no-flux boundary conditions are given by

$$\frac{\partial S}{\partial x} = 0, \quad \frac{\partial I}{\partial x} = 0, \quad \frac{\partial R}{\partial x} = 0 \text{ on } (\{x = 0\} \cup \{x = 30\}) \times (0, T), \tag{4.52}$$

$$\frac{\partial S}{\partial y} = 0, \quad \frac{\partial I}{\partial y} = 0, \quad \frac{\partial R}{\partial y} = 0 \text{ on } (\{y = 0\} \cup \{y = 20\}) \times (0, T). \tag{4.53}$$

Several rabies models include additional classes to track the exposed (infectious but not infected) raccoons as well as the dynamics of the baits on the ground [19, 25, 26]. For simplicity, we classify a raccoon to be susceptible if not previously exposed to rabies, infected if able to transmit rabies, or immune if vaccinated. Data and model predictions indicate low levels of natural immunity (1 - 5%) within raccoon populations, suggesting that raccoons develop little or no natural immunity to rabies [4]. Thus, we ignore natural immunity in our model and assume all rabies immunity is gained through vaccination.

We refer to Coyne *et al.* [25] and Clayton *et al.* [19] for values of birth and death rates. Raccoons give birth during the spring of each year, March 20 - June 21, an approximate 14 week period. Assuming a 50/50 sex rate within the population and that half the population are mature females, a reproductive rate of 1.34 $\frac{1}{\text{year}}$ is estimated [19, 25]. Dividing this yearly rate by 14 week period, we find $b(t) = 0.096\frac{1}{\text{week}}$ for t within the birthing period. When t is not within the birthing period, we let $b(t) = 0\frac{1}{\text{week}}$. Without loss of generality, we assume the birthing period to be weeks 13 through 27. The constant year-long natural death rate, $\mu_1 = 0.026\frac{1}{\text{week}}$, is calculated so that in absence of any disease or spatial spread, the susceptible population at $t = 0$ weeks and $t = 52$ weeks are approximately equal [19]. Rabies-related death rate is estimated to be $\mu_2 = 0.490\frac{1}{\text{week}}$ [19, 25].

Rabies incidence is formulated as a mass action term, βSI. The rate of infection βI is taken to be $0.03I$ [19, 78, 69]. The vaccine uptake rate a, a parameter with units $\frac{1}{\text{vaccine}\cdot\text{week}}$, is an indication of how successful the grounded baits are in vaccinating a raccoon. That is, to successfully vaccinate a raccoon, a bait must be first found and then eaten by a susceptible raccoon. This process can be inhibited by deterioration of the bait, human removal of the bait, or consumption of the bait by an animal other than a susceptible raccoon. The value of a is not straightforward and therefore we present results for a range

of values, from $a = 0.01 \frac{1}{\text{vaccine}\cdot\text{week}}$ to $a = 0.03 \frac{1}{\text{vaccine}\cdot\text{week}}$.

The set of admissible controls, denoted by U, will consist of all measurable functions $v(x, y, t)$ satisfying $0 \leq v(x, y, t) \leq v_{\max}$, for a.e. $(x, y, t) \in \Omega \times [0, T]$. Here, $v_{\max}$ is a large positive constant representing an upperbound on the density of baits placed at each location. The optimal control problem under consideration can be stated as follows. Find $v^*(x, y, t) \in U$ which minimizes the objective functional

$$\int_{\Omega \times [0,T]} (I(x, y, t) + cv^2(x, y, t)) dxdydt$$

subject to system (4.48) -(4.53). Above, T is the number of weeks over which we apply control and observe population dynamics. Note that we have simplified the original formulation of the objective functional so that $B = 0$ and, without loss of generality, $A = 1$. The cost of vaccination is expected to be a non-linear function of v. Here, we choose a quadratic function indicating additional costs associated with high levels of vaccination. The parameter c, with units $\frac{\text{raccoon/km}^2}{\text{vaccine}^2}$, balances the squared cost of vaccine with the cost associated with the infected population. Thus, an optimal control will be one which minimizes both the cost of vaccination and infectious population over the spatial domain Ω and a time period of length T weeks.

In our rabies S-I-R model, the diffusion coefficients and initial population sizes are spatially dependent. This allows us to distinguish between areas with inhibited raccoon movement and decreased raccoon densities (i.e. rivers, forests) as well areas with increased density due to raccoon translocation. The boundary conditions of our model assume raccoons neither enter nor exit the domain. Our goals are to (1) explore patterns of optimal vaccination in populations with simple, homogeneous rabies spread, (2) investigate how spatial heterogeneity, such as a river or forest cover, may affect optimal vaccination, and (3) address optimal allocation of vaccine given long-distance translocation of an infected raccoon.

The following two sections of results are differentiated by their pattern of disease propagation. In Section 4.4.1, we assume constant diffusion coefficients, $a_{11} = a_{22}$ and an initial homogeneous susceptible population. These assumptions yield a uniform wave of infection

traveling at constant speed across our domain. In Section 4.4.2, we consider spatially dependent diffusion coefficients, $a_{11}(x, y)$ and $a_{22}(x, y)$, and an initial heterogeneous susceptible population. This combination produces an irregular pattern of disease propagation. Also in this section, we incorporate LDT events within initial conditions. All numerical results are placed in Appendix C.

4.4.1 Spatial Homogeneity and Rabies Spread

In this section, we assume constant diffusion coefficients ($a_{11} = a_{22}$) and a homogeneous initial susceptible population. See Figure C.1 for the growth and decline of the susceptible population over a 52-week period in the absence of any infection. The changes in population density in Figure C.1 are attributed to the natural birth and natural death of susceptible raccoons.

Unless otherwise specified, we assume the infection is initiated at $t = 0$ week. Infection first appears on our domain as 2 infected raccoons/km^2 in 6 km^2 within the southwestern corner of our domain Ω. This site of initial infection is referred to as the subdomain Ω_I. See Figure 4.1 for it's exact placement. Table 4.1 displays all parameter values used in the homogeneous simulations. Constant diffusion coefficients, in combination with homogeneous initial conditions for the susceptible population, yield an expanding wave of infection, propagating from one corner of our spatial grid the the opposite corner within the 52-week time period. See Figure C.2 and Figure C.3 for six snapshots of the susceptible and infected populations, respectively, during the 52-week period in absence of intervention.

If the initial infection is detected during the first few weeks, a quickly implemented vaccination strategy should stop the spread of the infected population. To illustrate this, we assign a value to the balancing coefficient, $c = 0.10$, and display the optimal 10-week vaccination regime starting at $t = 0$ weeks in Figure C.4. Here, we have assumed a moderate vaccine uptake rate, $a = 0.02\frac{1}{\text{vaccine}\cdot\text{week}}$. The optimal vaccination strategy in Figure C.4 recommends placing the highest amount of vaccine at the location where rabies is initally detected with lesser amounts in all adjacent areas. Vaccination continues until $t = 6$ weeks, at which time we see little to no vaccine baits present in our domain. The corresponding susceptible, infected, and immune raccoon populations are displayed in Figure C.5, Figure

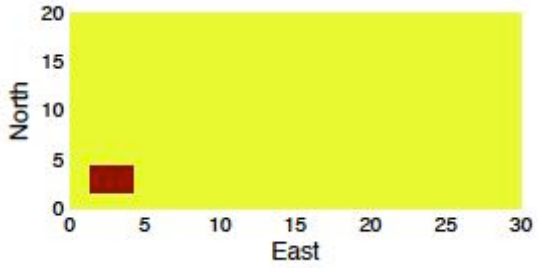


Figure 4.1: Our spatial domain Ω (yellow) and subdomain Ω_I (red). We assume infection originates in Ω_I.

Table 4.1: Parameters for homogeneous rabies spread and control

Parameter	**Notation**	**Value**
Spatial domain	Ω	30 km $\times$ 20 km
Spatial subdomain (site of initial infection)	Ω_I	3 km $\times$ 2 km
Initial susceptible population	$S_0(x, y)$	30 raccoons/km^2 for $(x, y) \in \Omega$
Initial infected population	$I_0(x, y)$	2 raccoons/km^2 for $(x, y) \in \Omega_I$ 0 raccoons/km^2 for $(x, y) \in \Omega \backslash \Omega_I$
Initial immune population	$R_0(x, y)$	0 raccoons/km^2 for $(x, y) \in \Omega$
Diffusion coefficient	a_{11}	0.2 km^2/week
Diffusion coefficient	a_{22}	0.2 km^2/week
Birth rate	b(t)	0.096 week^{-1} for $13 \leq t < 28$ 0 week^{-1} otherwise
Natural death rate	μ_1	0.026 week^{-1}
Rabies related death rate	μ_2	0.490 week^{-1}
Rabies transmission rate	β	0.03 $\frac{1}{\text{raccoon/km}^2 \cdot \text{week}}$
Vaccine uptake rate	a	0.01 - 0.03 $\frac{1}{\text{vaccine} \cdot \text{week}}$
Balancing coefficient	c	0.10 $\frac{\text{raccoon/km}^2}{\text{vaccine}^2}$

C.6, and Figure C.7 respectively. Without any vaccination, the density of the infected class at $t = 10$ weeks exceeded 4 raccoons/km^2 in some areas. With the optimal distribution of vaccine, the density of the infected class at $t = 10$ weeks is less than 1 raccoons/km^2 at all locations within our domain, thus we consider the infection successfully eliminated. By implementing the optimal 10-week vaccination strategy upon detection of disease, rabies is contained to the corner of our domain.

The timing of initial infection relative to the birth pulse can determine the intensity of infection and therefore cause variation in the intensity and duration of optimal vaccine allocation. To determine the significance of this variation, we now assume infection is initiated at the end of the birth pulse, at $t = 27$ weeks. At this time, the disease-free susceptible population is at its largest density, 55 raccoons/km^2 for a.e.$(x, y) \in \Omega$. Assuming the initial infection (I_0) occurs at $t = 27$ weeks, we compute the optimal 15-week vaccination regime starting at $t = 27$ weeks. Figure C.8 displays the optimal vaccination and Figure C.9, Figure C.10, and Figure C.11 display the corresponding susceptible, infected, and immune raccoon populations, respectively. Comparing the vaccination regime in Figure C.8 with that in Figure C.4, we find that an increased density of susceptible raccoons at the start of vaccination necessitates an increased quantity of vaccine at the site of infection. Note the increased upper bound in the color scale for Figure C.8. Both vaccination strategies end after the sixth week. However, in the case of infection entering at the end of the birth pulse, it is not until 15 weeks after the start of vaccination that we find a successfully eliminated infected class; Figure C.10. Thus, the timing of initial infection and control implementation in regards to the birth pulse should be considered when determining optimal strategies. For the purpose of comparing strategies, we herein assume initial infection always occurs at $t = 0$ weeks.

Realistically, immediate detection and action is not always possible. Assume infection progresses according the expanding wave displayed in Figures C.2, C.3 and vaccination is not implemented until $t = 21$ weeks. We use the spatial structure of our susceptible and infected populations at $t = 21$ weeks as our initial conditions and compute the optimal 20-week vaccination strategy. We choose to look at intervention over a longer duration ($T = 20$) because of the increased intensity of the infection at the start of intervention. We

simulate two optimal 20-week strategies of vaccination, one for $a = 0.01\frac{1}{\text{vaccine}\cdot\text{week}}$ and another for $a = 0.03\frac{1}{\text{vaccine}\cdot\text{week}}$. Recall, the parameter a is an indicator of how successful distributed baits are in vaccinating raccoons.

First, we assume $a = 0.01\frac{1}{\text{vaccine}\cdot\text{week}}$ and display the optimal 20-week vaccination strategy in Figure C.12. This strategy starts by placing the heaviest amount of vaccine east of the infected population, i.e. in front of the infectious wave. Surrounding areas receive a lesser amount of vaccine, while areas still further away from the infectious wave receive no vaccine. As the weeks progress, the number of baits at each location decreases, but the general pattern of distribution remains the same until $t = 33$ weeks at which time no vaccine is present within the domain. Thus, the optimal vaccine regime lasts approximately 12 weeks. The susceptible, infected, and immune raccoon population densities during this optimal vaccination are displayed in Figure C.13. Figure C.14 and Figure C.15 respectively. It is not until $t = 41$ weeks that we see the infectious class eliminated from our domain, an indication that the infection will not continue to spread.

Second, we assume $a = 0.03\frac{1}{\text{vaccine}\cdot\text{week}}$ and display the optimal 20-week vaccination strategy in Figure C.16. Here, the increased value of a produces a strategy focused on initially placing the greatest quantity of vaccine at the location of the infected population with less vaccine placed in surrounding areas. Instead of vaccinating in front of the infected population as in Figure C.12, this strategy heavily vaccinates the area where infection resides. This vaccination regime is shorter in duration than that in Figure C.12 by 4 weeks. Corresponding susceptible, infected, and immune raccoon populations are displayed in Figure C.17, Figure C.18, and Figure C.19 respectively. As a consequence of the high volume of baits and the increased value of parameter a, the size of the immune class within infected areas quickly increases. With the successful vaccination of most susceptible raccoons at infected locations, rabies has little opportunity to spread. By $t = 33$ weeks, just twelve weeks after the start of intervention, the infected population is eliminated from our domain.

4.4.2 Spatial Heterogeneity and Rabies Spread

We now investigate the effects of spatial heterogeneity and long-distance translocation (LDT) on our optimal strategies of vaccination. To do so, we assume that our domain

encloses a river (running north-south) and that part of our domain is lightly forested. See Figure 4.2 for the heterogeneous design of our spatial domain, Ω. Smith *et al.*'s analysis of Connecticut data indicates rabies spread is substantially slower in lightly forested areas and rivers can slow the spread by as much as 7-fold [78, 79]. For our parameters, we will assume areas without rivers or forest cover have the greatest rates of diffusion, $a_{11} = 0.50$ km^2/week and $a_{22} = 0.50$ km^2/week. Areas with light forest cover will have decreased rates of diffusion, $a_{11} = 0.20$ km^2/week and $a_{22} = 0.20$ km^2/week. Lastly, an area containing the river will have the smallest rate of diffusion in the x-direction, $a_{11} = 0.05$ km^2/week, but have unobstructed diffusion in the y-direction. That is, $a_{22} = 0.50$ km^2/week in areas with a river and no forest cover and $a_{22} = 0.20$ km^2/week in areas with a river and forest cover. These parameters were chosen so that slow movement across the river is possible. We further account for spatial heterogeneity in the initial density of susceptible raccoons. Studies have shown that the density of raccoons in urban areas can be as much as 20-times that of rural settings. Higher raccoon densities in urban areas are likely often related to greater survivorship. Raccoon populations in urban areas are free from intense harvest pressure and may also benefit from stable food and denning resources that mitigate the effects of severe winter conditions. We assume the initial density of susceptible raccoons to be 0 raccoons/km^2 in the areas containing the river, 10 raccoons/km^2 in forested areas, and 30 raccoons/km^2 in areas without forest or river [68]. Figure C.20 illustrates the dynamics of the susceptible population over a 52-week period in the absence of infection. Here, changes in the population density are attributed to natural birth and natural death of susceptible raccoons as well as diffusive movement.

As done previously, we initiate infection in our domain at $t = 0$ weeks with 6 infected raccoons uniformly distributed on the subdomain Ω_I. Table 4.2 displays all parameter values used for the heterogeneous simulations. We first simulate the disease dynamics on our heterogeneous domain in the absence of intervention. Displayed in Figure C.21 and Figure C.22 are eight snapshots of the susceptible and infected raccoon populations over a 52-week period. Additional frames are provided because of the complexity of the spatial spread. As expected, the infection spreads in an irregular pattern within Ω due to the varying rates of diffusion and the heterogeneous initial density of susceptible raccoons. We

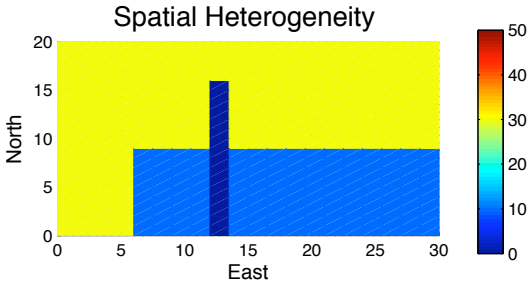


Figure 4.2: Spatial domain Ω is divided by areas with no forest cover (yellow), areas with light forest cover (light blue), and areas containing a river (dark blue). Colors represent initial density of susceptible raccoon population. Diffusive coefficients are chosen to that movement in areas with no forest is greater than movement in lightly forested areas. Movement up and down the river in unrestricted, but movement across the river is significantly inhibited.

Table 4.2: Parameters for heterogeneous rabies spread and control

Parameter	**Notation**	**Value**
Initial susceptible population	$S_0(x, y)$	30 raccoons/km^2 for (x, y) in areas with no forest, no river 10 raccoons/km^2 for (x, y) in areas with forest, no river 0 raccoons/km^2 for (x, y) in areas with river
Initial infected population	$I_0(x, y)$	2 raccoons/km^2 for $(x, y) \in \Omega_I$ 0 raccoons/km^2 for $(x, y) \in \Omega \backslash \Omega_I$
Initial immune population	$R_0(x, y)$	0 raccoons/km^2 for $(x, y) \in \Omega$
Diffusion coefficient (no river, no forest)	a_{11}	0.50 km^2/week
	a_{22}	0.50 km^2/week
Diffusion coefficient (river, no forest)	a_{11}	0.05 km^2/week
	a_{22}	0.50 km^2/week
Diffusion coefficient (no river, forest)	a_{11}	0.20 km^2/week
	a_{22}	0.20 km^2/week
Diffusion coefficient (river, forest)	a_{11}	0.05 km^2/week
	a_{22}	0.20 km^2/week
Birth rate	b(t)	0.096 week^{-1} for $13 \leq t < 28$ 0 week^{-1} otherwise
Natural death rate	μ_1	0.026 week^{-1}
Rabies related death rate	μ_2	0.490 week^{-1}
Rabies transmission rate	β	0.03 $\frac{1}{\text{raccoon/km}^2 \cdot \text{week}}$
Vaccine uptake rate	a	0.01 - 0.02 $\frac{1}{\text{vaccine} \cdot \text{week}}$
Balancing coefficient	c	0.10 - 0.25 $\frac{\text{raccoons/km}^2}{\text{vaccine}^2}$

find that, due to the forest cover and river placement, the infection spreads predominantly northward during the first 30 weeks. By $t = 30$ weeks, we find the bulk of the infected raccoons to be in the northwest corner of Ω. Over the next 7 weeks, the infected class moves eastward, traveling above the river. At $t = 44$ weeks, the infected population mostly resides in the northeast corner of our domain and by $t = 52$ weeks, the infected density is declining, along with the susceptible population. Rabies does not appear to greatly affect the forested area because of the decreased density of susceptible raccoons within this area.

As in the homogeneous case, if the initial infection is detected during the first few weeks, an immediate vaccination regime should stop the spread of rabies. To illustrate this, we assume the same balancing coefficient $c = 0.10$ within the objective functional and display the optimal 10-week vaccination regime starting at $t = 0$ weeks in Figure C.23. We assume $a = 0.02\frac{1}{\text{vaccine}\cdot\text{week}}$, as done in the corresponding homogeneous case (Figure C.4). In comparing Figure C.23 with Figure C.4, we find that with the decreased density and movement associated with the forest cover, the optimal initial quantity of vaccine is reduced and the area over which it is spread is curtailed. The largest quantity of vaccine is placed on the area of initial infection and reduced quantities are placed in adjacent unforested areas. Vaccine is not placed on the forested areas. Vaccination continues until $t = 6$ weeks, at which time we see no vaccine present in our domain. The corresponding susceptible, infected, and immune raccoon populations are displayed in Figure C.24, Figure C.25, and Figure C.26 respectively. By implementing the optimal vaccination strategy upon detection of disease within our spatial domain, the density of infected raccoons is less than 1 raccoons/km^2 by $t = 10$ weeks, indicating successful elimination.

We now consider the situation in which no intervention takes place before $t = 21$ weeks. Using the raccoon densities resulting from Figures C.21, C.22 at $t = 21$ weeks as our initial conditions, we simulate the optimal 20-week vaccination regime for the parameters in Table 4.2 and $a = 0.01\frac{1}{\text{vaccine}\cdot\text{week}}$. See Figure C.27 for the optimal vaccine allocation. Interestingly, we find that the river and forest cover aid the vaccination regime as a natural *cordon sanitaire*, i.e. barrier against disease spread. In the corresponding homogenous case (Figure C.16), vaccine is distributed in the shape of an arc complete from the vertical axis to the horizontal axis of our spatial domain. Now, with the addition of the river and

forest cover, the optimal vaccination regime is considerably curtailed, only forming a partial arc and allowing the natural barriers to act in place of vaccination. We find that initially fortifying the edges of forested areas and areas containing a river with small quantities of vaccine is optimal in reducing spread of infection By $t = 37$ weeks, sixteen weeks after the start of intervention, the optimal vaccination has ended and no vaccine appears on our domain Ω. The corresponding susceptible, infected and removed populations are displayed in Figure C.28, Figure C.29, and Figure C.30 respectively. The vaccination strategy utilizing the river and forest cover as a *cordon sanitaire* is successful in restricting the infected population to the west side of the river. By $t = 41$ weeks, we find no infectious raccoons among our population and have prevented numerous raccoon deaths, as reflected in the large immune and susceptible populations east of the river.

We now assess the influence of LDT in our optimal strategies of vaccination on a heterogeneous domain. LDT events are fairly common. Most are isolated events and do not spread. However, a few LDT events do become new foci of epidemic spread [79] and this is the case we consider. We assume that the infection originating in Ω_I at $t = 0$ weeks progresses without intervention until $t = 21$ weeks as in Figures C.21, C.22. At this time, one additional infected raccoon appears in the northeast corner of our spatial domain, Ω. In the continued absence of intervention, the existing epidemic and newly added epidemic focus will progress as in Figure C.31 and Figure C.32. The LDT event produces a separate expanding wave of disease. The two fronts converge by $t = 37$ weeks and all infection appears to be subsiding at $t = 41$ weeks due to a diminishing susceptible population.

Using the parameters in Table 4.2 with $a = 0.01$, we simulate the optimal 20-week vaccination regime beginning at $t = 21$ weeks, the time corresponding to the LDT event. Optimal allocation of vaccine is displayed in Figure C.33. Even though the LDT is small, i.e. one infected raccoon, it is an immediate concern. The optimal vaccination regime immediately applies the largest concentration of vaccine on the site of the LDT. For our objective functional and cost coefficients, it is cost-effective to prevent the new focus from becoming a widespread epidemic while also vaccinating near the existing infectious wave. Even with the intense local vaccination, the LDT event does produce several secondary infections. See Figure C.35 for the infectious population dynamics during vaccination.

However, with the two-part vaccination regime, the infectious wave produced by the LDT is limited. Further, the existing outbreak is mitigated by the barrier created by the vaccine in combination with the river and forest cover. By $t = 41$ weeks, successful elimination of the infected class is visible and we have maintained a viable susceptible and immune population. Corresponding susceptible and immune populations can be see in Figure C.34 and Figure C.36 respectively.

As a last exercise in addressing the role of LDT events in vaccination regimes, we alter the value of c, the parameter in the objective functional balancing the squared cost of vaccination with the cost associated with the size of the infectious population. In all previous runs, we have assumed the balancing coefficient to be $c = 0.10$. Now we will assume a higher cost associated with vaccination by taking $c = 0.25$. All other parameters stay the same as in Table 4.2 with $a = 0.01$ and we again assume the LDT event occurs at $t = 21$ weeks. The optimal 20-week vaccination regime starting at $t = 21$ weeks is displayed in Figure C.37. Corresponding susceptible, infected, and immune raccoon populations are displayed in Figure C.38, Figure C.39, and Figure C.40 respectively. We find that the increased cost of vaccination significantly constrains the optimal amount of vaccine used. However, the vaccine placement pattern in Figure C.37 is similar to that in Figure C.33. With the higher cost, it remains optimal to place largest amount of vaccine on the LDT event and place smaller quantities on the existing outbreak. However, with the reduced amount of vaccine, the optimal strategy in Figure C.37 does not contain disease spread and is unsuccessful in eliminating the infected population over the 20-week period. At $t = 41$ weeks, infected racoons still reside in the northeast corner of our domain. These results suggest the existence of a threshold level of population immunity necessary to contain rabies spread [68].

4.4.3 Conclusions of Numerical Simulations

General patterns of optimal vaccination are realized from the study of both homogeneous and heterogenous spatial domains. In studying the simple homogeneous domain, our rabies outbreak takes the form of a uniform infectious wave, expanding at a constant rate. The heterogeneous domain provides a more realistic scenario by including rivers and forest cover,

both of which inhibit a raccoon's movement and are associated with decreased raccoon density. On such a domain, disease spreads irregularly and at varying rates. Results for both domain types indicate that early detection of rabies along with a quickly implemented vaccination regime can prevent a large epidemic. Vaccination regimes implemented immediately after the birth pulse should be heightened to accommodate the increased density of susceptible raccoons.

Realistically, immediate detection and implementation are not always feasible. Vaccination regimes that begin several months after infection enters the area may take different forms depending on the vaccine uptake rate. In areas where baits are less likely to be removed by humans or consumed by animals other than raccoons, we find it optimal to place the heaviest amounts of vaccine on infected areas with decreased amounts nearby. In areas assumed to have a decreased vaccine uptake rate, it becomes optimal to vaccinate in front of the infectious wave. Rivers and forest cover, or any natural barrier inhibiting raccoon movement and/or density, can aid in reducing the optimal amount of vaccine used. Acting as a *cordon sanitaire*, forested areas and rivers fortified with small amounts of vaccine can be as successful in preventing widespread outbreaks as would large amounts of vaccine in a homogenous domain. Thus, in designing cost-effective strategies for intervention, heterogeneous domain characteristics that may influence movement and density should be regarded as possible aids.

Long-distance raccoon translocation (LDT) can impact the shape and speed of an epidemic front. Although initially small in size, LDT events can become new epidemic foci and their appearance should be incorporated into a vaccination plan. For best results, an LDT event occurring at the beginning of vaccination should be addressed as an infection separate from the existing infectious wave. Heavy amounts of vaccine placed on and around the new focus can prevent a secondary infectious wave. At the same time, a separate vaccine regime should be implemented to contain the existing front. An increase in the cost of control reduces the optimal quantity of vaccine present at any one location and time. Results show that reduced quantities of vaccine may not be effective in containing disease spread or eliminating the infected population.

4.5 Concluding Remarks and New Contributions

In this chapter, we present the first application of optimal control theory to spatiotemporal epidemic models described by a system of partial differential equations. We prove the existence and uniqueness of the solutions to our parabolic state system as well as prove the existence of an optimal control, sensitivity system, and adjoint system. For a given objective functional, the optimal control is characterized in terms of the states and adjoint variables. Lastly, uniqueness of the optimality system is shown for sufficiently small time.

We solve the optimality system numerically with parameters chosen to model the spread of rabies within a raccoon population. Although partial differential equations have previously been used to model rabies spread among raccoons, optimal control theory was not applied to these models. Our solutions indicate that the optimal timing and placement of rabies vaccine baits can vary with the spatial structure of a domain. Natural land features which deter raccoon movement (such as a river or forest cover) act as a barrier against the spread of rabies. In our simulations, these barriers aided the optimal vaccination strategies considerably. Vaccination regimes implemented immediately after the birth pulse should be heightened to accommodate the increased density of susceptible raccoons. We also simulate long distance translocation of an infected raccoon. Optimal vaccination strategies suggest that a new epidemic focus resulting from translocation should be addressed immediately with heavy amounts of vaccine. The methods we present in this chapter provide a new and useful tool in addressing the control of disease spreading in both space and time. In the future, we can consider similar results for systems of PDEs with alternative terms modeling disease incidence as well as results for systems which also include an additional state equation describing control dynamics.

Chapter 5

Summary and Future Directions

In summary, this dissertation contributes new results in the theory of optimal control as well as advancements in mathematical modeling and applications of optimal control. A novel existence result of an optimal control was proven in the case of ordinary differential state equations containing quadratic expressions of the control variable. In the case of partial differential state equations, we showed new existence results for both the state solutions and the optimal control. Furthermore, in the PDE case, we derived the sensitivity and adjoint systems and characterized a unique optimal control.

We applied our theoretical results to scenarios specific to the control of a harmful population. From the numerical results, a new framework for the design of intervention policies arises. In Chapter 2, we illustrated that our approach can be a useful tool to aid in the management of native and invasive populations. In Chapter 3, we presented the first application of optimal control theory to a cholera model. The model formulation, sensitivity analyses, optimal control problem, and associated solution provide a new methodology to evaluate trade-offs in multiple intervention schemes and assess which are most effective. In Chapter 4, we introduced the first application of optimal control theory to a parabolic PDE epidemic model. Numerical results illustrate this approach as an appropriate tool for determining when and where to place oral rabies vaccine in order to successfully contain the spread of rabies among raccoons.

Possible future work to be developed from Chapter 2 includes investigating and proving other existence results for state systems containing non-linear expressions of the control

variable. For the spatiotemporal epidemic model in Chapter 4, we will work to overcome the restriction of T being sufficiently small by proving existence of the state solution for all time T.

Epidemic models, including those for cholera, are constantly evolving. Thus, there is continued need for optimal control theory in determining appropriate strategies of intervention. The results presented in Chapter 3 will aid in formulating an age-structured or spatially explicit cholera model. Applying optimal control to these extended models will provide insight as to which age groups or locales should receive priority in treatment. Our spatiotemporal epidemic model in Chapter 4 was simplified to accommodate the rabies and raccoons example, but the inclusion of a different disease incidence term or a different vaccine uptake term can provide interesting dynamics applicable to other infectious diseases and can lead to different optimal allocations of control. We continue to explore these possibilities and to make advancements in mathematical modeling and optimal control theory.

Bibliography

[1] S. Alam and S. Bhatnagar, Current status of anti-diarrheal and anti-secretory drugs in the management of acute childhood diarrhea, Indian Pediatr., 73(2006), pp. 693-696.

[2] Ali, M, Emch, M, *et al.*, Herd immunity conferred by killed oral cholera vaccines in Bangladesh: A reanalysis, Lancet, July 2-8, 2005; 366, 9479; pp. 44-49.

[3] V. Arnautu, V. Barbu, and V. Capasso, Controlling the spread of a class of epidemics, Appl. Math. Optim., 20(1989), pp. 297-317.

[4] J. E. Childs, A. T. Curns, M. E. Dey, L. A. Real, L. Feinstein, O. N. Bjornstad and J. W. Krebs, Predicting the local dynamics of epizootic rabies among raccoons in the United States, Proc Natl Acad Sci USA V. 97, 25(2000), pp. 13666-13671.

[5] V. Barbu, Mathematical Methods in Optimization of Differential Systems, Kluwer Academic, Dordrecht, 1994.

[6] V. Barbu and M. Iannelli, Controlling the S-I-S age- structured epidemics, Mathematical Models and Health Sciences, Vanderbilt Press, 1998, pp. 29-34.

[7] V. Barbu and M. Iannelli, Optimal control of population dynamics, J. Optim. Theory Appl., 102(1999), pp. 1-14.

[8] B. Beckage, L. J. Gross, and W. J. Platt, Modeling responses of pine savannas to climate change and large-scale disturbance, Appl. Veg. Sci., 9(2006), pp. 75-82

[9] M. Bendahmane, M. Langlais, and M. Saad, Existence of solutions for reaction-diffusion systems with L^1 data, Adv. Differential Equations, 6(2002), pp. 743-768.

[10] M. Bendahmane, M. Langlais, and M. Saad, On some anistropic reaction-diffusion systems with L^1-data modeling the propagation of an epidemic disease, Nonlinear Anal., 54(2003), pp. 617-636.

[11] S. M. Blower and H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example, Int. Stat. Rev. 2(1994), pp. 229-243.

[12] "Bogra." The Encyclopaedia Britannica, 11th ed. (1911)

[13] W. J. Bond, G. F. Midgley, and F. J. Woodward, The importance of CO-2 and fire in promoting the spread of grasslands and savannas, Global Change Biol., 8(2003), pp. 973-982.

[14] J. M. Bony, Principe du nmaxinmunm dants les espaces de Sobole,. C. R. Acad. Sci. Paris, A265(1967), pp. 333-336.

[15] M. Brokate, Pontryagin's principle for control problems in age-dependent population dynamics, J. Math. Bio. 23(1985), pp. 75-101.

[16] M. L. Brooks and D. A. Pyke, Invasive plants and fire in the deserts of North America, IN: Galley, K.E.M. and T.P. Wilson (eds.), Proceedings of the Invasive Species Workshop: The Role of Fire in the Spread and Control of Invasive Species. Fire Conference 2000: The First National Congress on Fire Ecology, Prevention, and Management, Tallahassee, FL, Tall Timbers Research Station, 11(2001), pp. 1-14.

[17] V. Capasso, Mathematical Methods for the control of infectious diseases, Lecture Notes in Control and Information Sciences, Springer - Verlag, Heidelberg , Berlin, 1987, pp. 162-174.

[18] V. Capasso and S. L. Paveri-Fontana, A mathematical model for the 1973 cholera epidemic in the european mediterranean region, Rev. Eipém. et. Anté Publ. 27(1979), pp. 121-132.

[19] T. Clayton, S. Duke-Sylvester, L. J. Gross, S. Lenhart, and L. A. Real, Optimal Control of a Rabies Epodemic Model with a Birth Pulse, J. Bio. Dyn. (preprint).

[20] C. T. Codeço, Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir, BMC Inf. Dis. 1(2001).

[21] R. Collwell, Global climate and infectious disease: The cholera paradigm, Science, 274(Dec 20, 1996), pp. 2025-2032.

[22] R. Collwell, A. Huq, M. Islam, *et. al*, Reduction of cholera in Bangladeshi villages by simple filtration, PNAS, 100(2003), pp. 1051-1055.

[23] R. M. Conroy, M. E. Meegan, *et. al*, Solar disinfeciton of drinking water protects against cholera in children under 6 years of age, Arch. Dis. Child. 85(2001), pp. 293-295.

[24] F. Courchamp and S. Caut, Use of biological invasions and their control to study the dynamics of interacting populations, Conceptual Ecology and Invasions Biology, M.W. Cadotte, S.M. McMahon and T. Fukami (eds), Springer, Berlin, pp. 253-279, 2005.

[25] M. J. Coyne, G. Smith, and F. E. McAllister, Mathematic Model for the Population Bi- ology of Rabies in Raccoons in the Mid-Atlantic States, Am. J. Vet. Res., 50(1989), pp. 2148-2154.

[26] W. Ding, L. Gross, K. Langston, S. Lenhart, and L. Real, Rabies in raccoons: optimal control for a discrete time model on a spatial grid, J. Bio. Dyn., 1(2007), pp. 379-393.

[27] P. van den Driessche and J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci., 180(2002), pp. 29-48.

[28] C. Dunston, D. McAfee, R. Kaiser *et al.*, Colalboration, cholera, and cyclones: A profject to improve point-of-use water quality in Madagascar, Am. J. Public Health, 91(2001), pp. 1574-1576.

[29] I. Ekeland, On the variational principle, J. Math. Anal. Appl. 47(1974), pp. 324-353.

[30] I. Ekeland and R. Temam. Convex Analysis and Variational Problems, North-Holland Publishing Company, Amsterdam, 1944, SIAM, Philadelphia, PA, 1999.

[31] L. C. Evans, Partial Differential Equations, American Mathematical Society, Providence, RI, 1998.

[32] N. D. Evans and A.J. Pritchard, A control theoretic approach to containing the spread of rabies, IMA J. Math. Appl. Med., 18(2001), pp. 1-23.

[33] W. E. Fitzgibbon and J.J. Morgan, A diffusive epidemic model on a bounded domain of arbitrary dimension, Differ. Integral Equ., 1(1988), pp. 125-132.

[34] W. H. Fleming and R. W. Rishel, Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 2005.

[35] M. Ghosh, P. Chandra, P. Sinha, and J. B. Shukla, Modeling the spread of carrier-dependent infectious diseases with environmental effect, Appl. Math. Comp., 152(2004), pp. 385-402.

[36] E. Gotuzzo, C. Seas, *et al.* Ciprofloxacin for the treatment of cholera: A radnomized, double-blind, controlled clinicial trial of a single daily dose in Peruvian adults, Clin. Infect. Dis. 20(1995), pp. 1485-1490.

[37] D. M. Hartley, J. G. Morris, and D. L. Smith, Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics?, PLoS Medicine, 3(2006), pp. 63-69.

[38] T. R. Hendrix, The pathophysiology of cholera, Bull NY Acad Med 47(1971), pp. 1169 -1180.

[39] D. A. Jameson, Effects of burning on a Galleta-Black Grama range invaded by Juniper, Ecol., 43(1982), pp. 760-763.

[40] A. Kallen, P. Arcuri, and J. D. Murray, A simple model for the spatial spread and control of rabies. J. Theor. Biol., 116(1985), pp. 377-393.

[41] J. B. Kaper, J. G. Morris, and M. M. Levine, Cholera, Clin. Microbiol. Rev. 8(1995), pp. 48-86.

[42] A. King, E.L. Ionides, M. Pascual, and M. Bouma, Inapparent infections and cholera dynamics, Nature 454(14 August 2008) pp. 877-880.

[43] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, J. Theor. Bio. 254(2008), pp. 178-196.

[44] K. Koelle, M. Pascual, and M. Yunus, Pathogen adaptation to seasonal forcing and climate change, Proc. R. Soc. B. 272(2005), pp. 971-977.

[45] K. Koelle, X. Rodó, M. Pascual, *et al.* Refractory periods and climate forcing in cholera dynamics, Nature, 436(2005) pp. 696-700.

[46] N. V. Krylov, Nonlinear Elliptic and Parabolic Equations of the Second Order, D. Reidal, Holland, 1987.

[47] E. B. Lee and L. Markus, Foundations of Optimal Control Theory, John Wiley, New York, 1968.

[48] D. Legros, C. Paquet, W. Perea, *et al.*, Mass vaccination with a two-dose oral cholera vaccine in a refuge camp, Bull. of the World Health Organization, 77(1999), pp. 837-842.

[49] S. Lenhart and J. Workman, Optimal Control Applied to Biological Models, Boca Raton, Chapmal Hall/CRC, 2007.

[50] M. M. Levine, J. B. Kaper, D. Herrington, G. Losonsky, J. G. Morris, M. L. Clements, R. E. Black, B. Tall, and R. Hall, Volunteer studies of deletion mutants of *Vibrio cholerae* 01 prepared by recombinant techniques, Infect. and Immun. (1988), pp. 161-167.

[51] X. Li and J. Yong, Optimal Control Theory for Infinite Dimensional Systems, Birkhäuser, Boston, 1995.

[52] J. L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, New York, 1971.

[53] D. M. Lodge, Biological invasions: Lessons for ecology, TREE, 8(1993), pp. 133-137.

[54] D. L. Lukes, Differential Equations: Classical to Controlled, Academic Press, New York, 1982.

[55] D. S. Merrell, S. M. Butler, Host-induced epidemic spready of the cholera bacterium, Nature, 417(2002) pp. 642-645.

[56] J. G. Morris, *Vibrio cholerae* 0139 Bengal: Emergence of a new epidemic strain of cholera, Infect. Agents Dis. 4(2001), pp. 41-46.

[57] J. D. Murray and W. L. Seward, On the spatial spread of rabies among foxes with immunity, J. Theo. Biol., 130(1991), pp. 327-348.

[58] J. D. Murray, E. A. Stanley, and D. L. Brown. On the spatial spread of rabies among foxes, Proc. R. Soc. London B, 229(1986), pp. 111-150.

[59] J. H. Myers, D. Simberloff, A. M. Kuris, and J. R. Carey, Eradication revised: dealing with exotic species, TREE, 15(2000), pp. 316-320.

[60] A. Naficy *et al.*, Treatment and vaccination strategies to control cholera in sub-Saharan refugee settings: A cost-effectiveness analysis, JAMA 279(1998), pp. 521-525.

[61] L. Nian, J. Steiner, and S. Tang, Convergence and Stability Analysis of an explicit Finite Difference Method for 2-Dimensional Reaction Diffusion Equations, J. Austral. Math. Soc. 36(1994), pp. 234 - 241.

[62] M. Pascual, M. Bouma, and A. Dobson, Cholera and climate: Revisiting the quantitative evidence, Microbes Infect., 4(2002), pp. 237-245.

[63] M. Pascual, K. Koelle, and A.P. Dobson, Hyperinfectivity in cholera: A new mechanism for an old epidemiological model? PLoS Med 3(6): e280 doi:10.1371/journal.pmed.0030280.

[64] D. P. C. Peters, B. T. Bestelmeyer,J. E. Herrick, E. L. Fredrickson, H. C. Monger, and K. M. Havstad, Disentangling complex landscapes: new insights into arid and semiarid system dynamics, Bioscience, 56(2006), pp. 491-501.

[65] O. Pollack and T. Kan, The use of prescribed fire to control invasive exotic weeds at Jepson Prairie Preserve, IN, 1998, C. W. Witham, E. T. Bauder, D. Belk, W. R. Ferren, Jr., and R. Ornduff, [eds.], Ecology, Conservation, and Management of Vernal Pool Ecosystems, Sacramento, CA: Proceedings California Native Plant Society,1996, pp. 241-249.

[66] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelize, and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Wiley, New York, 1967.

[67] E. Pourabbas, A. d'Onofrio, and M. Rafanelli, A method to estimate the incidence of communicable diseases under seasonal fluctuations with application to cholera, Appl. Math. Comp., 118(2001), pp. 161-174.

[68] S. P. D. Riley, J. Hadidian, and D. A. Manski, Population density, survival, and rabies in raccoons in an urban national park, Can. J. Zool. 76(1998), pp. 11531164.

[69] C. A. Russell, D. L. Smith, J. A. Child, and L. A. Real, Predictive spatial dynamics and strategic planning for raccoon rabies emergence in Ohio, PLoS Biology 3(3): e88(2005).

[70] H. H. Risley and E.A. Gait, Report on the census of India, 1901, Superintendent of Government Printing, Calcutta (1903), pp. 30.

[71] S. P. Sethi and G. L. Thompson, Optimal Control Theory: Applications to Management Science and Economics, Kluwer, Boston, 2nd edition, 2000.

[72] A. A. Sher and D. L. Marshall, Seedling competition between native *Populus deltoides* (Salicaceae) and exotic *Tamarix ramoisissima* (Tamaricaceae) across water regimes and substrate types, Am. J. Bot., 90(2003), pp. 413-422.

[73] A. A. Sher, D. L. Marshall and S. A. Gilbert, Competition between *Populus deltoides* and invasive *Tamarix ramoisissima* and the implications for reestablishing flooding disturbance, Conserv. Biol., 14(2000), pp. 1744-1754.

[74] A. A. Sher, D. L. Marshall and J. P. Taylor, Establishment patterns of native *Populus* and *Salix* in the presence of invasive *Tamarix*, Ecol. Appl., 12(2002), pp. 760-772.

[75] N. Shigesada and K. Kawasaki, Biological Invasions: Theory and Practice, New York, Oxford University Press, 1997.

[76] J. Simon, Compact sets in the space $L^p(0,T;B)$, Annali di Matematica Pura ed Applicata, V. 146, 1(1986), pp. 65-96.

[77] D. Slate, C. E. Rupprecht, J. A. Rooney, D. Donovan, D. H. Lein, and R. B. Chipman, Status of oral rabies vaccination in wild carnivores in the United States, Virus Res. 111(2005), pp. 68-76.

[78] D. L. Smith, B. Lucey, L. A. Waller, J. E. Childs, and L. A. Real, Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes, PNAS (2002), pp. 3668-3672.

[79] D. L. Smith, L. A. Waller, C. A. Russell, J. E. Childs and L. A. Real, Assessing the role of long-distance translocation and spatial heterogeneity in the raccoon rabies epidemic in Connecticut, Prev. Vet. Med. (2005).

[80] J. Stromberg, Dynamics of Fremont cottonwood (Populus fremonti) and saltcedar (Tamarix chinensis) populations along the San Pedro River, Arizona, J. Arid Environ. 40(1998), pp. 133-155.

[81] N.G. Tallent-Halsell and L. R. Walker, Responses of *Salix Gooddingii* and *Tamarix Ramoissima* to flooding, WETLANDS, 22(2002), pp. 776-785.

[82] J. P. Taylor, L. M. Smith, and D. A. Haukos, Evaluation of woody restoration in the middle Rio Grande: Ten years after, WETLANDS 26(2002), pp. 1151-1160.

[83] D. Tilman, Plant Strategies and the Dynamics and Structure of Plant Communities, Princeton University Press, Princeton, 1988.

[84] M. H. Topps, Oral cohlera vaccine–for whom, when, and why?, Travel Medicine and Infectious Disease 4(2006), pp. 38-42.

[85] J. H. Vandermeer, The Ecology of Intercropping, Cambridge Press, Cambridge, 1989.

[86] C. K. Wallace, P.N. Anderson, T.C. Brown, Optimal Antibiotic Therapy in Cholera, Bull. Wld. Hlth Org, 39(1968), pp. 239-245.

[87] G. F. Webb, A reaction-diffusion model for a deterministic diffusive epidemic, J. Math. Anal. Appl., 84(1981), pp. 150-161.

[88] G. F. Webb, Theory of Nonlinear Age-Dependent Population Dynamics, Marcel Dekker, New York, 1985.

[89] R. Westbrooks, et al., Invasive Plants, Changing the Lands of America, Fact Book, Washington, D.C., Federal Interagency Committee for the Management of Noxious and Exotic Weeds, U.S. Dept of the Interior, 1998.

[90] World Health Organization, "Cholera" <http://www.who.int/topics/cholera>.

[91] World Health Organization, Cholera Fact Sheet No. 107, November 2008.

[92] World Health Organization, "Cholera in Zimbabwe." Disease Outbreak News, December 2008, <http://www.who.int/csr/don/2008_12_02>.

[93] World Health Organization Global Task Force on Cholera Control, Oral cholera vaccine use in complex emergencies: What next?, WHO Meeting, Cairo, Egypt (2005). <http://www.who.int/cholera/publications/cholera_vaccines_emergencies_2005.pdf>.

[94] J. Yong, Optimal Control for Nonlinear Abstract Evolution Systems, Diff. Int. Eqs. 6 (1993), pp. 1145-1159.

[95] K. Yosida, Functional Analysis, Springer-Verlag, Berlin, 1980.

Appendices

Appendix A

Numerical Methods

Runge Kutta Method

Consider the problem of approximating the solution $y(t)$ to the ordinary differential equation and initial condition

$$\frac{dy}{dt} = f(t, y) \qquad \forall t \in [0, T] \tag{A.1}$$

$$y(0) = y_0 \tag{A.2}$$

where f is a continuous function and y_0 is a real number.

We partition our time domain $[0, T]$ into K equal intervals of length $\Delta t = T/K$ and denote the approximate value of y at time step k as y_k for $0 \leq k \leq K$.

The Runge Kutta method of the fourth order determines the value of y_{k+1} by

$$y_{k+1} = y_k + \frac{1}{6}\Delta t \cdot (m_1 + 2m_2 + 2m_3 + m_4)$$

where

$$
\begin{aligned}
m_1 &= f(k\Delta t, y_k), \\
m_2 &= f(k\Delta t + \frac{1}{2}\Delta t, y_k + \frac{1}{2}\Delta t \cdot m_1), \\
m_3 &= f(k\Delta t + \frac{1}{2}\Delta t, y_k + \frac{1}{2}\Delta t \cdot m_2), \\
m_4 &= f((k+1)\Delta t, y_k + \Delta t \cdot m_3).
\end{aligned}
$$

Thus, the value of y_{k+1} is determined by the present value y_k plus the product of the size of the interval Δt and a weighted average of slopes, m_i.

The following is an algorithm using the Runge Kutta method of the fourth order to approximate the solution, $y(t)$, to the boundary value problem in (A.1)-(A.2).

Algorithm 1 Runge Kutta 4 Method

Require: y_0, T, K
Ensure: y_k for $1 \leq k \leq K$.

dt = T/K
$t_k = k * dt$

for $k = 0$ to $K - 1$ **do**
$m_1 = f(t_k, y_k)$
$m_2 = f(t_k + \frac{dt}{2}, y_k + \frac{m_1}{2})$
$m_3 = f(t_k + \frac{dt}{2}, y_k + \frac{m_2}{2})$
$m_4 = f(t_{k+1}, y_k + m_3)$

$y_{k+1} = y_k + \frac{1}{6} * dt * (m_1 + 2 * m_2 + 2 * m_3 + m_4)$

end for

Finite Difference Method

Consider the problem of approximating the function $u(x,t)$ satisfying

$$\frac{du}{dt} - \alpha \Delta u = f(x,y) \qquad \forall (x,y,t) \in \Omega = (0,1)\times(0,1)\times(0,T] \tag{A.3}$$

$$\frac{du}{dx} = 0, \qquad \frac{du}{dy} = 0 \qquad \forall (x,y,t) \in \bar{\Omega} \tag{A.4}$$

$$u(x,y,0) = \bar{u}(x,y) \tag{A.5}$$

where f and $\bar{u}$ are smooth functions and $\bar{u}$ satisfies the no-flux boundary conditions.

We discretize Ω such that there is an integer number of points in space and an integer number of points in time at which we will calculate u. We assume equal grid spacing of the variables so that $\Delta x = x_m - x_{m-1}$ for $1 \le m \le M$, $\Delta y = y_n - y_{n-1}$ for $1 \le n \le N$, and $\Delta t = t_k - t_{k-1}$ for $1 \le k \le K$. Approximate the functions u, f and $\bar{u}$ by $u_{m,n,k} = u(x_m, y_n, t_k)$, $f_{m,n} = f(x_m, y_n)$ and $\bar{u}_{m,n} = \bar{u}(x_m, y_n)$. Based on this discretization and function approximations, we then write the following approximations of derivatives in time and space

$$\frac{\partial u}{\partial t}|_{x_m,y_n,t_k} \approx \frac{u_{m,n,k+1} - u_{m,n,k}}{\Delta t}$$

$$\frac{\partial u}{\partial x}|_{x_m,y_n,t_k} \approx \frac{u_{m+1,n,k} - u_{m,n,k}}{\Delta x}$$

$$\frac{\partial u}{\partial y}|_{x_m,y_n,t_k} \approx \frac{u_{m,n+1,k} - u_{m,n,k}}{\Delta y}.$$

We can take the spatial derivatives one step further by taking the differences of the derivative approximations to arrive at an approximation for the second derivatives

$$\frac{\partial^2 u}{\partial x^2}|_{x_m,y_n,t_k} \approx \frac{u_{m-1,n,k} - 2u_{m,n,k} + u_{m+1,n,k}}{(\Delta x)^2},$$

$$\frac{\partial^2 u}{\partial y^2}|_{x_m,y_n,t_k} \approx \frac{u_{m,n-1,k} - 2u_{m,n,k} + u_{m,n+1,k}}{(\Delta y)^2}.$$

Inserting the finite difference approximations into (A.3) and rearranging, we get

$$\frac{u_{m,n,k+1} - u_{m,n,k}}{\Delta t} = \alpha(\frac{u_{m-1,n,k} - 2u_{m,n,k} + u_{m+1,n,k}}{(\Delta x)^2} + \frac{u_{m,n-1,k} - 2u_{m,n,k} + u_{m,n+1,k}}{(\Delta y)^2}) + f_{m,n}.$$

Solving for $u_{m,n,k+1}$, we have

$$u_{m,n,k+1} = u_{m,n,k} + \Delta t(\alpha(\frac{u_{m-1,n,k} - 2u_{m,n,k} + u_{m+1,n,k}}{(\Delta x)^2} + \frac{u_{m,n-1,k} - 2u_{m,n,k} + u_{m,n+1,k}}{(\Delta y)^2}) + f_{m,n}). \tag{A.6}$$

If we construct the 2D spatial mesh first, we can choose Δt to satisfy the stability criterion $\Delta t \leq \frac{(\Delta x)^2+(\Delta y)^2}{8\alpha}$. The following is an algorithm applying these techniques to approximate the solution to (A.3) - (A.5).

Algorithm 2 Explicit Finite Difference Method

Require: α, M, N
Require: $f_{m,n}$, $u_{m,n,0}$ for $0 \leq m \leq M$, $0 \leq n \leq N$
Ensure: $u_{m,n,k}$ for $0 \leq m \leq M$, $0 \leq n \leq N$, $1 \leq k \leq K$.

dx = 1/(M+1)
dy = 1/(N+1)
K = roundup($(8T\alpha)/(dx^2 + dy^2)$)
dt = T/(K+1)
for $k = 0$ to $K - 1$ **do**
  for $n = 1$ to $N - 1$ **do**
    for $m = 1$ to $M - 1$ **do**

$$u_{m,n,k+1} = u_{m,n,k} + dt * (\alpha * ((u_{m-1,n,k} - 2u_{m,n,k} + u_{m+1,n,k})/(dx^2) + ... (u_{m,n-1,k} - 2u_{m,n,k} + u_{m,n+1,k})/(dy^2)) + f_{m,n})$$

    end for
  end for
  for $m = 1$ to M **do**
    $u_{m,0,k+1} = u_{m,1,k+1}$
    $u_{m,N,k+1} = u_{m,N-1,k+1}$
  end for
  for $n = 1$ to N **do**
    $u_{0,n,k+1} = u_{1,n,k+1}$
    $u_{M,n,k+1} = u_{M-1,n,k+1}$
  end for
end for

Appendix B

Previous Cholera Models

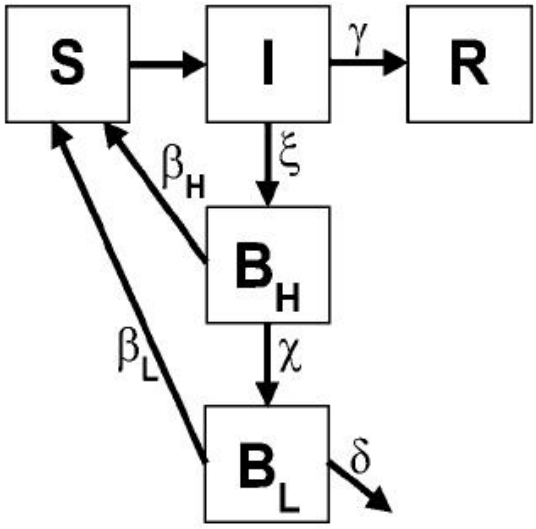


Figure B.1: Diagram illustrating the model of Hartley *et al.* [37] which distinguishes between hyperinfectious vibrios (B_H) and less-infectious vibrios (B_L).

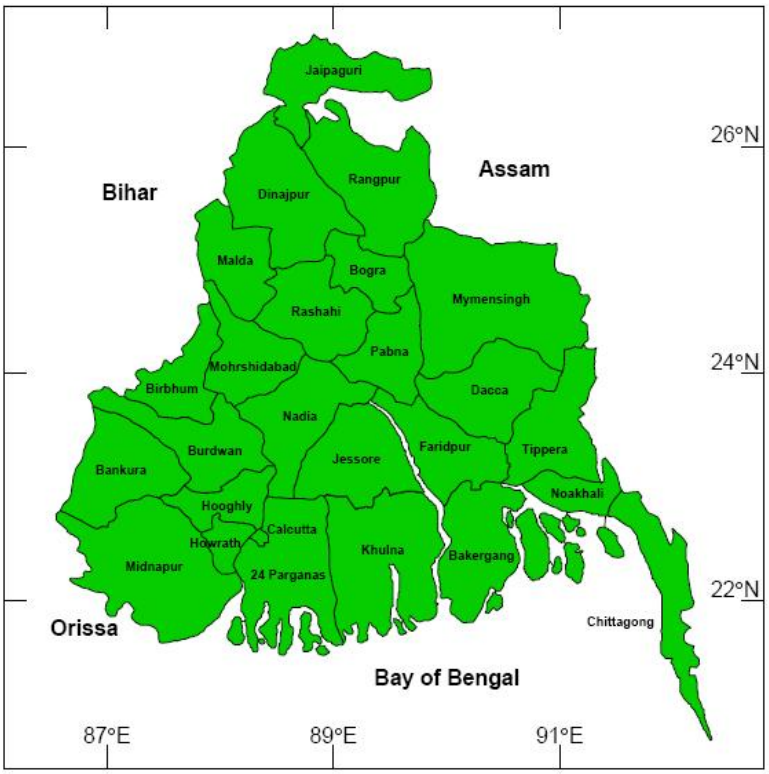


Figure B.2: Map of Bengal Bay districts from King *et al.* [42]

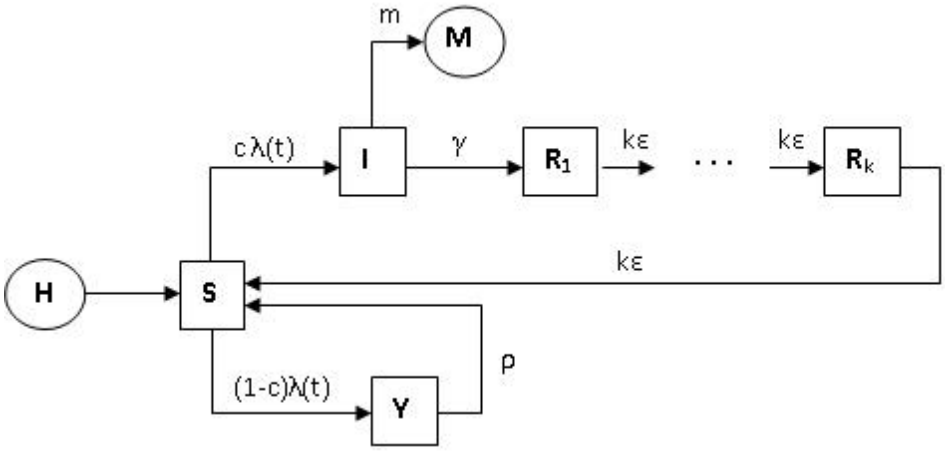


Figure B.3: Diagram illustrating the two-path model of King *et al.* [42] in which Y class represents mild, asymptomatic infections. Infected individuals die at an excess rate m year^{-1} and recover at a rate γ year^{-1}. An individual remains immune to reinfection for a duration gamma-distributed with mean $1/\epsilon$ years and variance $1/k\epsilon^2$ years^2. The mean duration of short term immunity is $1/\rho$ years.

Table B.1: Maximum likelihood estimates for two-path model in King *et al.* [42] for $k = 3$. Displayed are estimates for the six of the 26 Bengal districts which were best described by the two-path model when compared to 3 other models. The units of γ, m, ϵ, and ρ are year^{-1} and the other quantities are dimensionless.

District	$logL$	AIC_C	R_0	γ	m	ρ	ϵ	c
Bakergang	-3656.3	7354.0	1.38	7.1	2.740	8.5	0.5	0.0088
Bogra	-2837.5	5716.4	2.30	21.6	0.240	2.6	0.4	0.2400
Calcutta	-3035.4	6112.3	1.27	11.3	4.662	5.6	0.7	0.0104
Noakhali	-3471.7	6984.9	1.71	9.5	3.103	4.4	0.5	0.0108
Pabna	-3283.8	6609.0	1.47	12.3	13.756	4.4	0.8	0.0059
Rangpur	-3264.2	6569.8	1.55	17.3	0.134	14.0	1.4	0.3894

Appendix C

Rabies Simulations

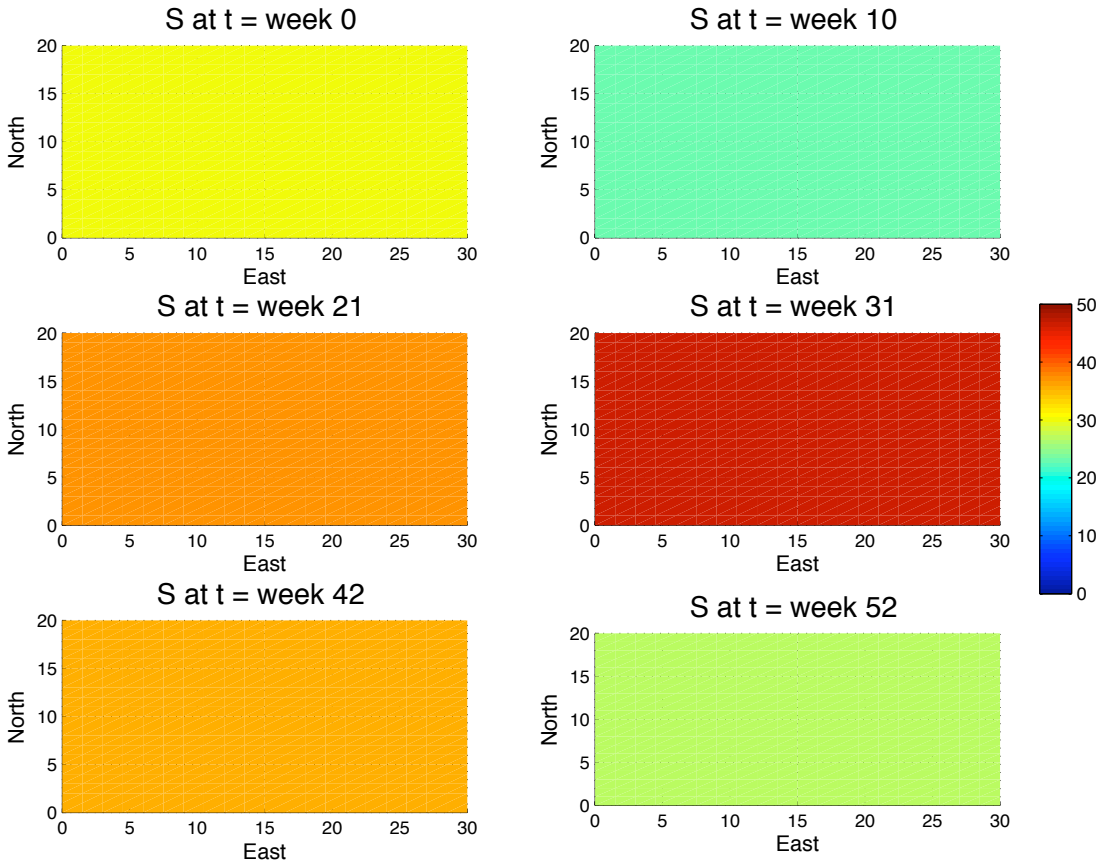


Figure C.1: Spatial Homogeneity: Susceptible population over 52 weeks in the absence of any infection. Growth and decline of population indicate natural birth and death of susceptible raccoons.

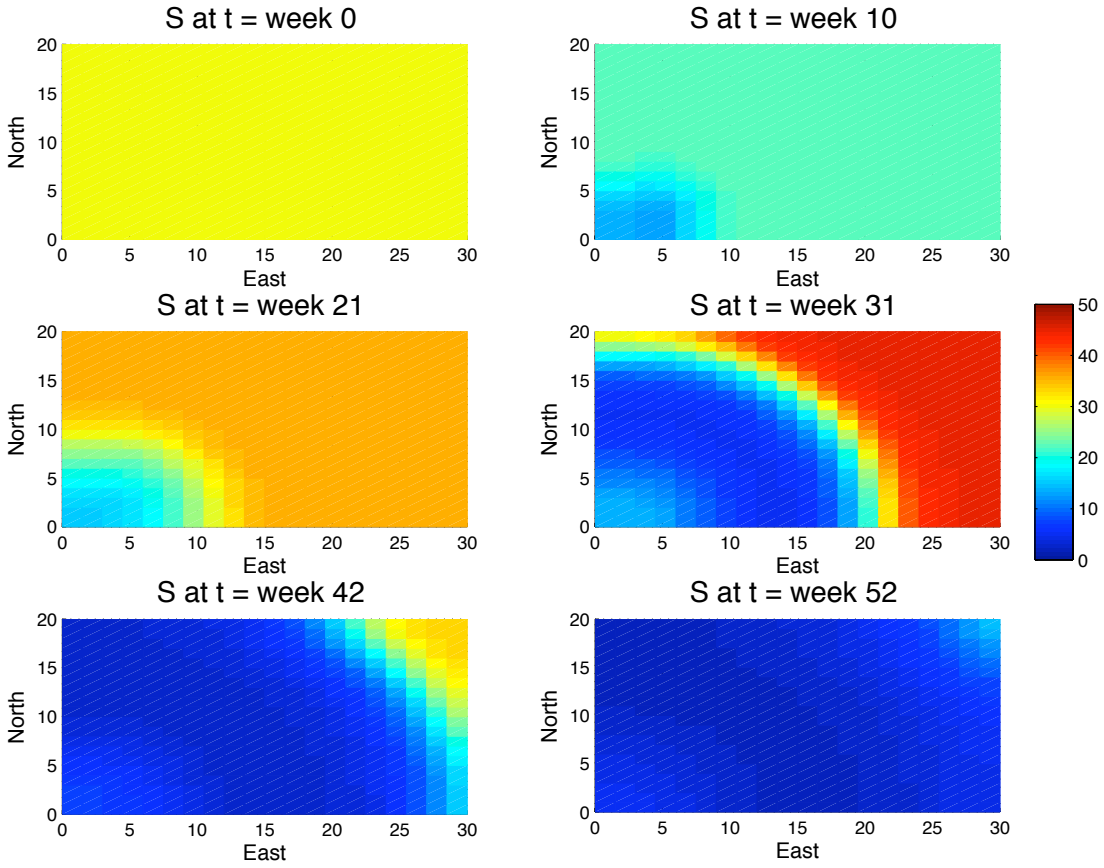


Figure C.2: Spatial Homogeneity: Susceptible population over 52 weeks in absence of vaccination. Infection is initiated in southwest corner at $t = 0$ weeks.

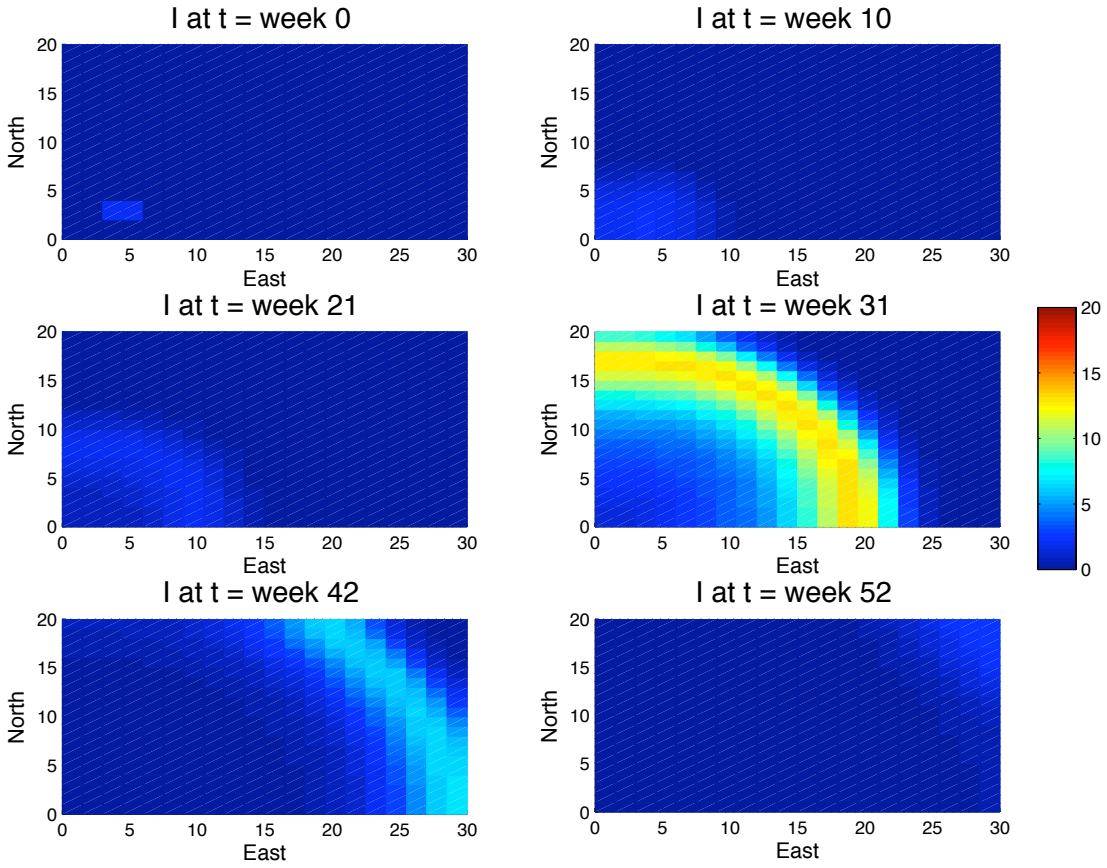


Figure C.3: Spatial Homogeneity: Infected population over 52 weeks in absence of vaccination. Infection is initiated in southwest corner at $t = 0$ weeks.

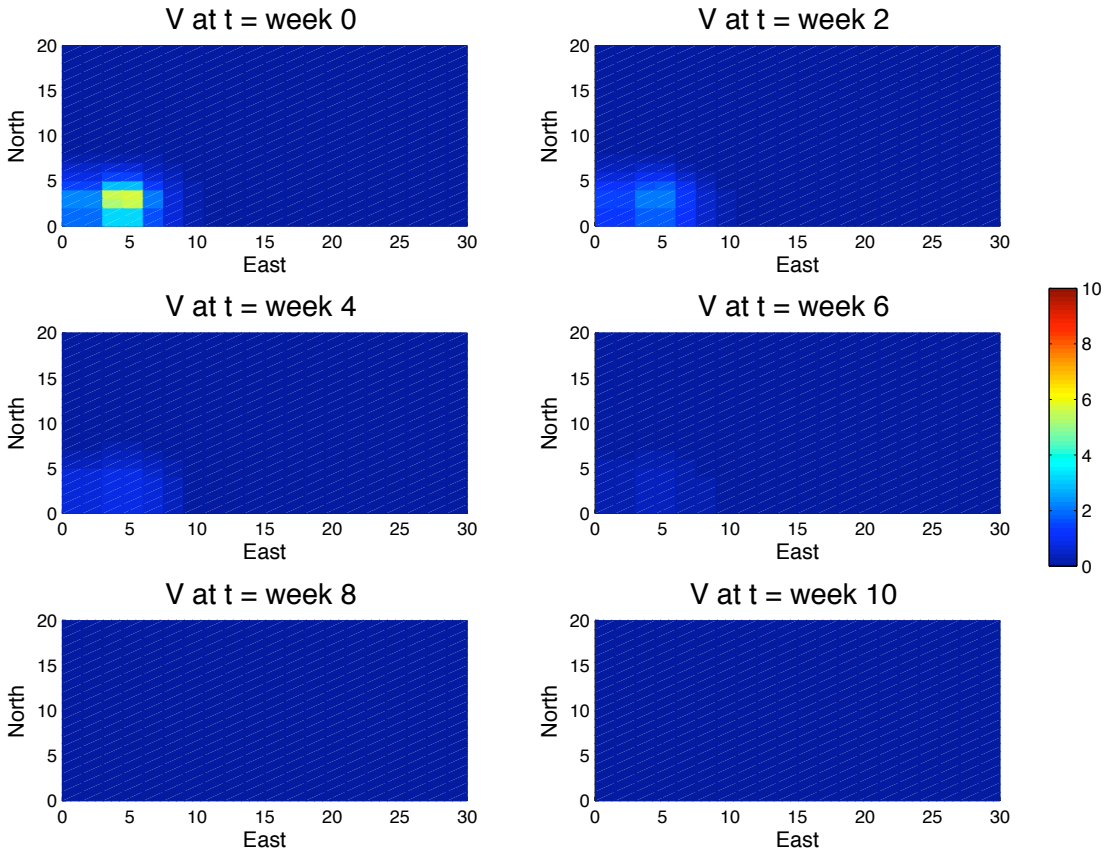


Figure C.4: Spatial Homogeneity: Optimal 10-week vaccination schedule starting at $t = 0$ weeks. Vaccine uptake parameter is assumed to be $a = 0.02$.

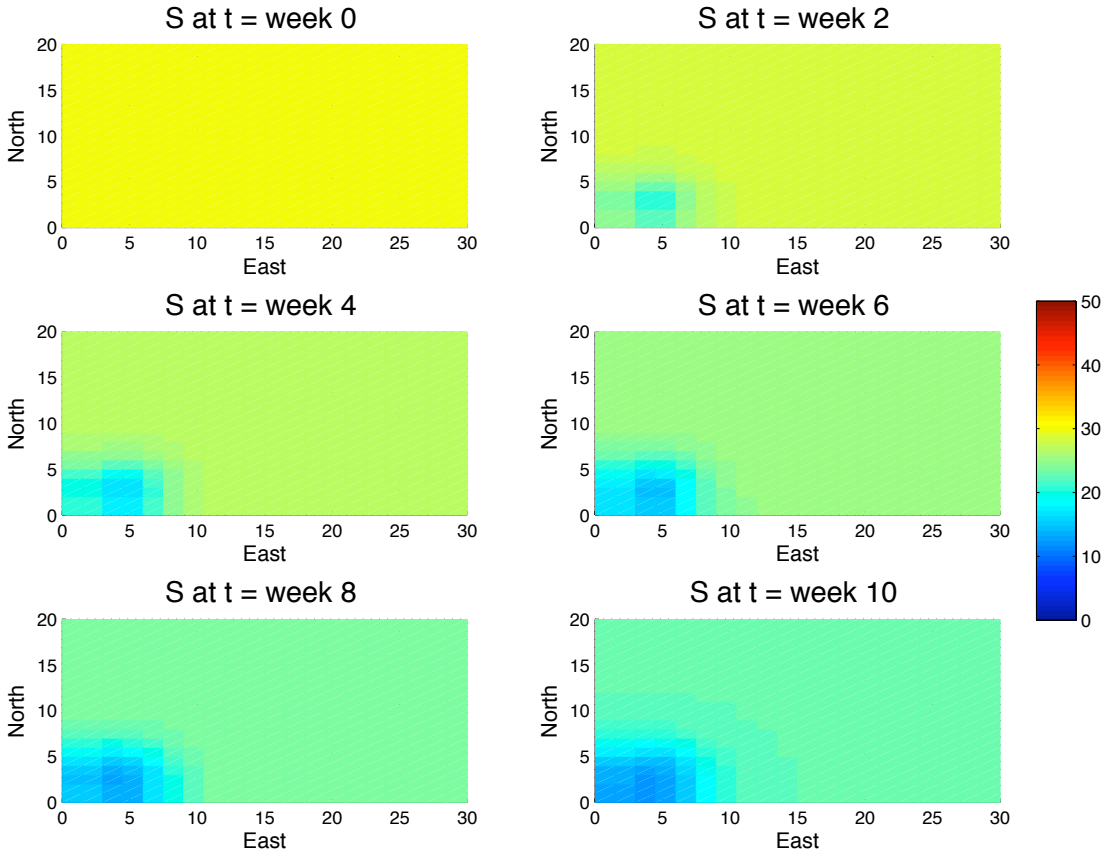


Figure C.5: Spatial Homogeneity: Susceptible population during optimal 10-week vaccination schedule starting at $t = 0$ week. Vaccine uptake parameter is assumed to be $a = 0.02$.

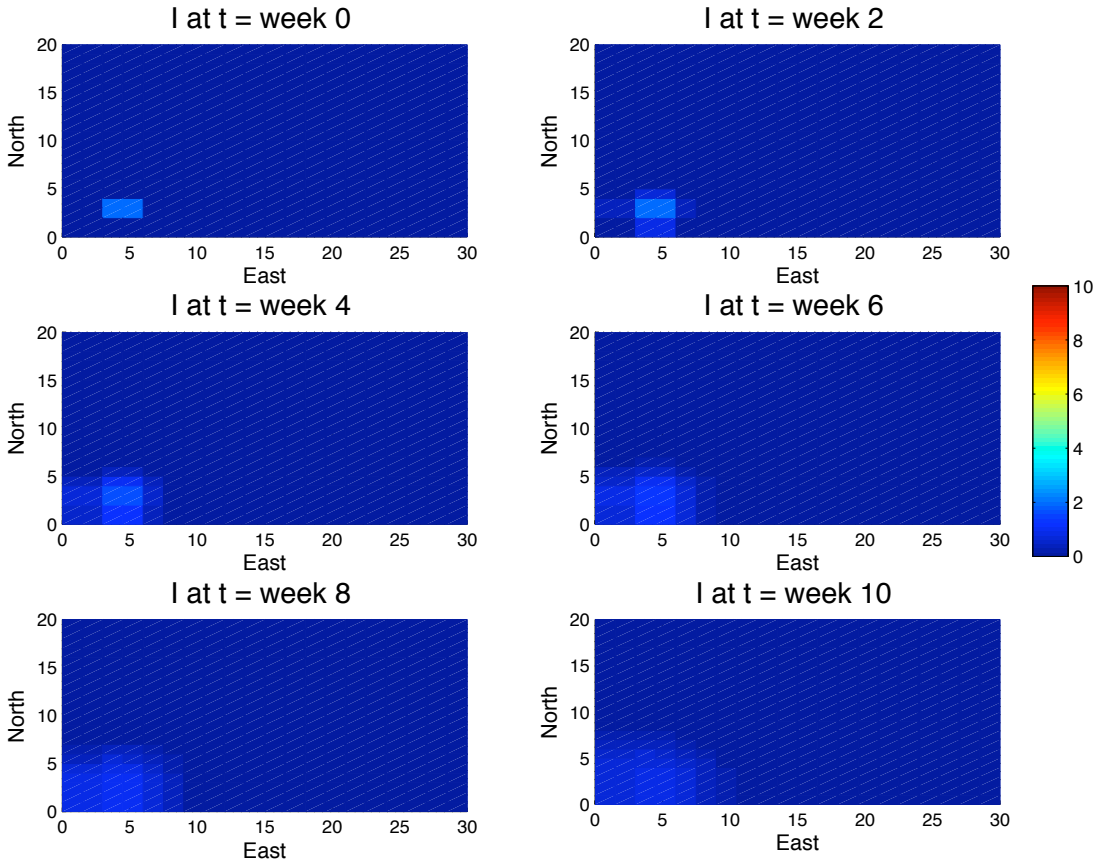


Figure C.6: Spatial Homogeneity: Infected population during optimal 10-week vaccination schedule starting at $t = 0$ week. Vaccine uptake parameter is assumed to be $a = 0.02$.

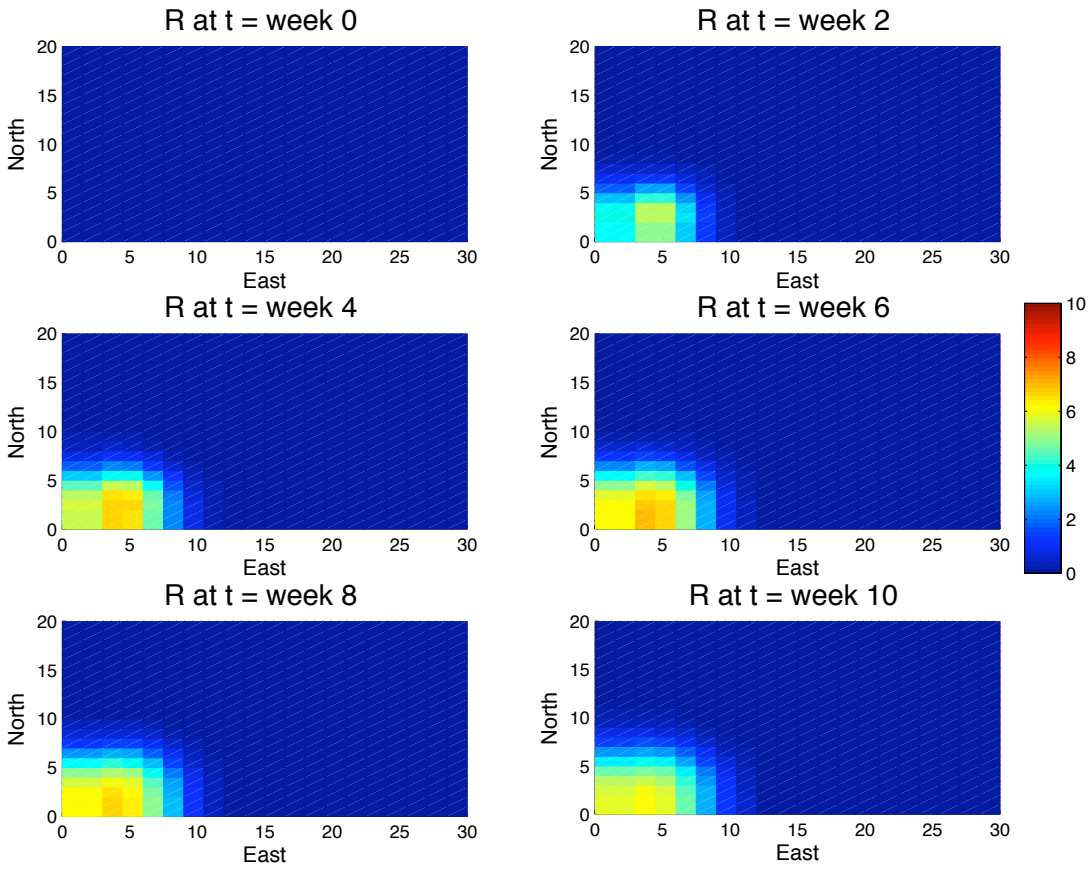


Figure C.7: Spatial Homogeneity: Immune population during optimal 10-week vaccination schedule starting at $t = 0$ week. Vaccine uptake parameter is assumed to be $a = 0.02$.

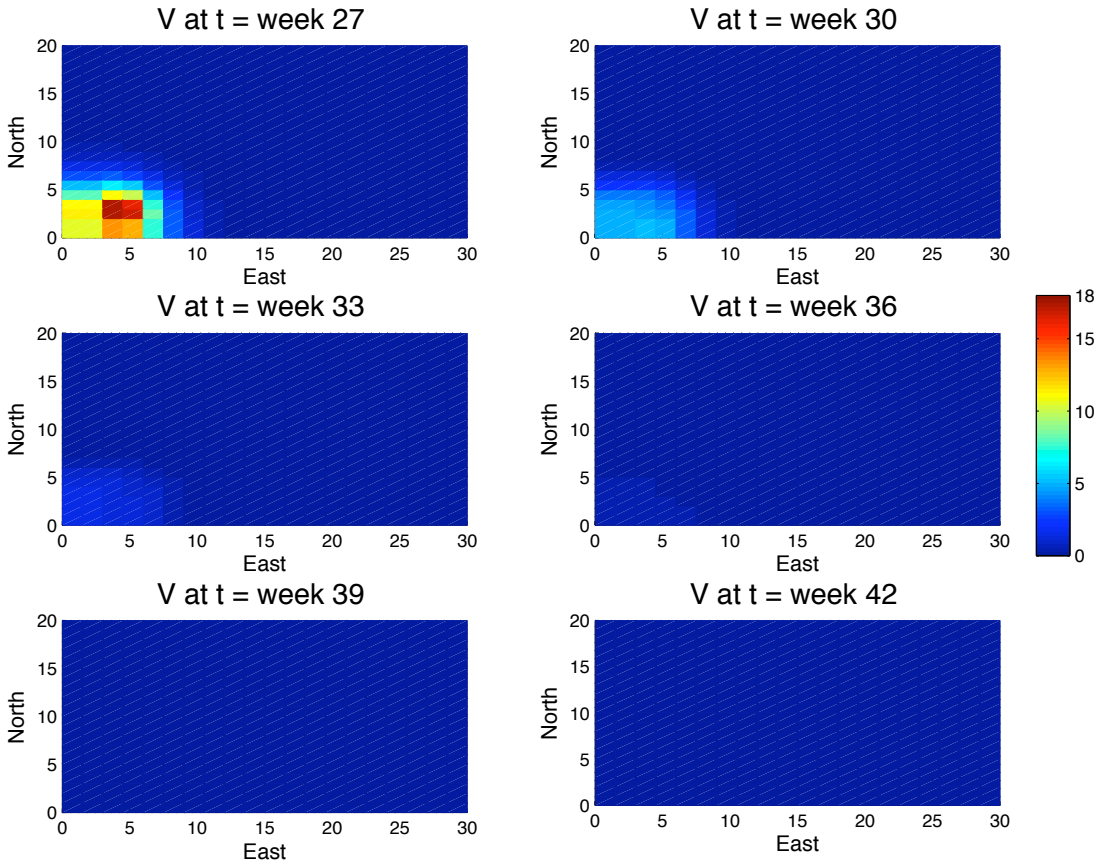


Figure C.8: Spatial Homogeneity: Optimal 15-week vaccination schedule starting at the end of the birth pulse, $t = 27$ week. Infection initiated at $t = 27$ week in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

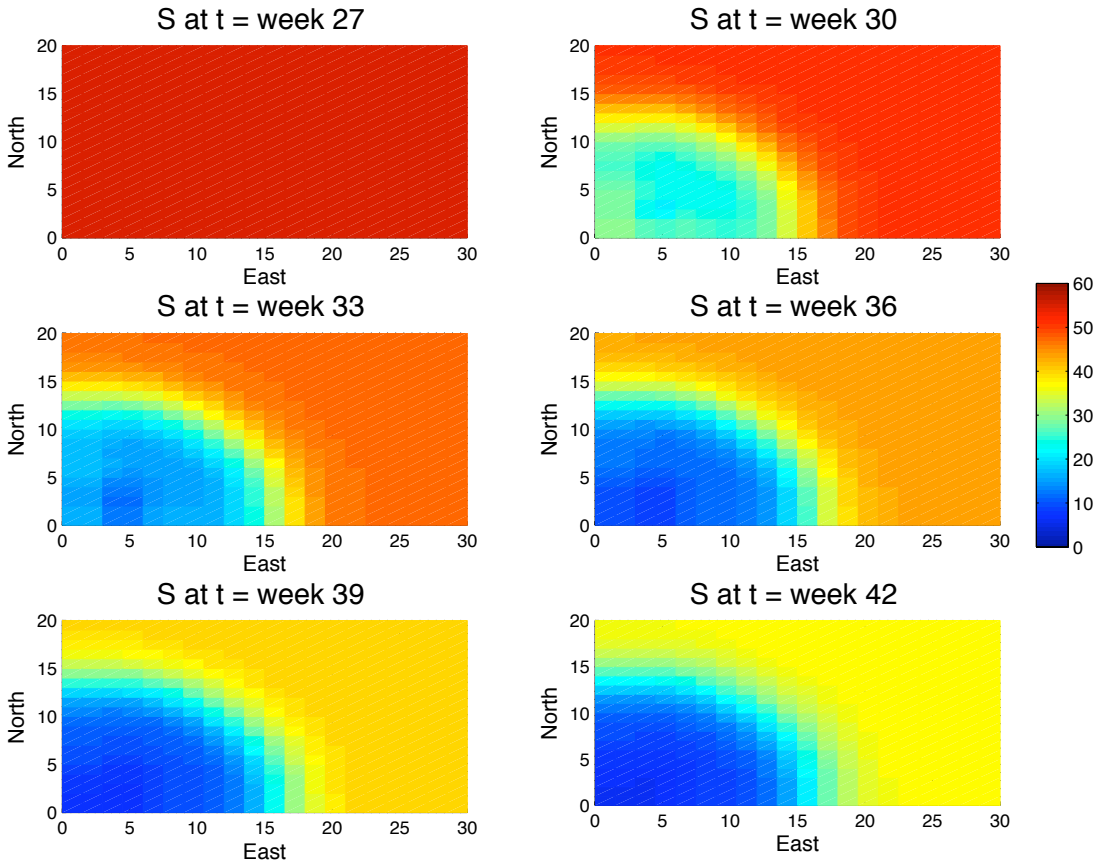


Figure C.9: Spatial Homogeneity: Susceptible population during optimal 15-week vaccination schedule starting at $t = 27$ week. Infection initiated at $t = 27$ week in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

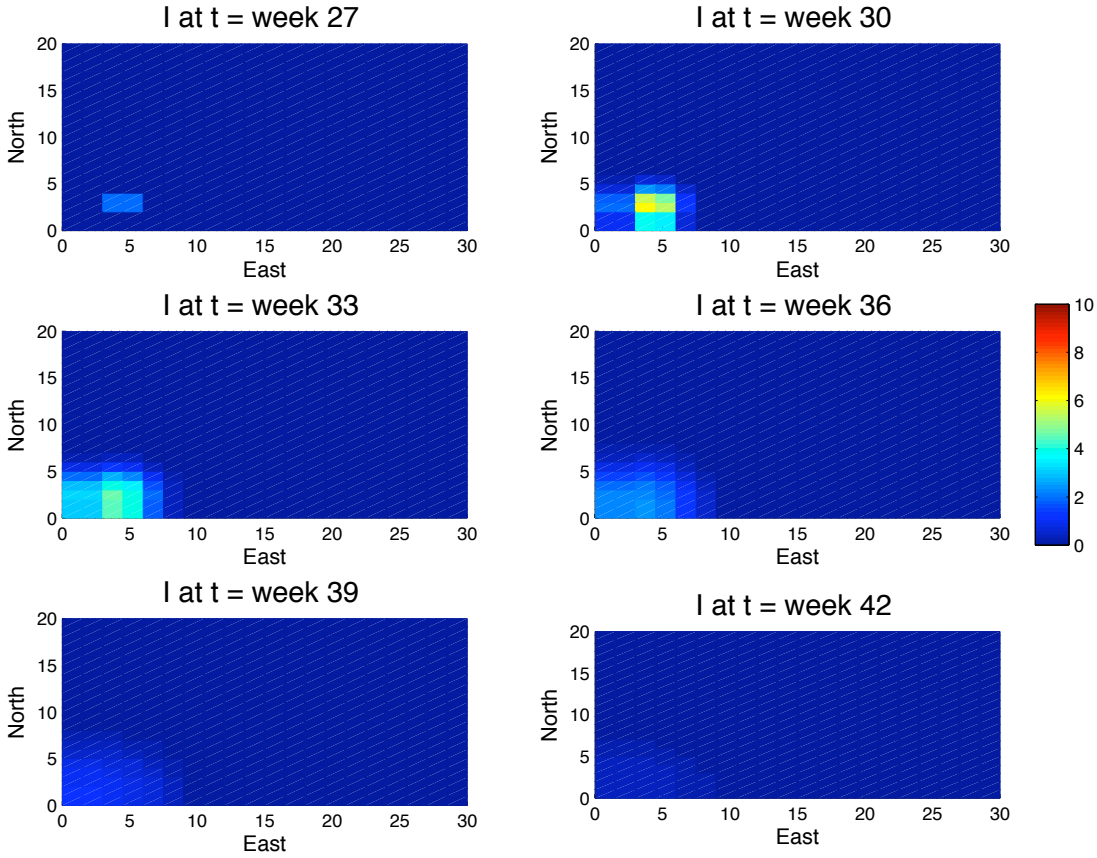


Figure C.10: Spatial Homogeneity: Infected population during optimal 15-week vaccination schedule starting at $t = 27$ week. Infection initiated at $t = 27$ week in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

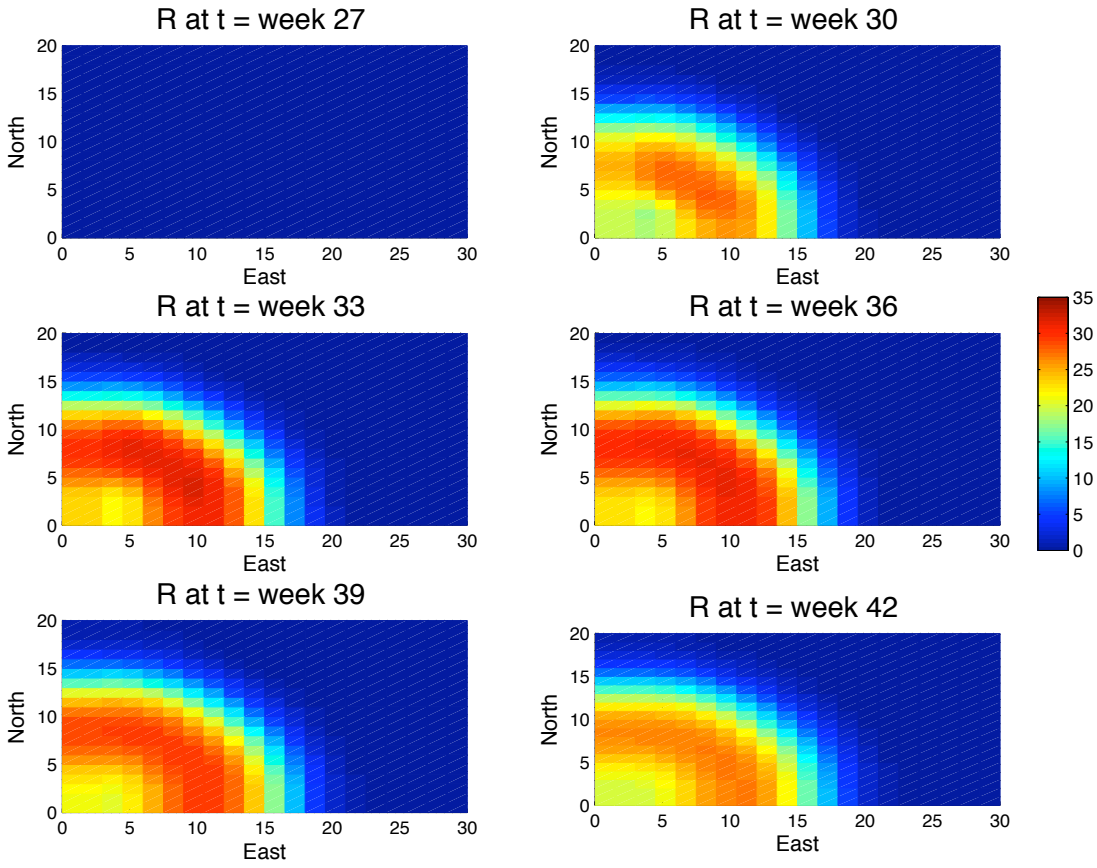


Figure C.11: Spatial Homogeneity: Immune population during optimal 15-week vaccination schedule starting at $t = 27$ week. Infection initiated at $t = 27$ week in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

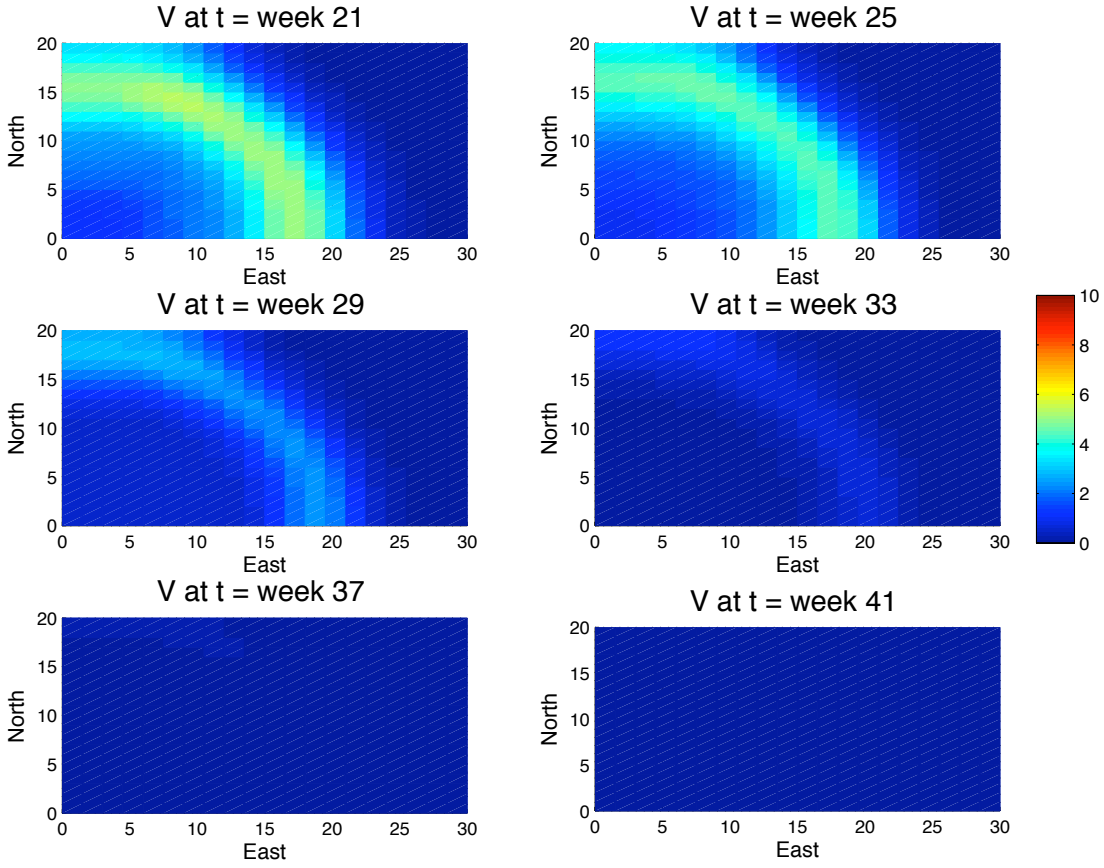


Figure C.12: Spatial Homogeneity: Optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

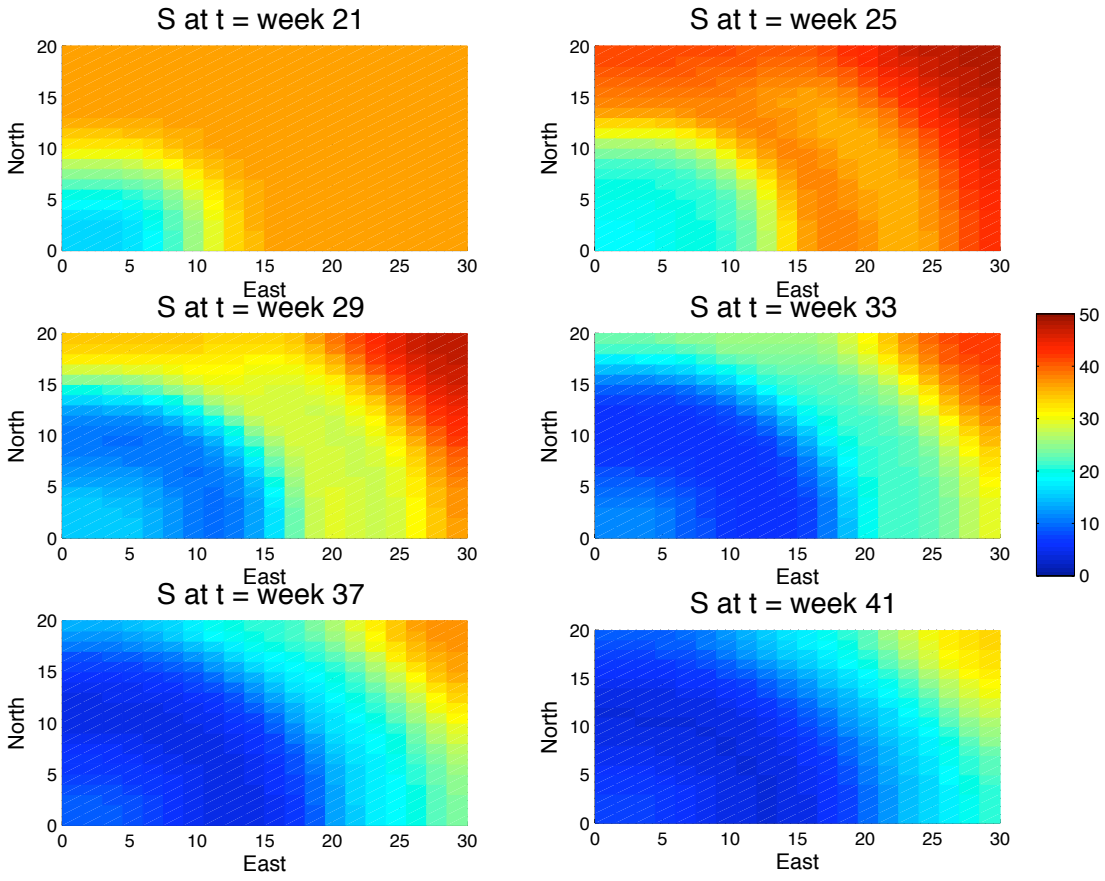


Figure C.13: Spatial Homogeneity: Susceptible population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

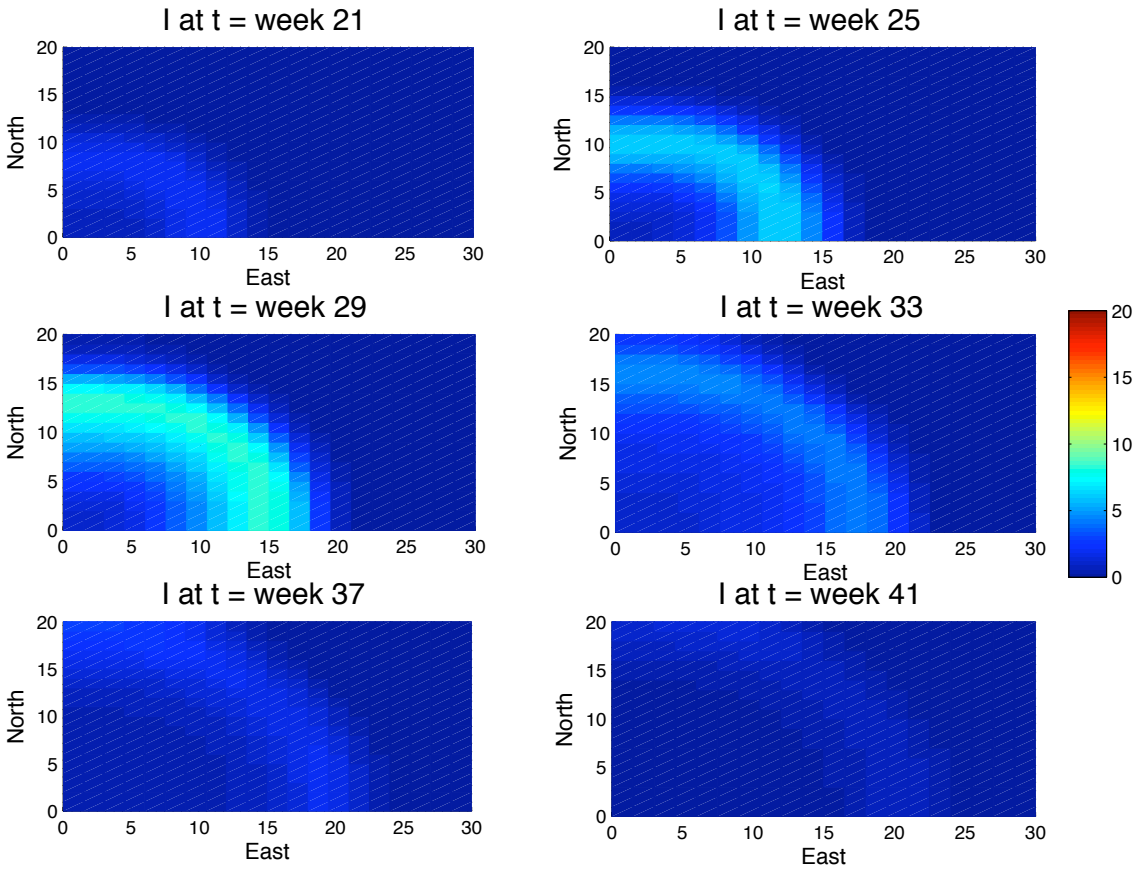


Figure C.14: Spatial Homogeneity: Infected population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

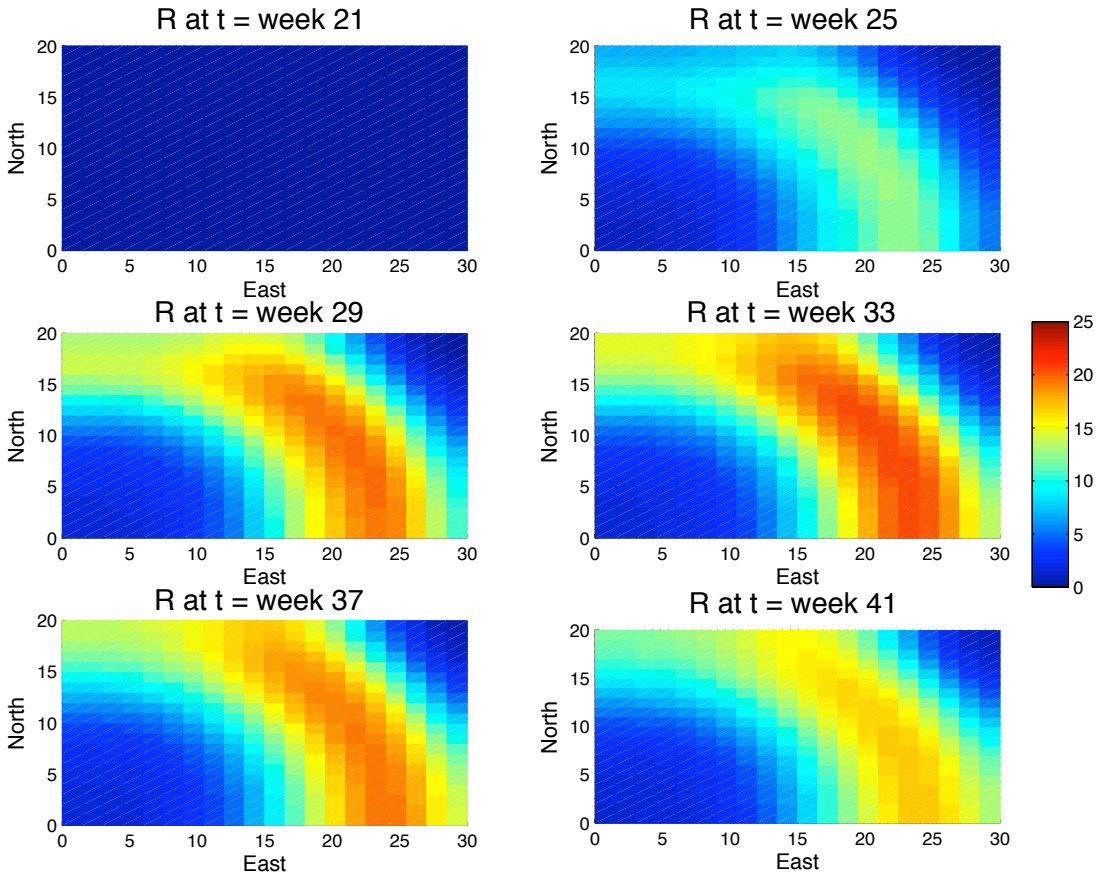


Figure C.15: Spatial Homogeneity: Immune population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection initiated at $t = 0$ week in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

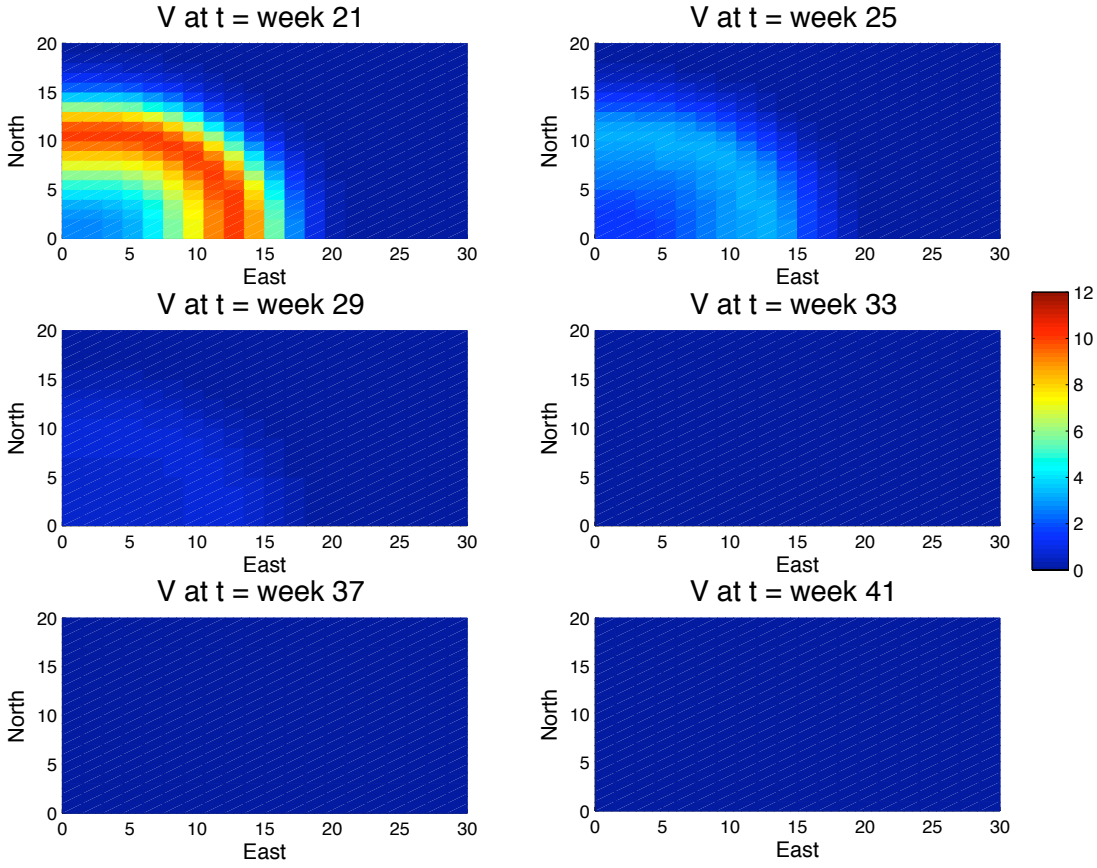


Figure C.16: Spatial Homogeneity: Optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.03$.

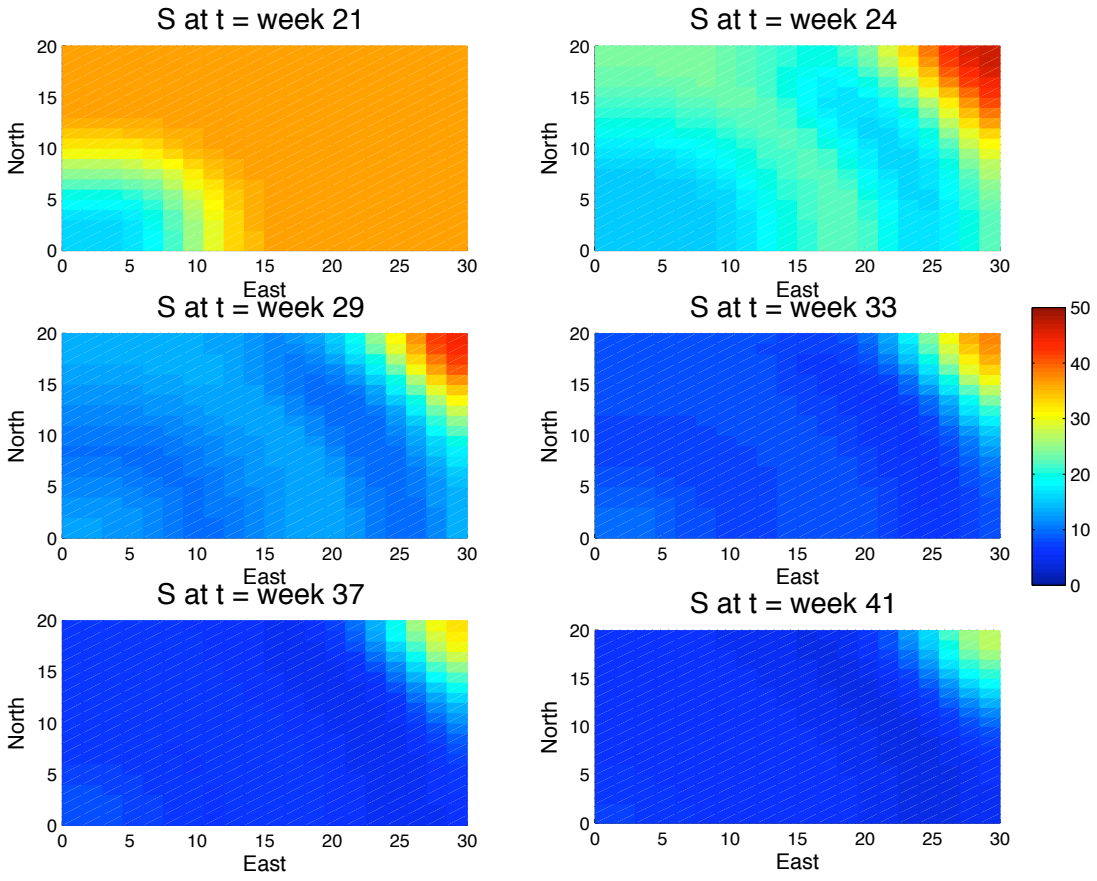


Figure C.17: Spatial Homogeneity: Susceptible population during optimal 20-week vaccination schedule starting at $t = 21$ week. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.03$.

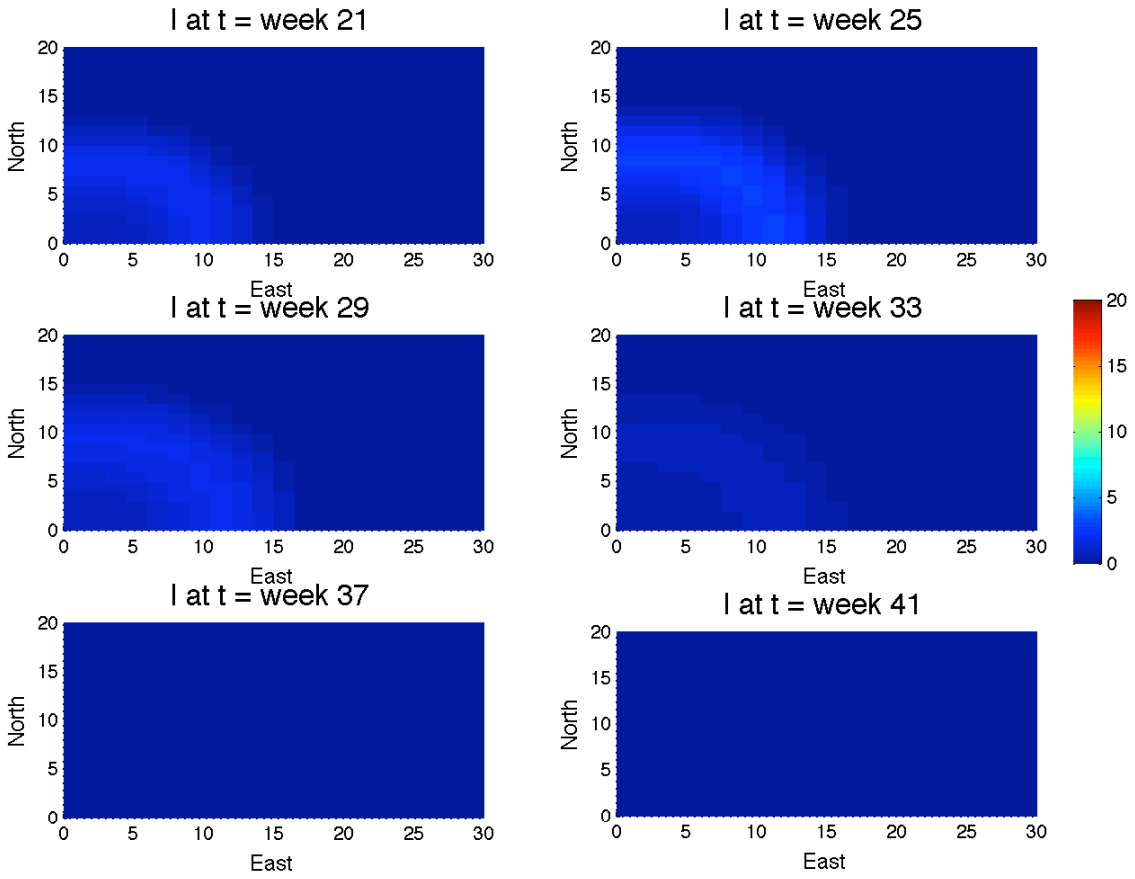


Figure C.18: Spatial Homogeneity: Infected population during optimal 20-week vaccination schedule starting at $t = 21$ week. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.03$.

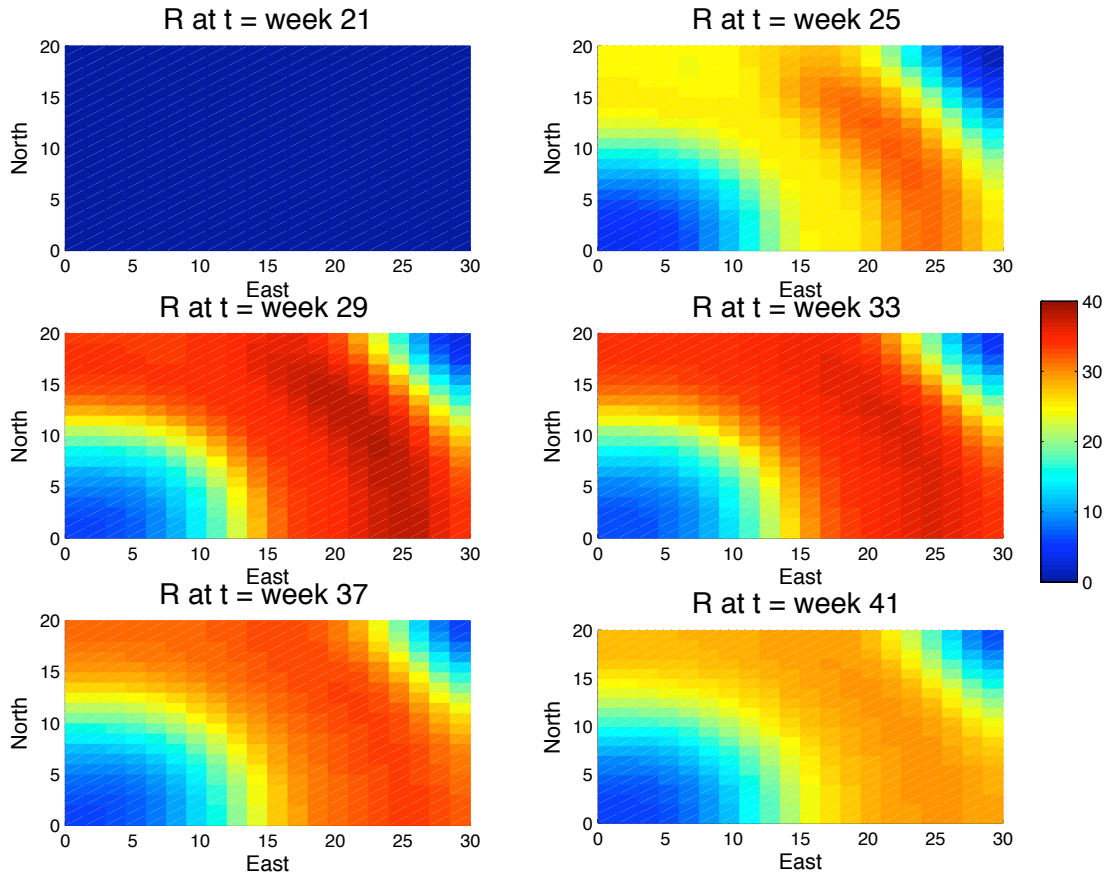


Figure C.19: Spatial Homogeneity: Immune population during optimal 20-week vaccination schedule starting at $t = 21$ week. Infection initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.03$.

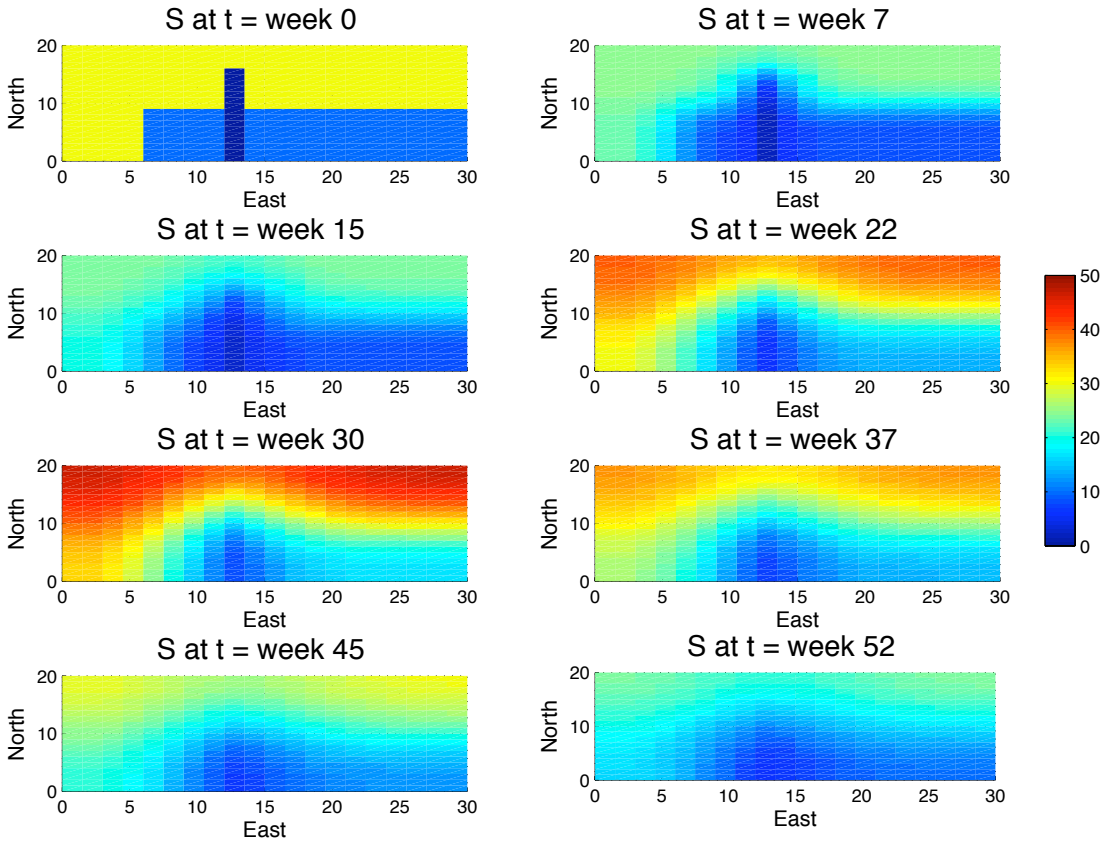


Figure C.20: Spatial Heterogeneity: Susceptible population over 52 weeks in the absence of any infection. Growth and decline of population indicate natural birth and death of susceptible raccoons.

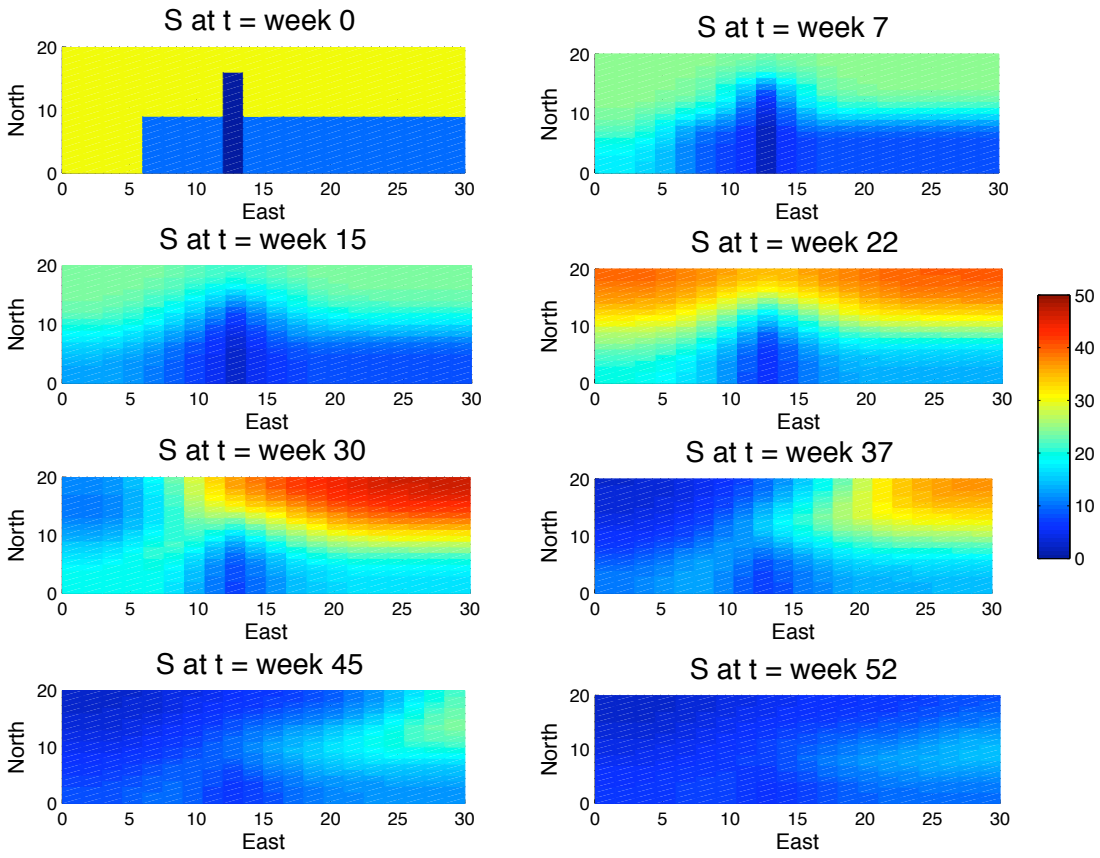


Figure C.21: Spatial Heterogeneity: Susceptible population over 52 weeks, in absence of vaccination. Spatial heterogeneity includes decreased movement and decreased initial density due to river and forest cover. Infection is initiated in southwest corner at $t = 0$ weeks.

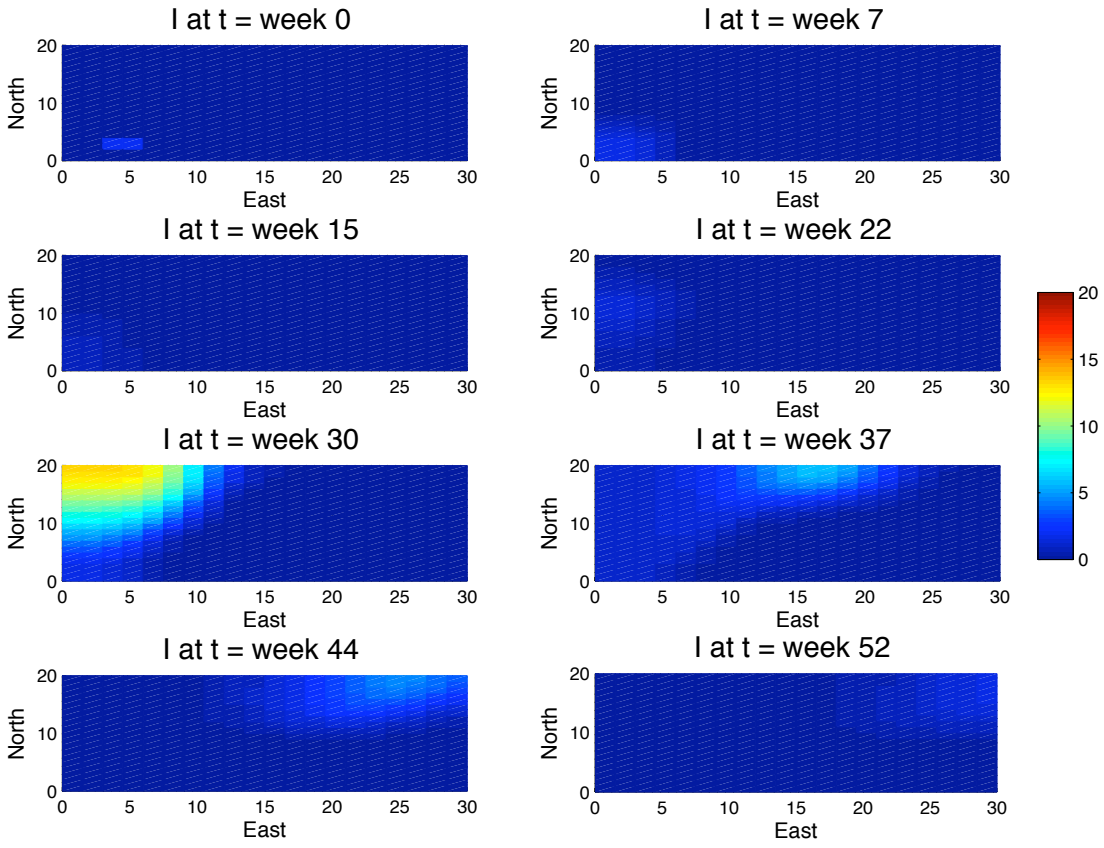


Figure C.22: Spatial Heterogeneity: Infected population over 52 weeks, in absence of vaccination. Spatial heterogeneity includes decreased movement and decreased initial density of susceptible raccoons due to river and forest cover. Infection is initiated in southwest corner at $t = 0$ weeks.

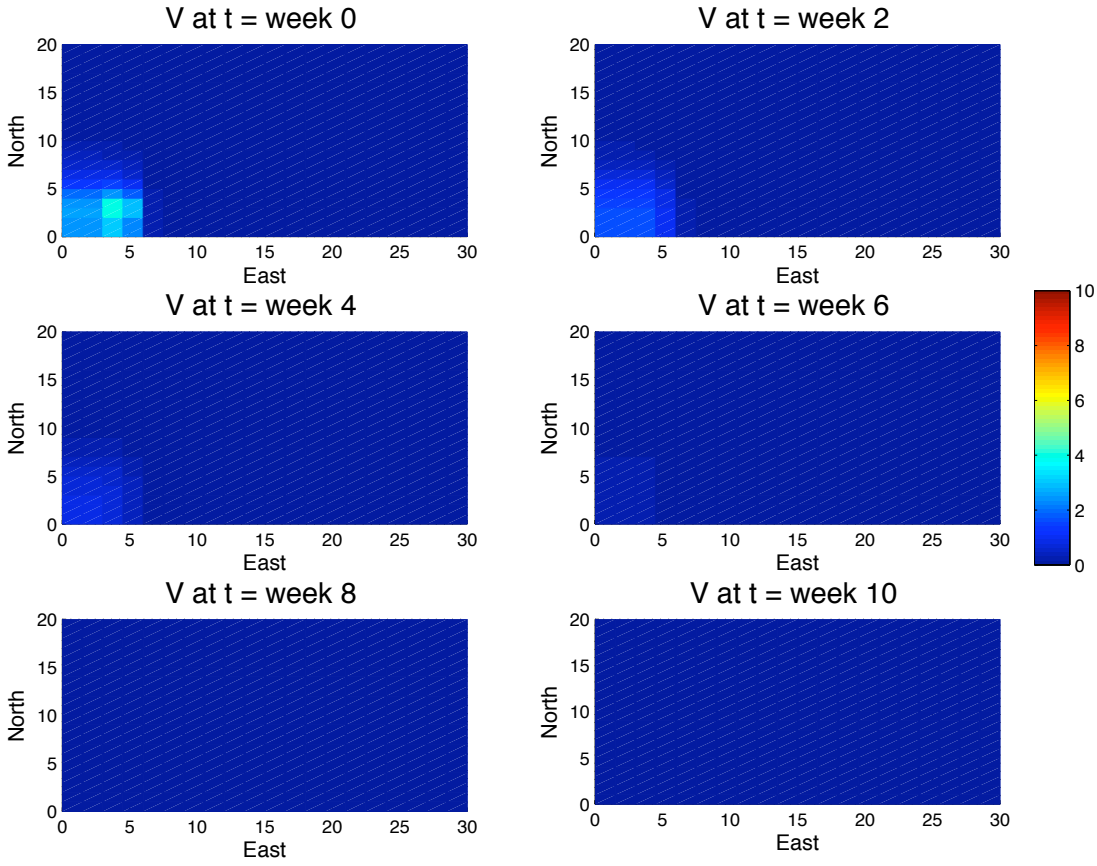


Figure C.23: Spatial Heterogeneity: Optimal 10-week vaccination schedule starting at $t = 0$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake is parameter assumed to be $a = 0.02$.

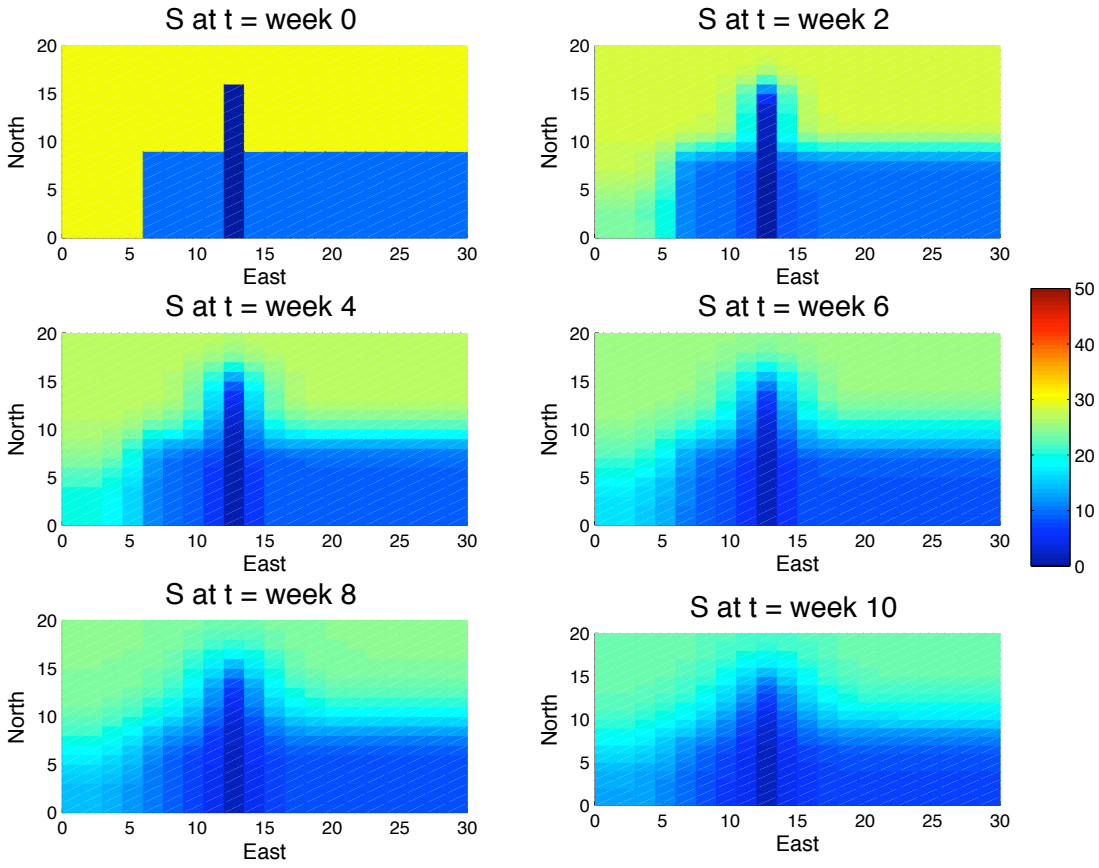


Figure C.24: Spatial Heterogeneity: Susceptible population during optimal 10-week vaccination schedule starting at $t = 0$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

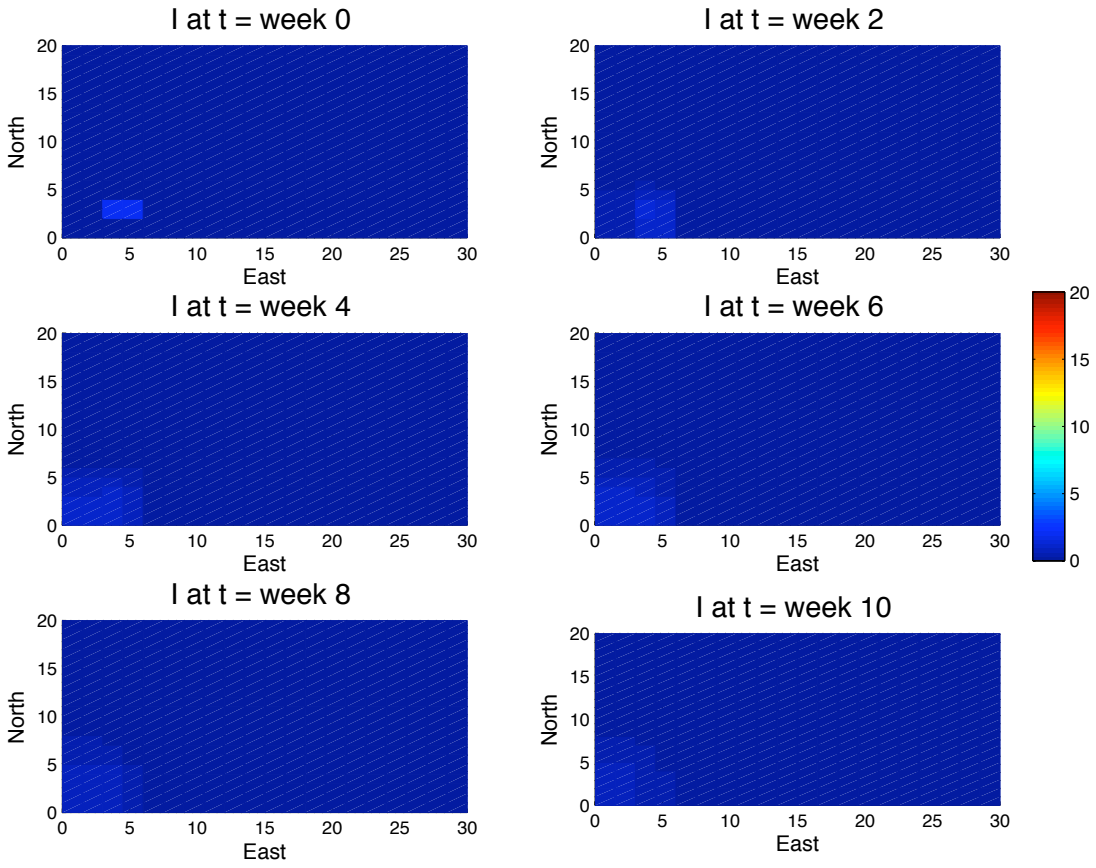


Figure C.25: Spatial Heterogeneity: Infected population during optimal 10-week vaccination schedule starting at $t = 0$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

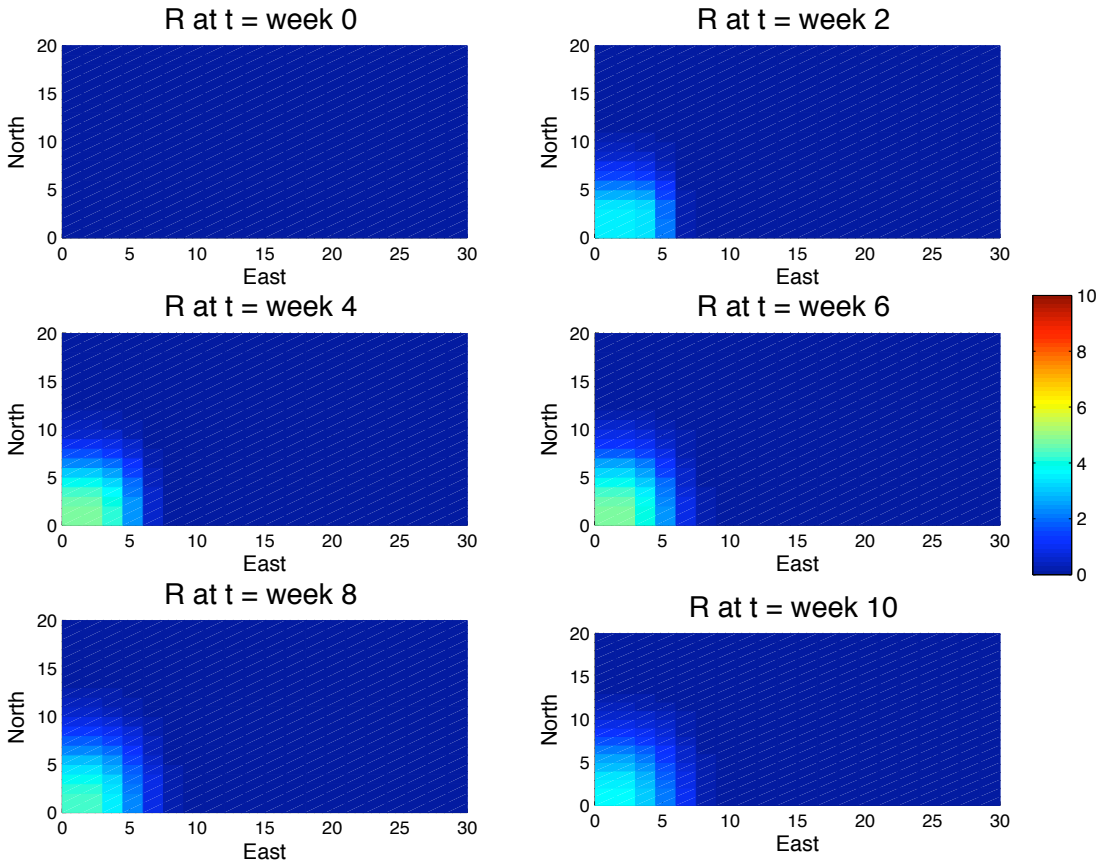


Figure C.26: Spatial Heterogeneity: Immune population during optimal 10-week vaccination schedule starting at $t = 0$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.02$.

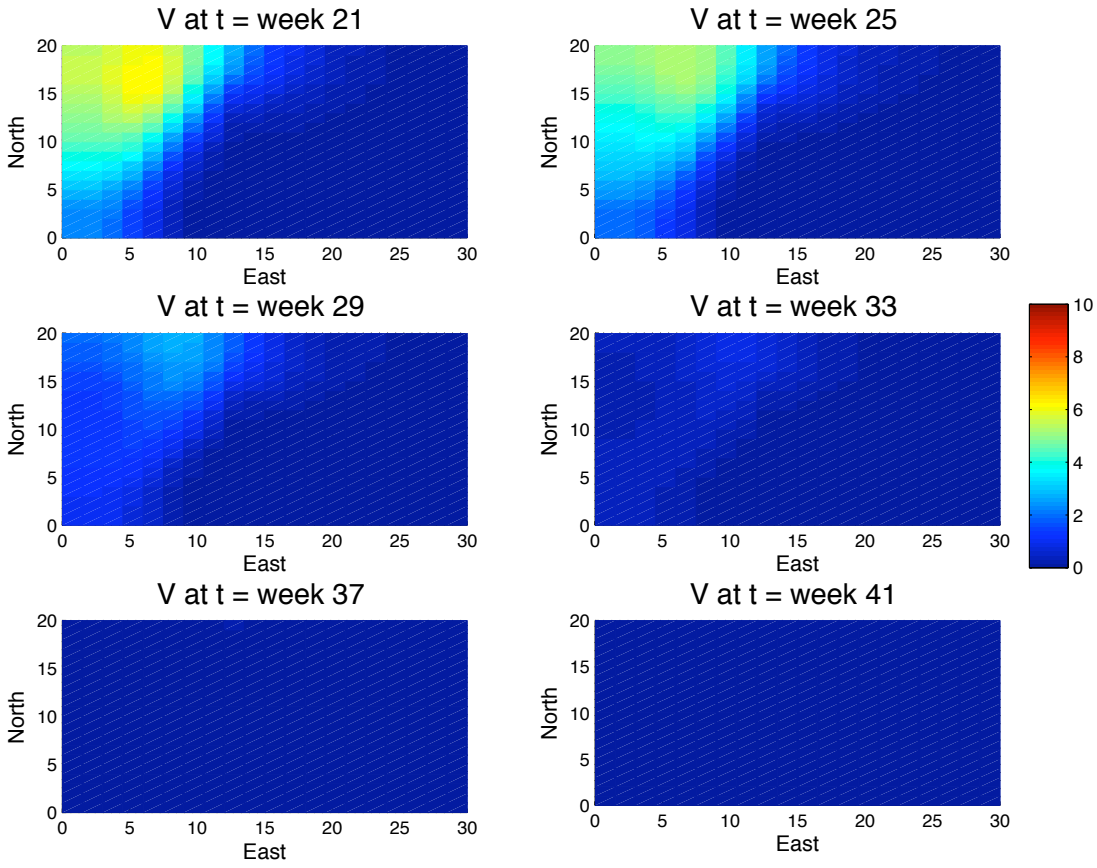


Figure C.27: Spatial Heterogeneity: Optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$

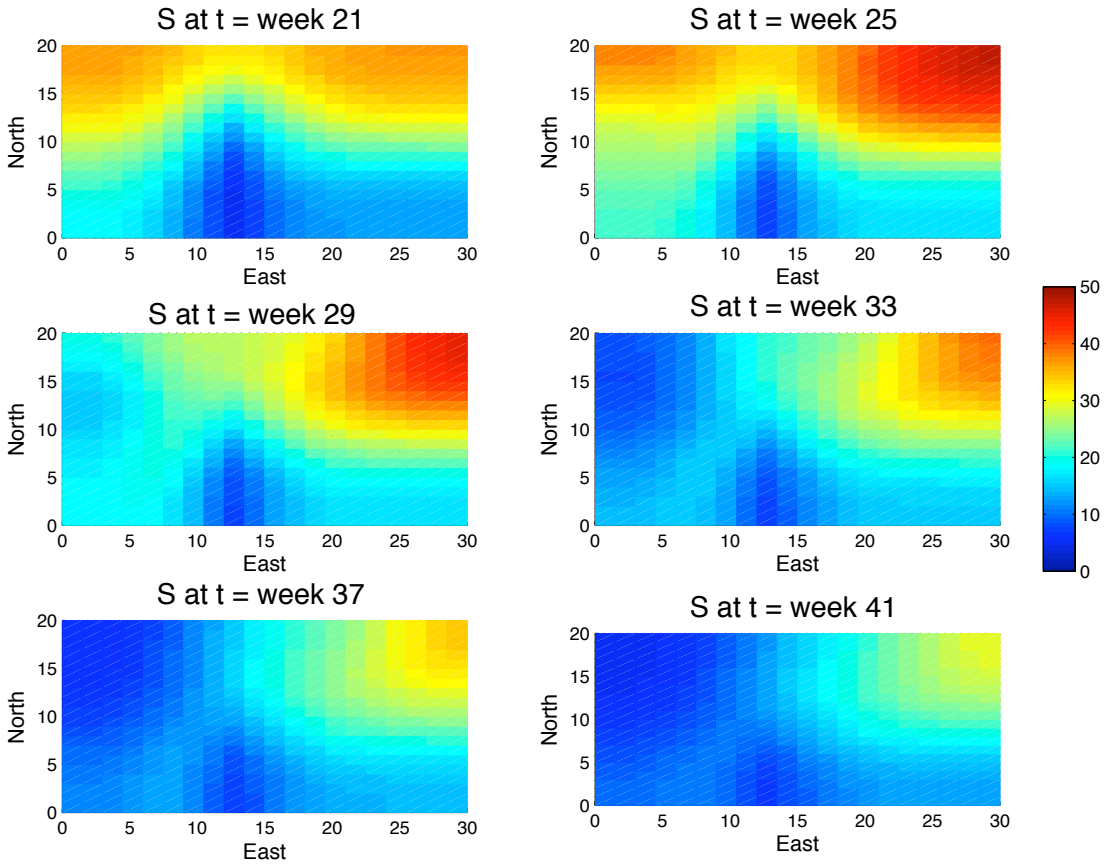


Figure C.28: Spatial Heterogeneity: Susceptible population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

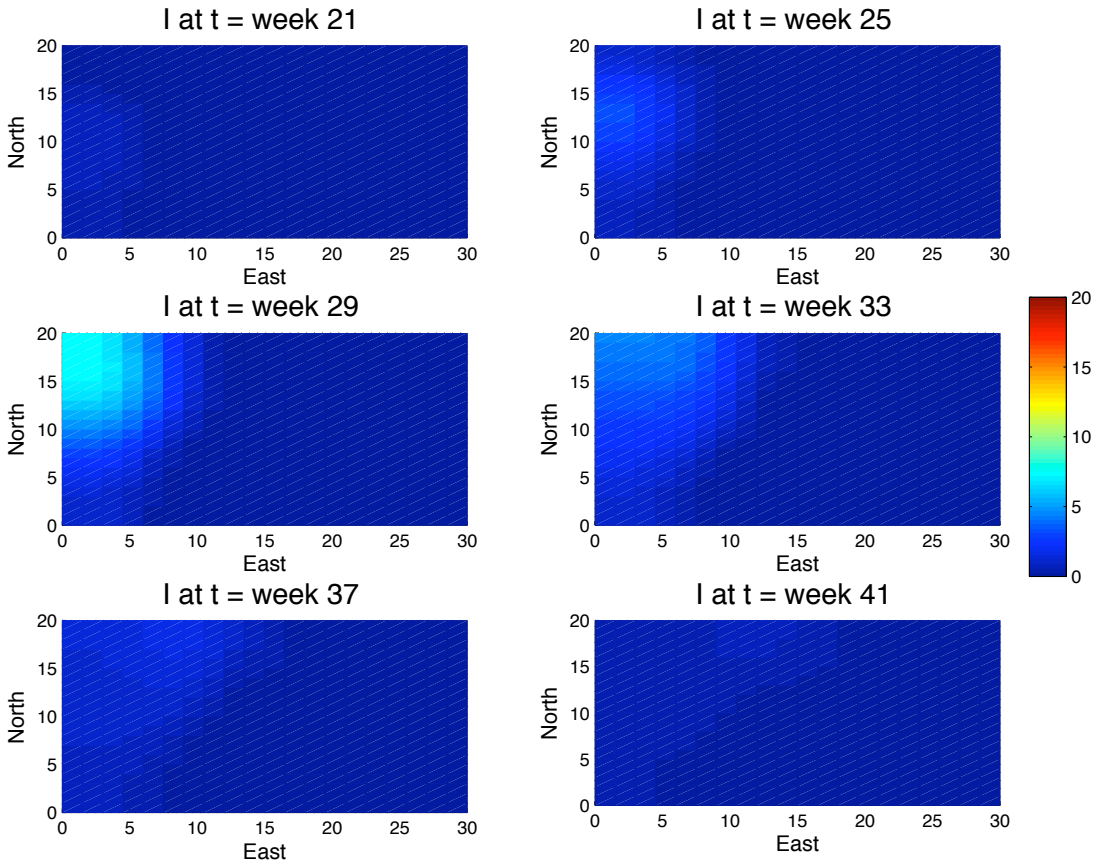


Figure C.29: Spatial Heterogeneity: Infected population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

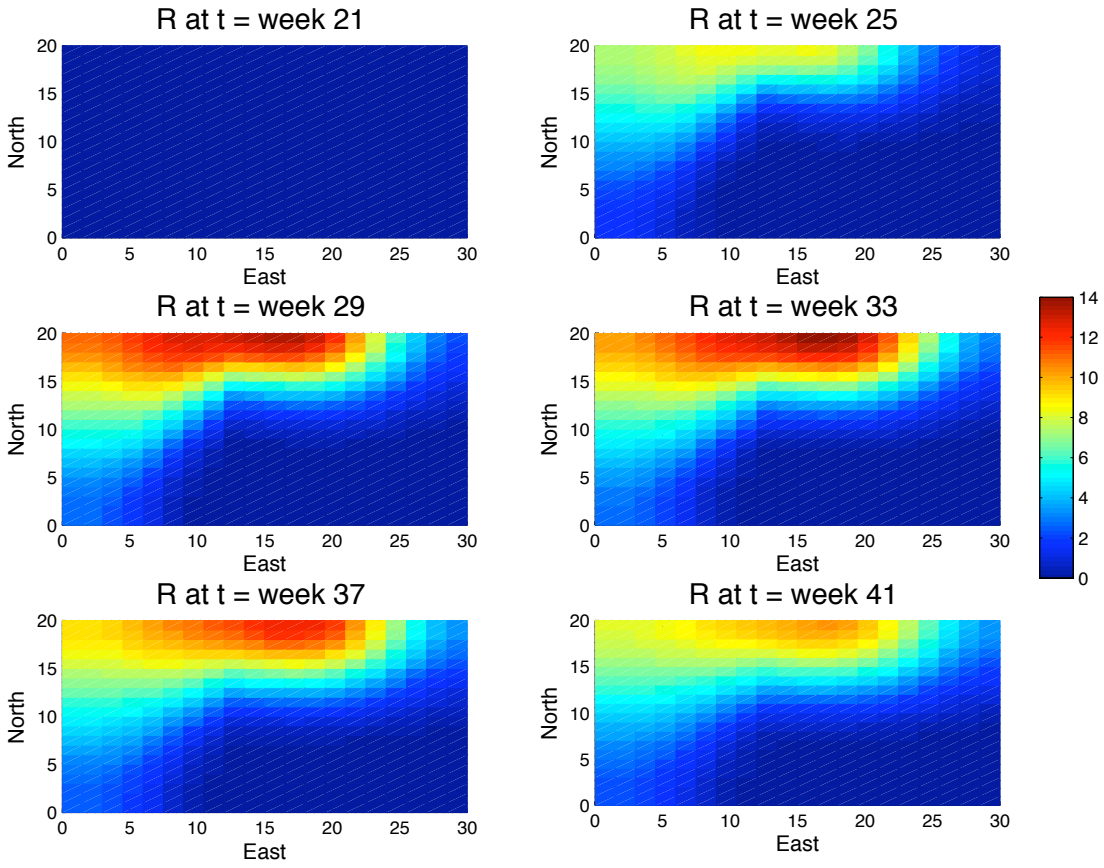


Figure C.30: Spatial Heterogeneity: Immune population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner. Vaccine uptake parameter is assumed to be $a = 0.01$.

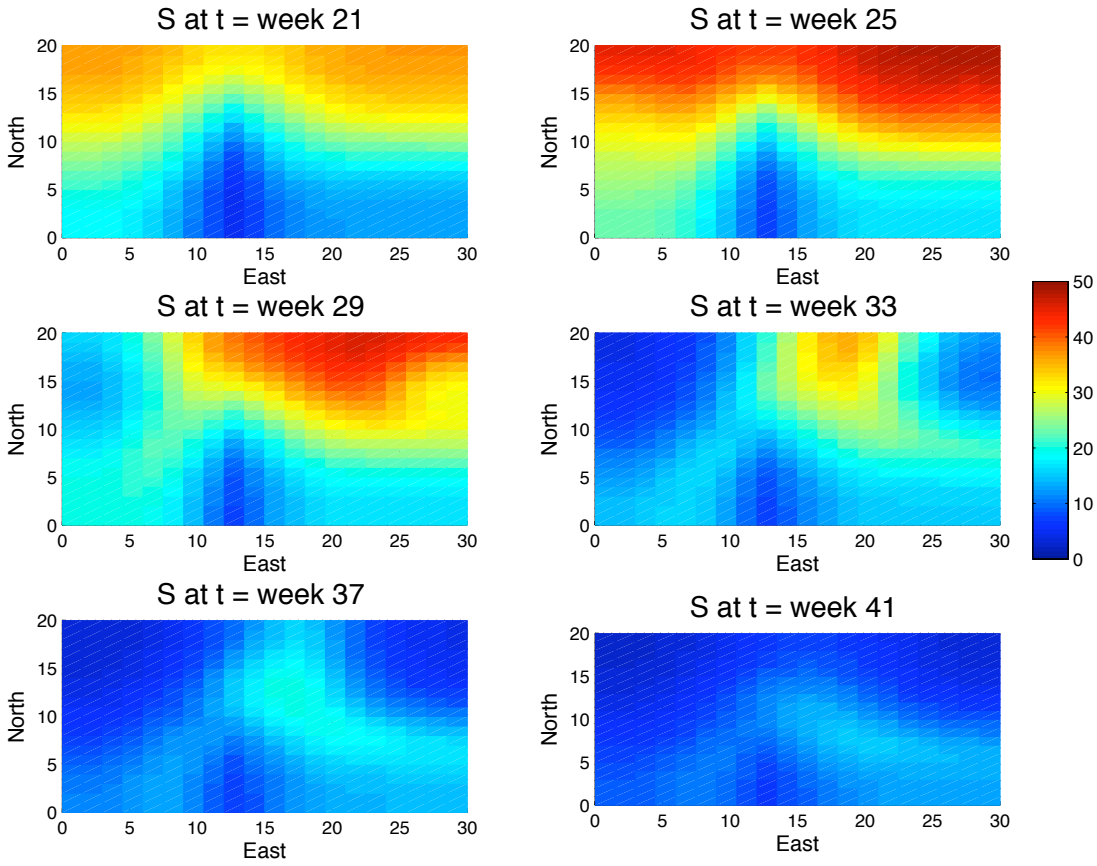


Figure C.31: Spatial Heterogeneity and LDT: Susceptible population starting at $t = 21$ weeks, in absence of vaccination. Infection initiated in bottom right corner at $t = 0$ weeks. Spatial heterogeneity includes decreased movement and decreased initial density due to river and forest cover. Long distance translocation of one infected raccoon appears in upper right corner at $t = 21$ weeks.

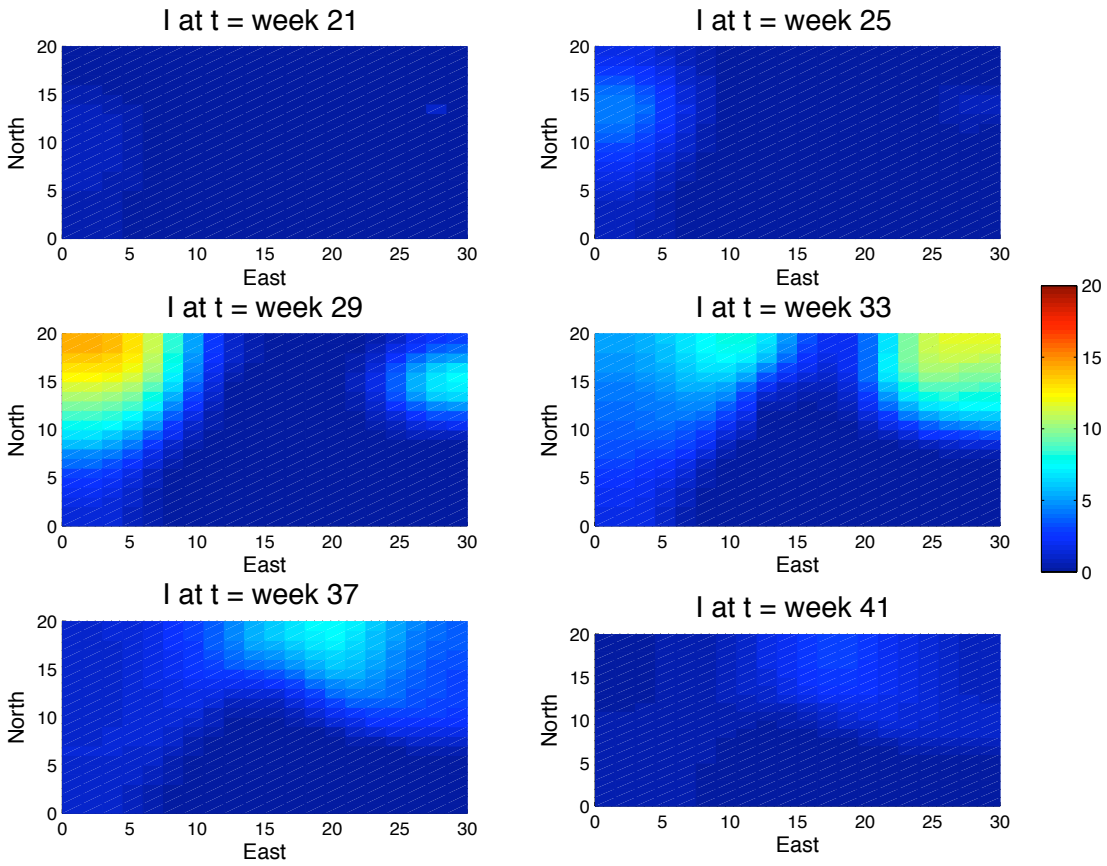


Figure C.32: Spatial Heterogeneity and LDT: Infected population starting at $t = 21$ weeks, in absence of vaccination. Infection initiated in bottom right corner at $t = 0$ weeks. Spatial heterogeneity includes decreased movement and decreased initial density due to river and forest cover. Long distance translocation of one infected raccoon appears in upper right corner at $t = 21$ weeks.

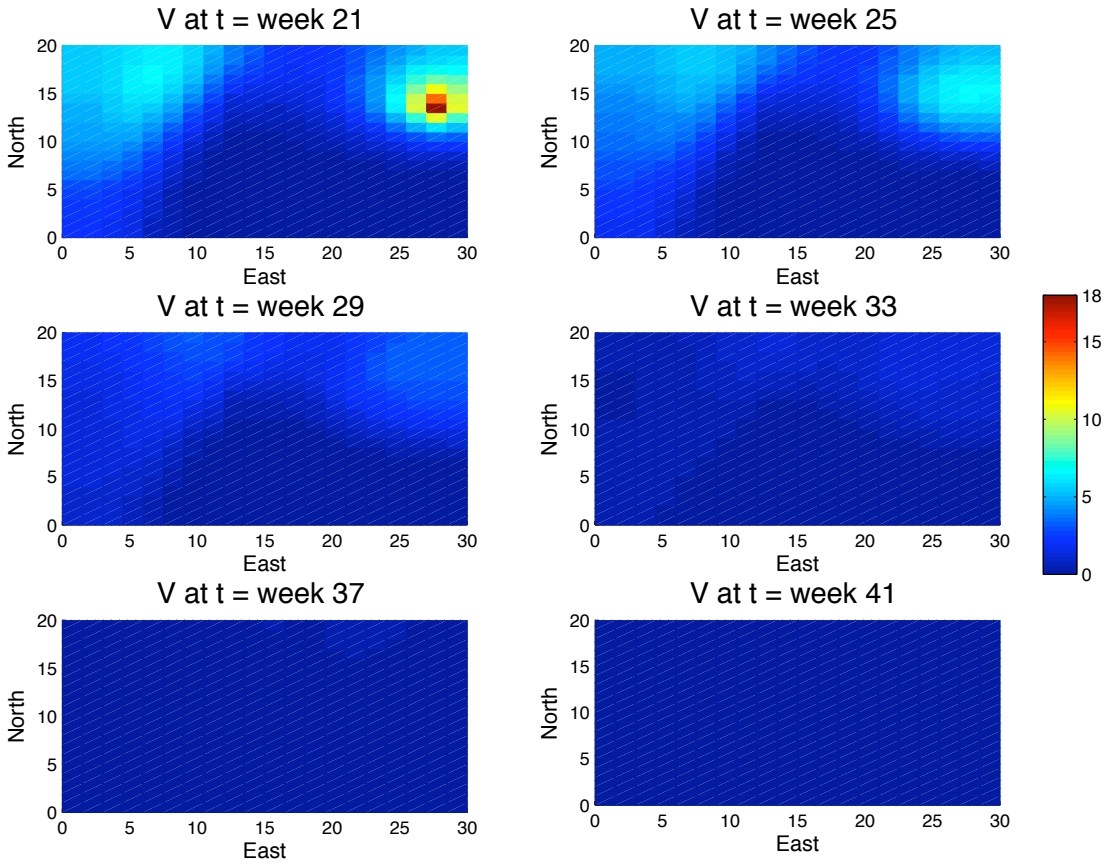


Figure C.33: Spatial Heterogeneity and LDT: Optimal 20-week vaccination schedule beginning on $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.10$.

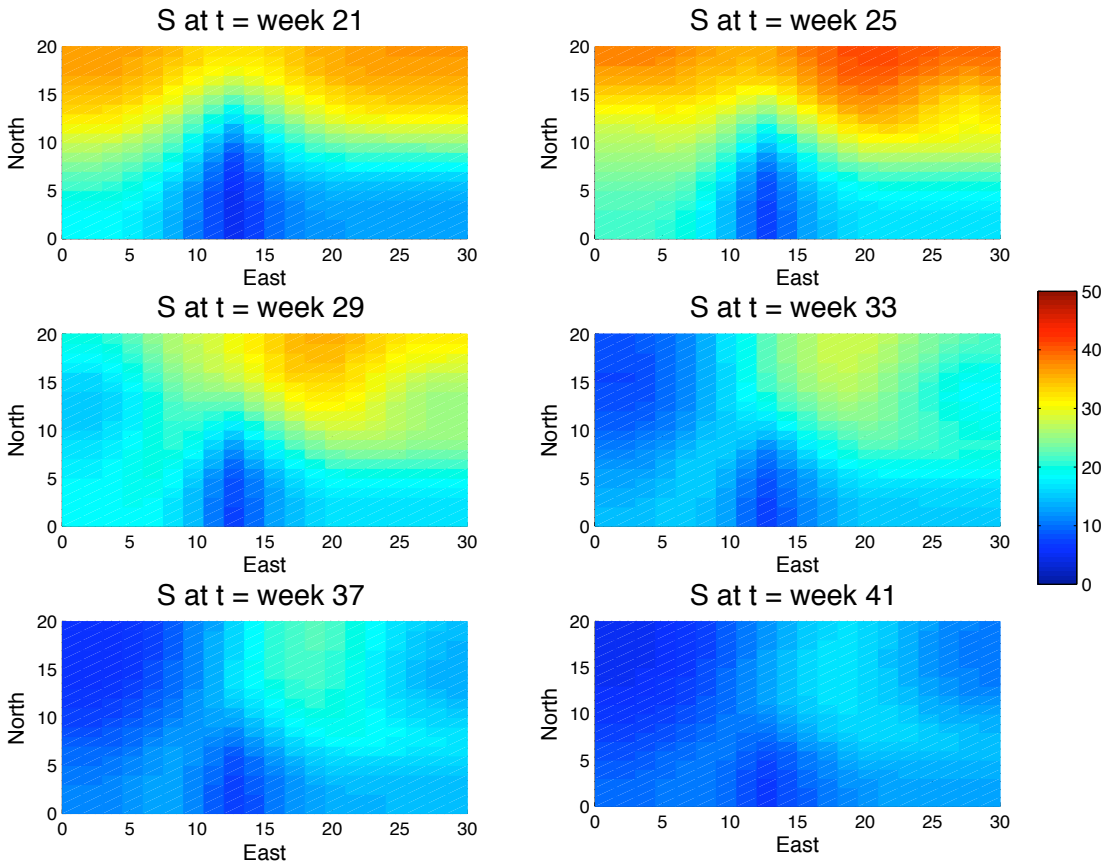


Figure C.34: Spatial Heterogeneity and LDT: Susceptible population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.10$.

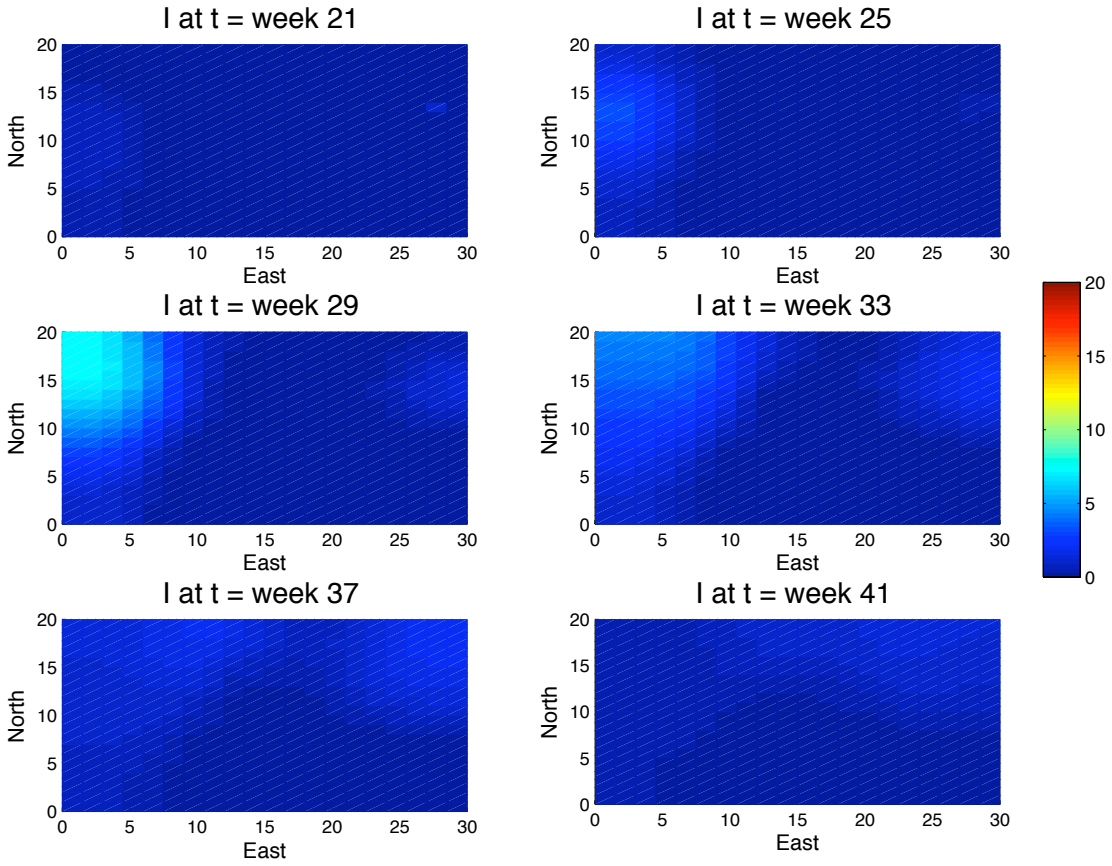


Figure C.35: Spatial Heterogeneity and LDT: Infected population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.10$.

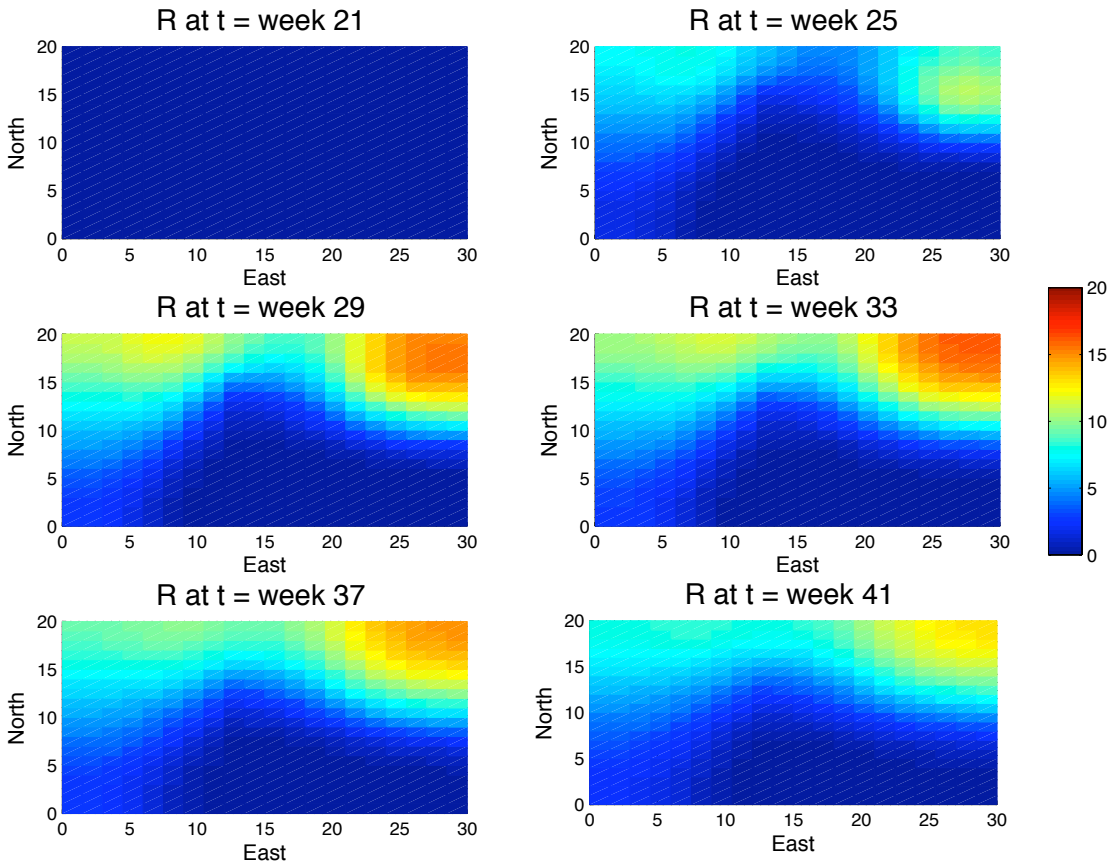


Figure C.36: Spatial Heterogeneity and LDT: Immune population during optimal 20-week vaccination schedule starting at $t = 21$ week. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.10$.

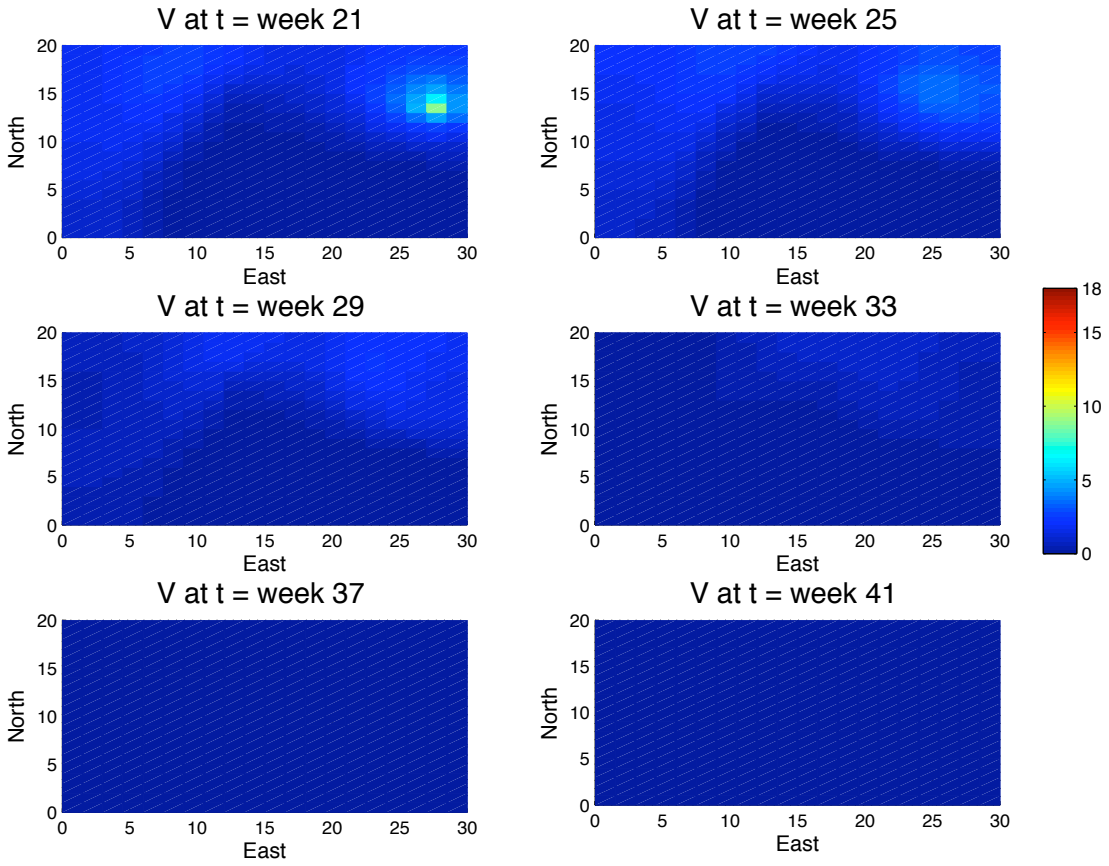


Figure C.37: Spatial Heterogeneity and LDT: Optimal 20-week vaccination schedule beginning on $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.25$.

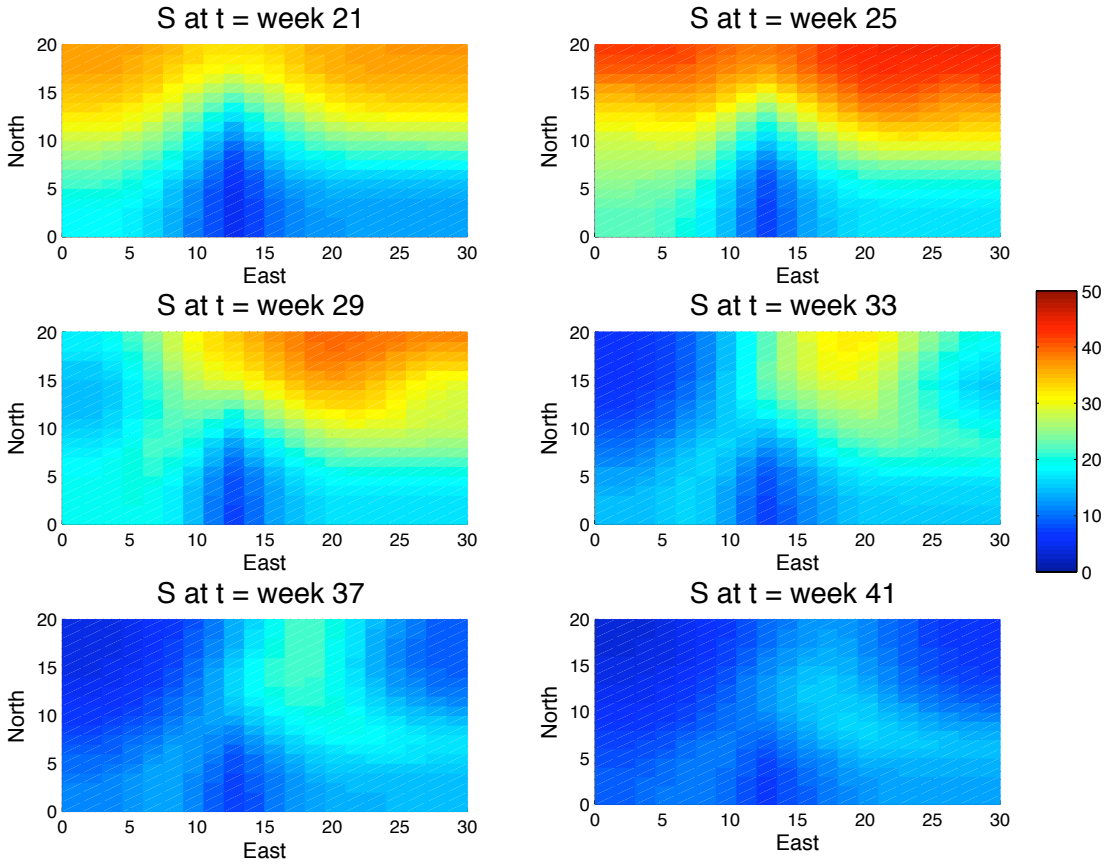


Figure C.38: Spatial Heterogeneity and LDT: Susceptible population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.25$.

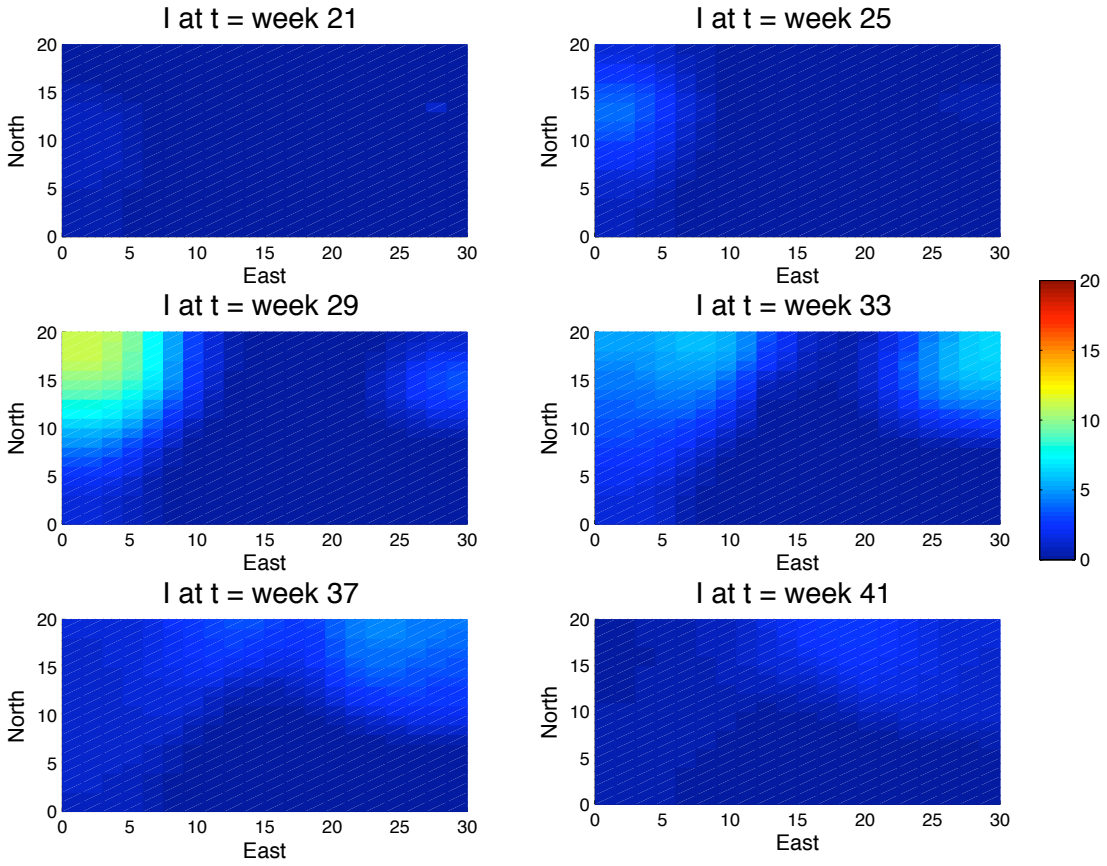


Figure C.39: Spatial Heterogeneity and LDT: Infected population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.25$.

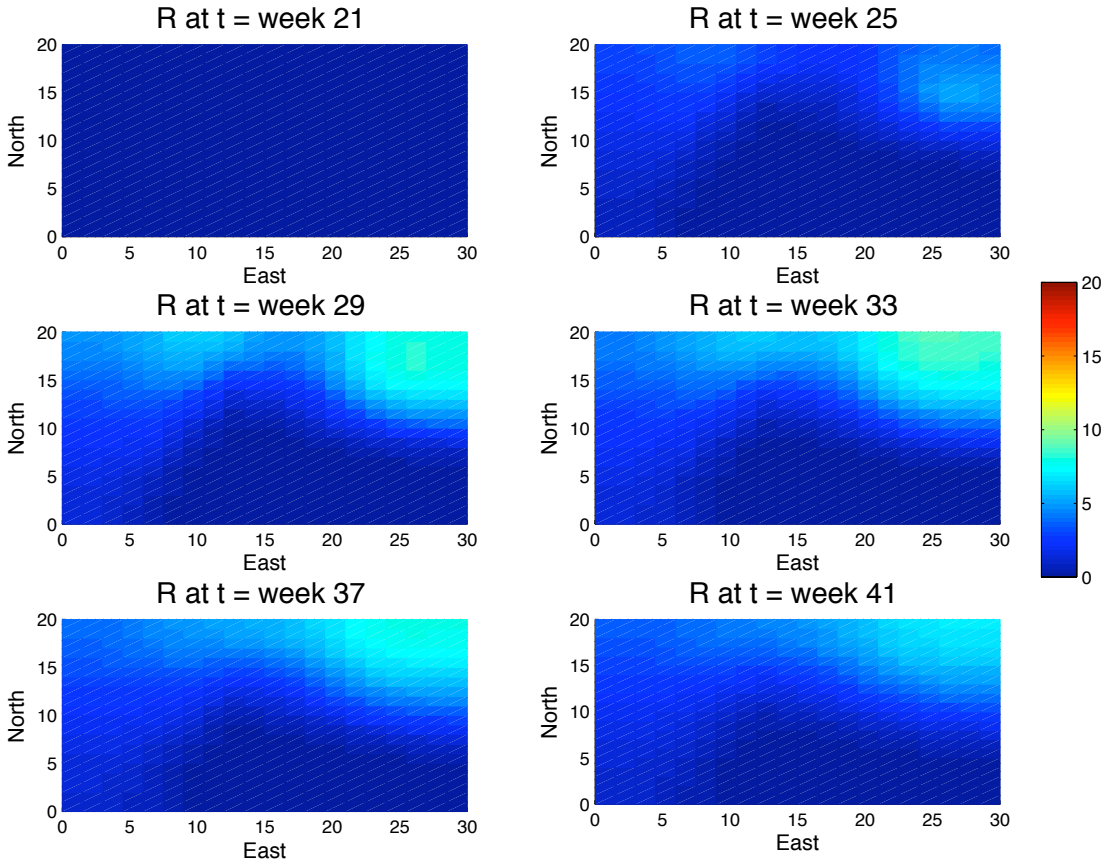


Figure C.40: Spatial Heterogeneity and LDT: Immune population during optimal 20-week vaccination schedule starting at $t = 21$ weeks. Infection is initiated at $t = 0$ weeks in southwest corner and LDT occurs at $t = 21$ weeks. Vaccine uptake parameter is assumed to be $a = 0.01$. Balancing coefficient is assumed to be $c = 0.25$.

Vita

Rachael Lynn Miller was born in Martinsburg, Pennsylvania, on October 10, 1982. She is the daughter of Larry and Susie Miller. After graduating from Central High School in 2001, she went on to Drexel University in Philadelphia where she studied mathematics and received her Bachelor of Science in December 2004. In January 2005, Rachael began her graduate career in the Department of Mathematics at the University of Tennessee, Knoxville. It was here that she met her husband, Michael Neilan, also a graduate student in Mathematics at the time. The two were married in August 2008 and both graduated the following year with doctorates in Mathematics.

Rachael Miller Neilan will continue her work in mathematical biology as a postdoctoral researcher in the Oceanography & Coastal Sciences Department at Louisiana State University in Baton Rouge, LA.