Protecting valuable resources using optimal control theory and feedback strategies for plant disease management

Elliott Hamilton Bussell

Department of Plant Sciences

University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Jesus College

September 2019

Declaration

This thesis is the the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

This thesis is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

This thesis does not exceed the prescribed word limit of 60,000 words (excluding bibliography, figures and appendices) as specified by the Degree Committee of the Faculty of Biology.

Elliott Bussell

September 2019

Abstract

Protecting valuable resources using optimal control theory and feedback strategies for plant disease management

Elliott Hamilton Bussell

Mathematical models of tree diseases often have little to say about how to manage established epidemics. Models often show that it is too late for successful disease eradication, but few study what management could still be beneficial. This study focusses on finding effective control strategies for managing sudden oak death, a tree disease caused by *Phytophthora ramorum*. Sudden oak death is a devastating disease spreading through forests in California and southwestern Oregon. The disease is well established and eradication is no longer possible. The ongoing spread of sudden oak death is threatening high value tree resources, including national parks, and culturally and ecologically important species like tanoak. In this thesis we show how the allocation of limited resources for controlling sudden oak death can be optimised to protect these valuable trees.

We use simple, approximate models of sudden oak death dynamics, to which we apply the mathematical framework of optimal control theory. Applying the optimised controls from the approximate model to a complex, spatial simulation model, we demonstrate that the framework finds effective strategies for protecting tanoak, whilst also conserving biodiversity. When applied to the problem of protecting Redwood National Park, which is under threat from a nearby outbreak of sudden oak death, the framework finds spatial strategies that balance protective barriers with control at the epidemic wavefront. Because of the number of variables in the system, computational and numerical limitations restrict the control optimisation to relatively simple approximate models. We show how a lack of accuracy in the approximate model can be accounted for by using model predictive control, from control systems engineering: an approach coupling feedback with optimal control theory. Continued surveillance of the complex system, and re-optimisation of the control strategy, ensures that the result remains close to optimal, and leads to highly effective disease management.

In this thesis we show how the machinery of optimal control theory can inform plant disease management, protecting valuable resources from sudden oak death. Incorporating feedback into the application of the resulting strategies ensures control remains effective over long timescales, and is robust to uncertainties and stochasticity in the system. Local management of sudden oak death is still possible, and our results show how this can be achieved.

Acknowledgements

Firstly, I would like to thank my supervisor, Nik Cunniffe. Thank you Nik for your guidance, encouragement and trust over the last few years. I have really appreciated your sharp eye and quick wit, and I've enjoyed working with you. I'd also like to thank my co-supervisor Chris Gilligan—it has been a pleasure working with you.

Thank you to the BBSRC for funding my PhD through a DTP studentship, and to the programme for the cohort events. I'm grateful to the Department of Plant Sciences, not forgetting Annie in the tea room. I'd also like to acknowledge my college, Jesus, for making me feel at home over the last 8 years. Particular thanks go to Noel Rutter, Jim Bellingham and Siân Dutton for your support during both my undergraduate course and PhD. Thanks also for supervising and interviewing opportunities.

I'd like to thank all the members, both past and present, of the Computational and Theoretical Epidemiology group, and the Epidemiology and Modelling group. You have all made this experience much more enjoyable. I'd particularly like to thank Clare Allen, Andrew Craig, Eleftherios Avramidis, Sally Hames and especially Ciara Dangerfield for your support, encouragement and friendship.

Last, but certainly not least, I'd like to thank all my friends and family. Thanks to my housemates for your fun and friendship. A big thank you to Lisa for proofreading. To my parents Zelma and Brian, my brother Luke, and Lisa: thank you all for your love and support over the years. I really appreciate your belief in me.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Infectious diseases of plants pose a serious economic and environmental threat across the globe. New pathogens are being introduced into novel environments at ever increasing rates, driven by increased international trade, climate change and agricultural intensification, causing significant damage to crops and natural environments (Anderson et al., 2004; Brasier, 2008). Yield losses of 20 to 30 % are seen globally across the major food crops of wheat, rice, maize, potato and soybean, with major implications for food security (Strange and Scott, 2005; Oerke, 2006; Savary et al., 2019). The economic cost from all crop losses to plant pathogens in the US has been estimated at $33 billion USD (Pimentel et al., 2005).

Plant pathogens also affect wild plants in the natural environment, with increasing numbers of new diseases reported in forest ecosystems (Stenlid et al., 2011). Important current examples include ash dieback in Europe (caused by *Hymenoscyphus fraxineus*; Kowalski and Holdenrieder, 2009; DEFRA, 2014b), sudden oak death (caused by *Phytophthora ramorum*) in the USA (Rizzo and Garbelotto, 2003) and Europe (Brasier and Webber, 2010), olive quick decline syndrome in southern Europe (caused by *Xylella fastidiosa*; Sicard et al., 2018), and sweet Chestnut blight in Europe (caused by *Cryphonectria parasitica*; Milgroom and Cortesi, 2004). Forests are a key part of a landscape, and provide important ecosystem services (Bateman et al., 2013). The economic value of forests is difficult to quantify, but the UK government has estimated that healthy forests contribute at least £5 billion to the UK economy per year, through forestry and social/environmental value (UK government, 2018). The long generation times of trees mean that resistance to disease develops slowly, or not at all, and so disease impacts have long-term implications (Boyd et al., 2013). The biodiversity of forests and the ecosystem services provided by trees are under severe threat from pests and diseases (Freer-Smith and Webber, 2017), and disease management approaches that take more consideration of these services are urgently required (Boyd et al., 2013)

1.1.1 Management efforts

There are a number of options available to decision makers for plant disease management, broadly grouped around four principles: exclusion, eradication, protection, and resistance (Maloy, 2005). With ever-increasing numbers of pathogens, decisions must be made about which management methods, plant species and diseases to prioritise. Efforts have been

made across the globe to help inform disease management, through schemes such as the UK Plant Health Risk Register (Baker et al., 2014), and regional plant protection organisations such as the European and Mediterranean Plant Protection Organisation (EPPO) that develop and advise governments on disease management and surveillance strategies (Maloy, 2005). But with very large numbers of plant species to protect, and increasing threats from pests and diseases, making these strategic decisions is not straightforward.

Mistakes in disease management decisions can be extremely costly, and major failures of management have been fairly common. In the UK, management of the Dutch elm disease outbreak in the 1970s was ultimately unsuccessful, with the loss of almost 30 million elm trees (Tomlinson and Potter, 2010). Initially it was believed that the epidemic would soon die out, and so the disease was left to run its course. In 1970 advice changed to recommend removing trees killed by the disease, but the need for larger scale efforts was not acknowledged until 1972. By this time disease containment would have been very expensive and unlikely to be successful, and so was not carried out. More recently, ash dieback has become established in the UK and it is now acknowledged that management is unlikely to make much difference to the long-term outcome across the country (Thomas, 2016). The total economic cost of the ash dieback outbreak, including management costs and loss of ecosystem services, has recently been estimated at £15 billion (Hill et al., 2019).

The citrus canker epidemic in Florida highlights the potential costs of these failures. The most recent introduction of the disease was discovered in Florida in 1995, after which an eradication program was quickly initiated. This program removed and destroyed all citrus trees within a certain radius of a known infection. This radius was initially 125 ft (38 m) to remove asymptomatic trees that had been exposed to the disease. The radius was increased in 1998 to 1,900 ft (579 m) as the initial radius did not remove enough trees to slow disease spread (Gottwald et al., 2001). In 2006, the ten year eradication program was abandoned once the disease was deemed endemic, after becoming widespread in commercial and residential citrus trees (Gottwald, 2007). A total of $1 billion USD had been spent on the program.

Management of disease epidemics is most successful when the scale of control matches the scale of the epidemic (Gilligan et al., 2007). Outbreaks have inherent spatial and temporal scales of spread, and control strategies that match these scales are the most effective. However, plants can be cryptically infected—where the host is infectious before symptoms appear—making estimation of future spread challenging (Thompson et al., 2016). Coupled with delays in disease reporting due to the high costs of surveillance (Parnell et al., 2014) as well as uncertainty surrounding rates of spread, determining this scale of management can be problematic. Also for these reasons, early detection of a new invasion is unlikely, but rapid deployment of resources is necessary for successful disease control

(Cunniffe et al., 2015). Successful management is therefore costly, and so decision makers require robust decision making processes.

1.2 Model-based disease management decisions

Increasingly, mathematical models are used to predict the impacts of disease, and assess potential management across human, animal and plant diseases (Thompson and Brooks-Pollock, 2019). Quantitative methods can be used to assist decision making by improved descriptive analysis of outbreaks, risk factors and response needs, as well as through forecasting and optimisation of interventions (Morgan, 2019). As examples, models of outbreaks have informed ebola vaccination campaigns in humans (Bellan et al., 2015), animal culling during the UK foot and mouth epidemic (Keeling et al., 2001), management of citrus canker in Florida (Gottwald and Irey, 2007), and of ash dieback in the UK (DEFRA, 2014a).

1.2.1 Simulation models

Models that simulate the expected course of an epidemic and explicitly capture the effects of interventions can quantify the impact of a potential management strategy (Lofgren et al., 2014). These simulation models, as we will refer to them, are designed to accurately forecast disease progression under a number of intervention scenarios being considered by a decision maker. As a result, simulation models have become important tools for assessing policy decisions relating to real-time management responses, as well as to increased preparedness for future threats. This allows decision makers to examine 'what if' scenarios, with all possible information available about disease impacts and uncertainty. However, to capture enough realism to be useful for guiding policy, simulation models must often be very complex (Basu and Andrews, 2013; Savary and Willocquet, 2014). This complexity ensures that the simulation model incorporates the many factors impacting on patterns and rates of epidemic spread, for example spatial heterogeneity and variation in host susceptibility (Keeling and Rohani, 2008; Anderson et al., 1986; Smith et al., 2002).

The complexity of simulation models can limit the extent to which management can be optimised. With many possible interventions that can be combined and potentially vary in space, time or according to disease risk, it becomes computationally infeasible to unambiguously determine the optimal strategy. This problem with optimisation in high-dimensional space is known as Bellman's 'curse of dimensionality' (Bellman, 1957). As a result, for most simulation models the only viable option is to test a small subset of plausible management strategies. This 'strategy testing' approach may be able to scan over a single

parameter, but the set of strategies to test cannot span the entire space of control options. This makes it difficult to have high confidence in the best-performing strategy.

The approach is nevertheless commonly used to inform policy decisions across human, animal and plant disease management. In human health, Jit et al. (2008) assess potential vaccination policies for human papillomavirus (HPV) in the UK, testing strategies that vary in vaccine coverage, age at vaccination, and whether the vaccine is given to boys as well as girls. The economic analysis carried out by Jit et al. (2008) was used by the Department of Health to inform decisions about distribution of the HPV vaccine. Similarly, models of Ebola virus disease outbreaks in west Africa have been used to assess potential interventions including hygienic burial and contact-tracing (Pandey et al., 2014), and more recently vaccination strategies for health care workers (Robert et al., 2019). In animal epidemiology, models were used to inform the response to the 2001 foot-and-mouth disease (FMD) outbreak (Keeling, 2005). The simulation model developed (Keeling et al., 2001) was used to assess different animal culling strategies, and later potential vaccination strategies (Keeling et al., 2003; Tildesley et al., 2006). Finally, in plant disease, simulation models have been used to assess potential host removal strategies for tree diseases of citrus (Cunniffe et al., 2014; Cunniffe et al., 2015; Hyatt-Twynam et al., 2017; Adrakey et al., 2017; Craig et al., 2018), and sudden oak death (Cunniffe et al., 2016).

In human disease outbreaks, for which modelling has played a prominent role, integration into the decision-making process can be slow because models are often built in reaction to ongoing epidemics (Rivers et al., 2019). The problem is amplified in plant disease management where limited funding, a lack of data, and poor surveillance means pathogens and their spread characteristics are almost always only identified once the disease is well-established. As a result, simulation models used in plant disease often simply show that it is too late for effective widespread eradication to remain a realistic proposition. This was the case with citrus canker in Florida (Gottwald and Irey, 2007), ash dieback in the UK (DEFRA, 2014b), and sudden oak death in California (Cunniffe et al., 2016). However, these models—and plant disease modellers in general—have said very little about how smaller-scale management, designed to achieve local goals rather than widespread eradication, could be made to be effective. Whilst studies might say what cannot be achieved, these simulation models have not been used to study what is still possible, for example through localised control or management to protect valuable resources.

1.3 Sudden oak death

The disease case study we will use throughout this thesis is the sudden oak death (SOD) epidemic in California. SOD is caused by the oomycete *Phytophthora ramorum*, which can

infect a very broad host range. Hosts most notably affected by the pathogen include oak, tanoak, larch, bay laurel and rhododendron, but over 100 plant species are susceptible to the disease (Grünwald et al., 2012). The disease effects vary depending on the host, but broadly split into two groups: lethal trunk infections and non-lethal foliar infections (Rizzo and Garbelotto, 2003). In tanoak and oak species the disease causes large cankers to form on the main stem, eventually leading to tree death. On 'spreader species', including rhododendron and bay laurel, *P. ramorum* can infect the host and sporulate, but this does not lead to host death. The pathogen spreads predominantly through short-distance rain splash dispersal of spores, but spores can be dispersed over longer distances by turbulent air currents, rivers and streams, or when carried by animals or human activity (Grünwald et al., 2012).

The disease was first detected in California in 1995 and has since spread widely along the west coast of the USA as shown in Figure 1.1 (Rizzo and Garbelotto, 2003; Meentemeyer et al., 2011). It has significantly impacted the nursery trade, and devastated populations of coast live oak and tanoak in California. SOD is currently found in areas covering over 2,000 km^2 in California (Grünwald et al., 2019), with an estimated $135M USD loss in property values attributed to the disease (Kovacs et al., 2011). In 2001 an isolated outbreak was identified in Curry County, Oregon, and in 2009 the pathogen was discovered in the UK where it is causing extensive mortality of larch (Brasier and Webber, 2010). In the UK the disease is known as ramorum disease or sudden larch death. The European and North American outbreaks are caused by different lineages of the pathogen, designated NA1 and NA2 for the North American, and EU1 and EU2 for the European pathogens (Grünwald et al., 2012). In 2016 the EU1 lineage was discovered in Oregon forests, which is problematic as it is of a different mating type to the NA1 and NA2 lineages (Grünwald et al., 2019). Whilst sexual reproduction of the pathogen has not yet been observed, this could lead to the much more rapid evolution of more aggressive forms of the pathogen.

1.3.1 Control of sudden oak death

Multiple scales of management are possible in any attempt to control SOD spread. For protection of individual high-value trees, for example at the urban-wildland interface, protective chemicals can be applied through sprays or by trunk injection (Garbelotto and Schmidt, 2009). At the landscape scale, the only management that has proved effective is the removal of hosts (Hansen et al., 2008). Modelling work has shown that state-wide eradication of SOD in California has long been impossible (Cunniffe et al., 2016). The method used by Cunniffe et al. (2016) to demonstrate this was to use a complex spatially-explicit simulation model of SOD spread, and test a number of state-wide management strategies: an example of the strategy testing approach we introduced earlier. They showed that the most effective strategies prioritised control at the epidemic wavefront, but needed

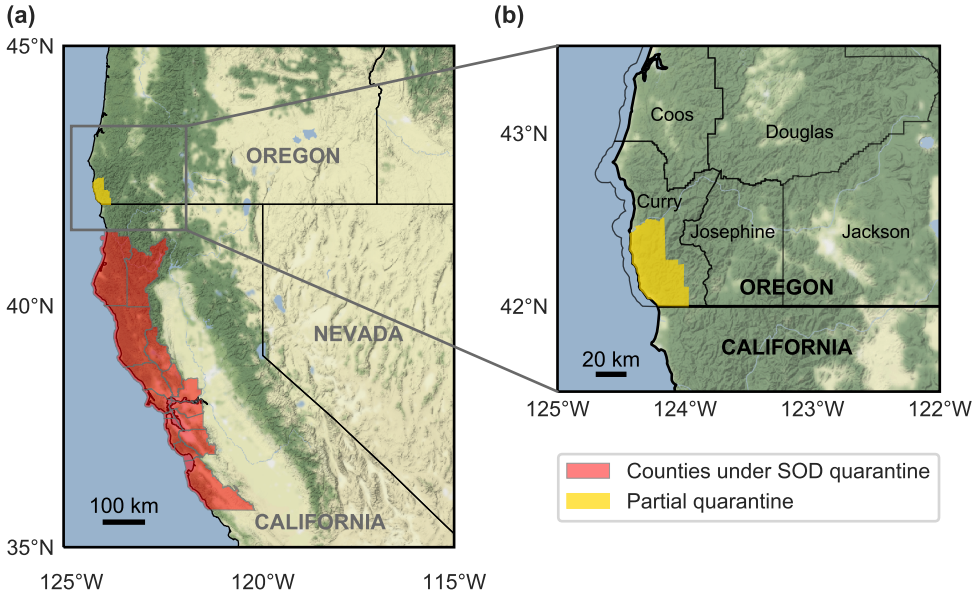

Figure 1.1: The current state of SOD spread in California and Oregon. Counties in California with confirmed SOD infestations are quarantined (shown in red in **(a)**). A partial quarantine of Curry county in Oregon has been implemented, as shown in **(b)**.

to have been implemented much earlier in the epidemic for widespread control to have been possible.

Despite eradication being unachievable, smaller-scale local management can still be beneficial. Since the Oregon outbreak was discovered in 2001, the disease has been actively managed, with $22.7M USD spent on identifying and treating infested sites (Grünwald et al., 2019). The management has been effective at slowing the spread of the infestation and containing the disease within Curry county (Figure 1.1(b)), with 2028 and 2038 being the estimated years of arrival into Coos county with and without control, respectively (Oregon Department of Forestry, 2019). In some locations in Oregon, control has shown that local eradication, whilst difficult, is possible (Hansen et al., 2019). These less ambitious local goals remain practically-relevant and achievable, but mathematical models have had little to say about how to deploy such management.

One reason for this may be that the objective of local management is less clear than that of eradication, and will vary depending on the wider forest management goals in each region. Local goals could involve slowing disease spread or protection of valuable resources (either particular hosts or regions), for example protecting culturally and ecologically important tanoak populations or slowing spread into national parks. These objectives must be considered alongside wider goals such as fire risk management and conservation. To date simulation models have only considered large-scale eradication goals, but how can models

be used to optimise local control? What strategies should decision makers deploy to manage SOD and protect local resources?

1.3.2 Models of sudden oak death

Many different models have been built to capture aspects of SOD spread. In this thesis we focus on SOD management strategies, and therefore require dynamic models that capture the drivers of disease spread into new regions and how these drivers are affected by possible management interventions. Much of SOD modelling, at least to start with, focussed on building risk maps. These maps show which areas are most likely to become infected, with the potential to be used to allocate control resources appropriately. Meentemeyer et al. (2004) used an expert informed, rule-based model to find high risk areas in California, based on weighted combinations of host distribution, temperature and moisture data. Later work by Kelly et al. (2007) compared environmental niche models like the model in Meentemeyer et al. (2004), with other classifiers including logistic regression and support vector machines, and similar models have been used in Oregon (Václavík et al., 2010). All these risk models predict the chance of future spread, but not the dynamics of those invasions into new regions. These models cannot therefore be used to investigate the dynamics of disease spread, and importantly what effect control would have on disease progression.

Further development of these ecological niche models incorporated dispersal estimation into the risk mapping (Meentemeyer et al., 2008). This in effect increases the risk of invasion in areas close to known infestations. Whilst this still did not capture the dynamics of the system, it begins to capture these dynamic effects. Models were also being developed to model the spread of *P. ramorum* in the UK. Analysis of susceptible host movement in the UK nursery trade showed a similarity to small-world and scale-free networks, suggesting that identifying and targeting key nodes in the network could manage the disease more effectively (Pautasso et al., 2008; Jeger et al., 2007). Harwood et al. (2009) developed a stochastic network model to capture the full dynamics of pathogen spread across the whole of the UK. The simulations however, did not directly model different host species, so could not be used to model the differing effects on multiple species, nor were they fitted to data.

Larger scale models of SOD spread seek to capture invasion dynamics at the landscape scale. Meentemeyer et al. (2011) developed a model of SOD invasion to predict spread across California through to 2030. This model was later used to assess different control strategies (Cunniffe et al., 2016). Another similar model (Tonini et al., 2018) integrates with the LANDIS-II forest simulation model (Scheller et al., 2007), designed to simulate forest disturbances. However for reasons of computational efficiency, as well as pragmatism in making very large scale predictions, both of these landscape scale models group host species together. In Meentemeyer et al. (2011) each simulation grid cell has a 'host index' that

captures the susceptibility and infectivity of the host composition in that cell. In LANDIS-II the disease model can only remove all hosts in a cohort of a given age in each cell. This means that small scale changes to host structure cannot be captured easily. However, both of these models capture sufficient dynamics to assess management strategies at the landscape scale.

Models of disease at the smaller scale of a forest stand are very limited in number. Brown and Allen-Diaz (2009) use 'stand reconstruction' to predict mortality within a forest stand. By looking for dead trees and symptomatic hosts in study plots, they estimate mortality rates and use these to predict future changes to stand structure. Again, dynamics are not captured here. Cobb et al. (2012) developed a differential equation model of SOD spread within a forest stand, capturing both invasion dynamics and differing mortality and infection rates by species and age class. In its current form this model does not include controls, but sufficient host dynamics are included that, with changes to the model, management strategies could be tested.

1.4 Aims and overview of the thesis

In this thesis, we seek to develop methods for using mathematical models to optimise local SOD management. More broadly, we aim to develop frameworks that can be used to optimise control on complex simulation models, improving on the strategy testing approach currently widely used. We will make use of the mathematical field of optimal control theory: a method for optimising time-dependent controls of dynamical systems. We will couple the predictive power of simulation models with mathematical results from applying optimal control theory, elevating abstract mathematical results into practical management strategies in a framework that could be used for policy.

We begin in the next chapter with an introduction to optimal control theory, and how it can be used to optimise time-dependent control strategies. In Chapter 3, we will apply this to a simple example case of protecting a high value region from a spreading epidemic, as a proxy for SOD invading an economically or culturally important region. This will demonstrate the optimisation capabilities of optimal control theory, but also its limitations for real-world practical management as a result of the simplicity of the underlying models. We will then develop frameworks for coupling such optimal control results with simulation models in Chapter 4, using the predictive capabilities of simulation models to account for inaccuracies in the optimal control model. We show that model predictive control—a framework incorporating feedback between a simulation model and a simpler, approximate model—finds the best disease management strategies. The feedback framework is applied to the problem of tanoak protection in Chapters 5 and 6, showing that the framework provides

robust strategies that limit the impacts of the worst-case scenario epidemics. In Chapter 7, we apply optimal control theory to the protection of a valuable region, Redwood National Park, showing how complex, spatially resolved control strategies can be identified. In the Discussion (Chapter 8) we explain how our frameworks could be used to improve plant disease outbreak responses, and where there are avenues for further study. Overall the work demonstrates how the frameworks we develop allow the insights of optimal control theory to be applied in a practical setting, with relevance to disease management across human, animal and plant health. Our unique contribution is to couple disease models with optimal control theory and systems engineering to find practical local management strategies for SOD.

2 Optimal control theory

2.1 Introduction

In Chapter 1 we highlighted how the optimisation of plant disease management can be mathematically and computationally challenging. In this chapter we will introduce the mathematical field of optimal control theory (OCT), and how it can be used to optimise epidemiological controls. Poorly designed management strategies can lead to expensive failures of control, for example as happened with Dutch elm disease in the UK in the 1970s (Tomlinson and Potter, 2010), and with citrus canker in Florida in the 2000s (Gottwald, 2007). OCT provides a framework for finding time-dependent, optimal strategies for simple models, which could be used to systematically generate potential management strategies for use in the real-world. Alongside expert-informed strategies this could enable a more mathematically robust system for making disease management policy decisions.

In this chapter we review the use of OCT in the epidemiology literature. We first describe the formulation of the optimal control problem, and give an overview of the mathematical background. We focus here on a review of applications of OCT in epidemiology and related fields, rather than the underlying mathematical theory. To set the scene for analysing a simple disease model in the next chapter, we then describe two main numerical methods that are commonly used to find solutions to the optimal control problem.

2.2 Background

Optimal control theory (OCT) is the field of mathematics concerned with finding time-varying inputs to dynamical systems, optimised to maximise (or minimise) some performance metric. Whilst the standard approach applies to ODEs, the dynamical system could also be a system of PDEs, or difference equations. The system is controlled by an input variable, or combination of variables, that can be varied from outside the system to adjust the dynamics. The field developed as an extension to the calculus of variations, with roots also in classical control theory, random processes, and linear and nonlinear programming (Bryson, 1996). In the 1950s, Bellman developed dynamic programming, a method for finding optimal controls by recursively solving smaller problems (Bellman, 1957). These methods struggle to solve realistic problems because of the 'curse of dimensionality', where, as the number of state and

control variables increases, the memory required to solve the recursive problem becomes impractical.

Later work by Pontryagin et al. (1962) extended work from the calculus of variations to find a necessary condition for the maximising path. Since whether an individual control-state trajectory satisfies the condition can be tested, this method avoids the 'curse of dimensionality' with the Bellman equation. There is no need to recurse over the entire state space. The necessary condition is contained in Pontryagin's famous 'maximum principle', the main theory underlying much of OCT. We will now summarise this principle as it applies to ODEs, but we do not describe the mathematics here. Further details can be found in Lenhart and Workman (2007) or Hocking (1991).

The maximum principle relates to a differential equation system for a state vector $\mathbf{x}$, described as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t); \quad \mathbf{x}(0) = \mathbf{x}_0; \quad t \in [0, T] \tag{2.1}$$

where $\mathbf{u}(t)$ is a time dependent control vector. The control function $\mathbf{u}(t)$ must be chosen in order to maximise the objective function J:

$$J = \Psi(\mathbf{x}(T)) + \int_0^T L(\mathbf{x}(t), \mathbf{u}(t), t)\ dt \tag{2.2}$$

where Ψ is a salvage term at the final time T, or payoff function, and L is the Lagrangian of the problem. The overall optimisation problem is therefore:

$$\max_{\mathbf{u}(t)} \quad \Psi(\mathbf{x}(T)) + \int_0^T L(\mathbf{x}(t), \mathbf{u}(t), t)\ dt \tag{2.3a}$$

$$\text{subject to} \quad \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \tag{2.3b}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{2.3c}$$

where the state dynamics are constraints to the maximisation problem.

The maximum principle states that if $\mathbf{u}^*(t)$ and $\mathbf{x}^*(t)$ are optimal for this system, then there must exist an adjoint system of variables $\boldsymbol{\lambda}(t)$ such that:

$$H(t, \mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t)) \leq H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \tag{2.4}$$

for all time $t \in [0, T]$. The Hamiltonian H is given by:

$$H(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t)) = L(\mathbf{x}(t), \mathbf{u}(t), t) + \boldsymbol{\lambda}(t) \cdot \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \tag{2.5}$$

and the adjoint system satisfies:

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t))}{\partial \mathbf{x}} \tag{2.6a}$$

$$\boldsymbol{\lambda}(T) = \frac{\partial \Psi(\mathbf{x}(T))}{\partial \mathbf{x}} \, . \tag{2.6b}$$

The adjoint system is a set of time-dependent variables that are equivalent to Lagrange multipliers. In economic terms they are shadow prices, representing the cost in the objective function associated with a change in the relevant state variable. This is useful in the maximisation problem as, through maximisation of the Hamiltonian, the adjoint system sets the relative importance of the control functions. The terminal constraint in the adjoint system is known as the transversality condition.

This maximum principle can then be used to find solutions to the optimisation problem described above. In Example 1 we illustrate this problem formulation for a simple SIR model, before reviewing how OCT has been used in epidemiology in the next section.

Example 1: Setting up a simple SIR optimal control problem.

As a more concrete illustration of how the problem formulation can be applied in an epidemiological setting, we here formulate the optimal control problem for a simple SIR model. The model tracks numbers of susceptible and infectious individuals, and the control consists of removing infected hosts (roguing). The state system (Equation 2.1 previously) is given by:

$$\dot{S} = -\beta SI \tag{2.7a}$$

$$\dot{I} = \beta SI - \mu(t) I \tag{2.7b}$$

where β is the infection rate, and $\mu(t)$ is the time dependent roguing rate. In the mathematical formulation this is $u(t)$.

We seek to find an optimal roguing strategy ($\mu^*(t)$) that minimises the number of infected hosts after the final time T, combined with the total cost of roguing. The objective we choose is given by:

$$J = I(T) + \int_0^T c\mu(t) I \; dt \tag{2.8}$$

where c is the relative cost of roguing. Comparing with Equation 2.2, the salvage term is $I(T)$, and the Lagrangian is the roguing cost $c\mu(t)I$. Note here that this is a minimisation problem, but the framework is exactly the same as for the maximisation problem considered previously, apart from the Hamiltonian being minimised when the control is optimal.

The Hamiltonian is formed by introducing adjoint variables for each state in the system. We define these as λ_S and λ_I for S and I respectively. Using Equation 2.5 the Hamiltonian is found to be:

$$H = c\mu(t)I + \lambda_S(-\beta SI) + \lambda_I(\beta SI - \mu(t)I) \; . \tag{2.9}$$

Using Equation 2.6(a), differentiating the Hamiltonian with respect to each state variable gives the dynamics of the adjoint variables:

$$\dot{\lambda}_S = \lambda_S \beta I - \lambda_I \beta I \tag{2.10a}$$

$$\dot{\lambda}_I = -c\mu(t) + \lambda_S \beta S - \lambda_I \left(\beta S - \mu(t)\right) \; . \tag{2.10b}$$

The optimal control problem is completed by the transversality constraints. These give the terminal time conditions for the adjoint variables. Using Equation 2.6(b) and the payoff term we have already defined, these are given by:

$$\lambda_S(T) = 0 \tag{2.11a}$$

$$\lambda_I(T) = 1 \; . \tag{2.11b}$$

This completes the formulation of the optimal control problem. We have a state system with initial conditions, an adjoint system with terminal conditions, and a Hamiltonian that is minimised when the roguing strategy is optimal. We will demonstrate how to numerically solve this problem in Example 2 later in the chapter.

2.3 OCT within epidemiology

When deciding how to manage human, animal or plant epidemics, public health or environmental decision makers must take into account economic constraints. OCT optimises control strategies whilst ensuring adherence to constraints and minimisation of total costs, and is therefore a potentially useful tool for designing epidemic interventions. Successful disease control relies on applying the most effective control methods at the correct time, and at a sufficient level. OCT is an effective tool for balancing disease management objectives with economic and logistic constraints, provided that the overall goal of the management program is well defined. OCT has been applied to models of many diseases, including HIV (Kirschner et al., 1997), mosaic virus in *Jatropha curcas* plants (Al Basir et al., 2017), and sudden oak death (Ndeffo Mbah and Gilligan, 2010b). In this section we will review how OCT has developed within the field of epidemiology, discussing outstanding limitations and problems. The OCT literature is extensive, and so we will focus on the key areas relevant to this thesis.

Early work & general principles

Early work using OCT on epidemiological models was carried out by Sethi and Staats (1978), who optimised levels of vaccination and treatment in simple SEIR type models. The work is analytic and finds cases where the control can be classified as 'bang-bang'. Bang-bang controls occur when the state dynamics and the Lagrangian, and so the Hamiltonian (Equation 2.5), are linear in the control functions. The Hamiltonian is then maximised (or minimised) by controls that switch between their maximum and minimum values, giving the so-called bang-bang strategies, where each control is either on or off. In bang-bang strategies there are often important switching times, when some or all of the controls switch from on to off or vice versa. These strategies have arisen repeatedly in disease control, for example by Forster and Gilligan (2007) for a plant disease spreading through an agricultural landscape, and Panetta and Fister (2003) for cancer drug treatments.

As OCT was increasingly used within epidemiology, the underlying models increased in complexity. Behncke (2000) developed models with more complex forms of control, including vaccination, quarantine and screening, and health-promotion campaigns. The wider range of interventions modelled are more realistic than those used in previous studies, but the models are analytic and abstract in nature. The models used a general infection interaction $f(S, I)$, so that the results are valid in the general case for frequency or density dependence, or any other interaction term. As well as this, the effects of screening and health-promotion campaigns are also included through generic functions, so do not capture any specific method of control. However, the control strategies found here do introduce the important concept of 'front-loading', where additional resources are allocated early in the epidemic when control can be more effective. The importance of this early timing of control for successful management has also been shown for control of sudden oak death (Cunniffe et al., 2016).

Bioeconomics

Throughout its history, OCT has been closely connected with the field of economics (Weber and Kryazhimskiy, 2011), for example in fisheries economics (Clark and Munro, 1975). Consideration of economics is important for optimising epidemiological controls (Perrings et al., 2014), since interventions must be cost-effective, and resources are often limited. In one study, OCT was used to find cost-effective human papillomavirus vaccination strategies (Brown and White, 2011), showing how time-dependent strategies can be found that balance the costs of administering vaccines with the costs of treating infected individuals. In plant epidemiology, a recent study shows how control strategies for vector transmitted diseases can balance multiple economic costs (Bokil et al., 2019), including control costs, yield loss

and long-term costs from insecticide use. Work on sudden oak death has shown how OCT can find an optimal balance when allocating limited resources between surveillance and eradication measures (Ndeffo Mbah and Gilligan, 2010a).

Bioeconomic studies, such as the ones considered in the previous paragraph, rely on a clearly defined objective of control, with all costs quantified in the same units. By setting the relative contributions for each cost term, for example surveillance, vaccination and treatment costs, the objective function chosen defines what is to be considered cost-effective. Epanchin-Niell (2017) argues that in bioinvasion management a key gap in current understanding is how to value environmental benefits, such as ecosystem services and biodiversity, alongside the costs and effectiveness of controls. The same difficulties are present in the epidemiological literature, but improved valuation and consideration of these services in determining control strategies are urgently required (Boyd et al., 2013). Choosing an objective function that balances costs meaningfully is important, since the choice of the objective function can make significant changes to the optimal strategy (Probert et al., 2016). In human disease, cost-effectiveness analyses are often based on quality adjusted life years (Whitehead and Ali, 2010). A similar concept could perhaps be used for plant and animal diseases, including calculations of yield losses (Savary et al., 2012; Savary et al., 2019) as well as effects on welfare, biodiversity and tourism for example (Boyd et al., 2013).

State of the art and outstanding questions

As the use of OCT in epidemiology progressed, the complexity of the underlying control models increased to capture more realism and ask more applied questions. As examples, OCT has been used in models of vector-borne diseases applied to malaria (Blayneh et al., 2009), vaccination rates against *Clostridium difficile* in a hospital setting (Stephenson et al., 2017), and, as previously mentioned, balancing of detection and eradication controls for sudden oak death (Ndeffo Mbah and Gilligan, 2010a). Despite the additional complexity and realism, the underlying models in these and other studies are still relatively simplistic, and cannot be expected to capture complex realistic dynamics. OCT results have therefore rarely been tested in the field.

Occasionally, though, OCT results have been tested using more complex models to relax the assumptions in the simplistic OCT models. Forster and Gilligan (2007) find that in a mean field epidemic model a switching strategy is optimal. Forster and Gilligan also test their mean field results on a spatial contact process model, and find that if the switch time is not known precisely then the OCT strategy can be worse than a simple constant strategy. Another study finds optimal control strategies for chlamydia, and tests the non-spatial results on a spatial network simulation (Clarke et al., 2013). The strategies found using OCT result

in rapidly switching controls that perform much worse in the simulation model than in the OCT model. For effective control in the real world, in this study the OCT model should be constrained to avoid these rapid switches. Whilst a few examples do test OCT results on more realistic systems, this is far from the norm, and understanding how epidemiological OCT results can be mapped to more realistic models, or even the real world, is a significant gap in the literature.

Within plant epidemiology, the effect of space is highly important in both modelling the spread of disease, and designing effective management strategies (Ostfeld et al., 2005; Plantegenest et al., 2007). Some OCT studies have used simple metapopulation models to capture some element of spatial dynamics. Ndeffo Mbah and Gilligan (2014) optimise control on a one-dimensional lattice, showing that optimal control tracks the wavefront of a spreading epidemic. An alternative approach to finding optimal spatial strategies is to use partial differential equation (PDE) models, as used by Neilan and Lenhart (2011) for a model of rabies in raccoons. Other studies resort to significantly simplified epidemiological models and optimisations, for example Epanchin-Niell and Wilen (2012) use a nearest neighbour spread model. The spread model is formulated as an inequality system, allowing the control optimisation to be carried out using integer programming. Spatial control strategies can be difficult to optimise, but use of simpler models could result in strategies that are not robust when additional realism is included. Without explicit testing of the OCT strategies on a realistic spatial model, how can a policy maker be sure the strategies are robust?

2.4 Optimisation methods

There are many numerical methods for solving the optimisation problem described in Equations 2.3 (p. 28). These methods can be grouped into two main classes: indirect and direct methods. Indirect methods find roots of the necessary condition given by the Pontryagin maximum principle, i.e. state-control trajectories that satisfy the necessary condition. Direct methods on the other hand, find a sequence of controls that minimise the objective function without using the adjoint system (Betts, 2010). We will briefly describe a formulation using each of these methods, but refer to Betts (2010) for a comprehensive discussion of numerical methods for optimal control problems. Example 2 continues the previous example, showing how the system can be optimised using the direct and indirect methods.

2.4.1 Indirect formulation

Indirect methods construct the necessary conditions given by the Pontryagin maximum principle for the system in question. This involves forming the Hamiltonian and the adjoint

system. The necessary conditions are then used to find an expression for $\mathbf{u}^*(t)$ in terms of the optimal state ($\mathbf{x}^*(t)$) and the adjoint ($\boldsymbol{\lambda}(t)$). This can be used to solve for $\mathbf{x}^*(t)$ and $\boldsymbol{\lambda}(t)$ with the boundary conditions, finally allowing $\mathbf{u}^*(t)$ to be calculated. The process effectively optimises the system, then discretises to solve for the optimal control.

More specifically, since the initial conditions for the state system, and the terminal conditions for the adjoint system are known (Equations 2.3c and 2.6b, p. 28), then the necessary conditions become a two point boundary problem. This state-adjoint system will be solved when the control function is optimal ($\mathbf{u}^*(t)$). The forward-backward sweep method (FBSM) is a simple numerical algorithm for solving this formulation of the optimal control problem (Lenhart and Workman, 2007). The method is described in Algorithm 1.

The FBSM is based in the necessary conditions for optimality. This means that the results are closely connected to the underlying mathematics of the optimal control theory. This can help to give more insight into the meaning of the resulting controls, since the underlying dynamics that influence the control are analytically described. In some cases this may help to make general statements about the nature of optimal control that could lead to a rule of thumb. For example, Sethi and Staats (1978) showed that there can be at most a single switch in the optimal strategy, a general statement that could not be found from numerical solutions. There are however, a number of limitations to using the indirect formulation. Firstly, this close connection to the optimality conditions means that significant work is necessary to set up the optimisation problem. The Hamiltonian, adjoint dynamics, and transversality conditions must all be derived. This can be mathematically challenging, and is inherently inflexible since the equations must be derived for each new problem (Betts, 2010). Further to this, there can be issues with the convergence of the FBSM. The convergence can be highly sensitive to the initial guess for the optimal control, and the method may never converge to a solution (Betts, 2010).

Algorithm 1: The forward-backward sweep method (FBSM) algorithm from Lenhart and Workman (2007), for solving optimal control problems using the indirect formulation.

1. Make an initial estimate for the control function, $\mathbf{u}(t)$.
2. Using this control and the initial state condition $\mathbf{x}(0) = \mathbf{x}_0$, solve the state system forward in time.
3. Using the adjoint terminal condition (Equation 2.6b), and the values for control and state, solve the adjoint system backwards in time.
4. Update the control $\mathbf{u}(t)$ using the new state and adjoint values, by maximising the Hamiltonian.
5. Check for convergence. If the system has not converged to an optimal control, return to step 2 using the updated control.

2.4.2 Direct formulation

The direct formulation differs from the indirect in that it does not rely on the derivation of the Hamiltonian or adjoint system. Direct formulation methods discretise the dynamic system first, and then optimise the discretised variables, rather than optimising using the optimality conditions. One method for solving a direct formulation problem is the direct transcription process. This method discretises the state and control in time, and treats the values of the state and control at these discrete times as optimisation variables in a nonlinear programming (NLP) problem. Within the NLP problem, the state dynamics and initial conditions are included as constraints on the NLP variables, and the optimisation is carried out to maximise (or minimise) the objective value.

It may seem excessive to be directly optimising all the state variables, but because each variable only directly influences variables that are close in time, the problem becomes large but sparse. This sparsity can be exploited by numerical optimisation routines, and is often simpler to solve than a small, dense problem in which the state is not discretised and optimised (Betts, 2010). The state dynamics constraints are constructed as a discretisation of the state ODE system, often using an implicit Runge-Kutta scheme. For example, following the standard approach in Betts (2010), if the state-control system is discretised on M grid points, the NLP variables are given by:

$$\mathbf{y}^{\mathrm{T}} = (\mathbf{x}_1, \mathbf{u}_1, \ldots \mathbf{x}_M, \mathbf{u}_M) \ . \tag{2.12}$$

Using an Euler discretisation scheme, the state ODEs ($\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$) can be approximated by the following NLP constraints:

$$\mathbf{0} = \mathbf{x}_{k+1} - \mathbf{x}_k - h_k \mathbf{f}_k \equiv \mathbf{c}(\mathbf{y}) \ . \tag{2.13}$$

The NLP variables $\mathbf{y}$ are then optimised subject to the constraints $\mathbf{c}(\mathbf{y})$ using an NLP solver.

The advantages of the direct over the indirect formulation are in its flexibility and robustness. Since the Hamiltonian and adjoint are not required, and gradients can be estimated using finite differences, the method can be used for any system without analytic derivatives. Furthermore, there is extensive literature and software available for solving large, sparse NLP problems. This means the optimisations can be fast, efficient and robust. Whilst the method does still require an initial guess for both the control and the state variables, the NLP problem has a much larger region of convergence than the root finding in indirect methods (Betts, 2010). Note that neither the direct nor the indirect method guarantee a globally optimal solution, and using different initial estimates of the optimal control in the optimisation process can converge to different solutions. Different solutions can be compared to find the optimal strategy.

Example 2: Applying the indirect and direct formulations to the simple SIR model.

In Example 1 we formulated the optimal control problem for roguing in a simple SIR model. Here we will numerically solve this problem, using both indirect and direct methods. First, as a reminder, the optimal control problem was given by:

$$\min_{\mu(t)} \quad J = I(T) + \int_0^T c\mu(t)I \; dt \tag{2.14a}$$

$$\text{subject to} \quad \dot{S} = -\beta SI \tag{2.14b}$$

$$\dot{I} = \beta SI - \mu(t)I \; . \tag{2.14c}$$

We use the initial conditions of $S(0) = 99$ and $I(0) = 1$, and the parameterisation β =0.1 host^{-1} t^{-1}, $c = 1.2$, and we set a maximum roguing rate $\mu_{\max}$ of 3 t^{-1}.

Indirect method

First we solve the problem above using the FBSM described in Algorithm 1. In Example 1 we derived the adjoint system dynamics and terminal conditions, and the Hamiltonian. The FBSM solves the state dynamics forward in time from the initial conditions, then uses this state to solve the adjoint dynamics backwards in time from the terminal conditions. The adjoint dynamics were given by:

$$\dot{\lambda}_S = \lambda_S \beta I - \lambda_I \beta I \tag{2.15a}$$

$$\dot{\lambda}_I = -c\mu(t) + \lambda_S \beta S - \lambda_I \left(\beta S - \mu(t)\right) \; . \tag{2.15b}$$

The state and adjoint values are then used to minimise the Hamiltonian from Equation 2.9. Because the Hamiltonian is linear in the controls, it is minimised by the following control:

$$\mu(t) = \begin{cases} 0 & \text{if } \lambda_I < c \\ \mu_{\max} & \text{if } \lambda_I > c \; . \end{cases} \tag{2.16}$$

Note that if $\lambda_I = c$ the the control is singular, and undefined by the equations we have derived here. In some cases singular controls can be derived, but here this will not be necessary. We use an initial estimate of the roguing strategy to start the iterative FBSM. Here we use zero control at all times, i.e. $\mu(t) = 0$. The results in Figure 2.1 show the converged solution, finding that the optimal control is bang-bang and switching. Control is maximal for approximately 90 % of the time. After this the benefits from roguing infected hosts are outweighed by the costs of control, and no further roguing is carried out.

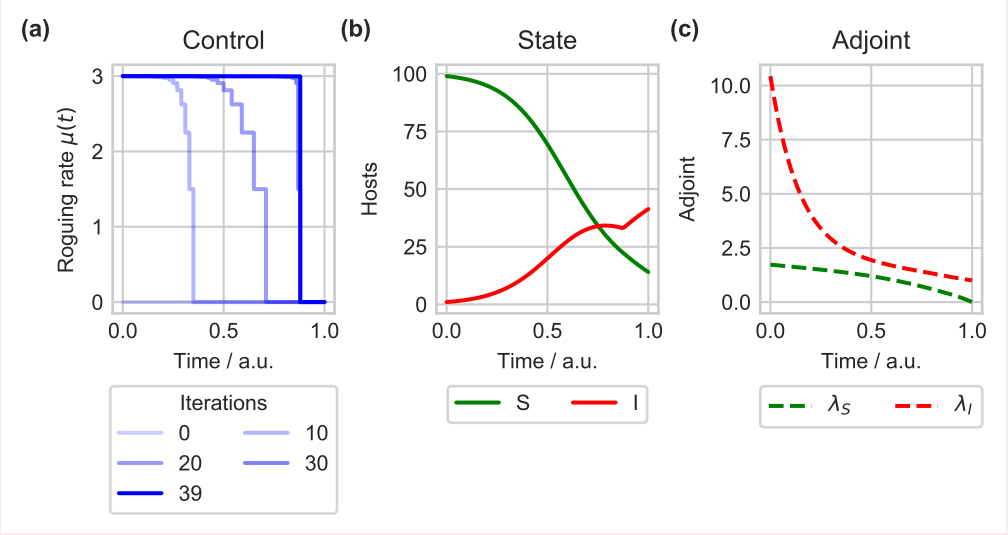

Figure 2.1: Solution to the optimal roguing problem using the FBSM approach. Control is initiated at zero, and in 39 iterations converges to a switching, bang-bang control as shown in **(a)**. **(b)** and **(c)** show the state and adjoint systems corresponding to the optimal control.

Direct method

Next we will show the solution to the same problem using the direct transcription method. First the state system and control must be discretised. Discretising onto $M+1$ grid points, we have the following variables to optimise:

$$\mathbf{y}^{\mathrm{T}} = (S_0, I_0, \mu_0, \dots S_M, I_M, \mu_M) \tag{2.17}$$

where:

$$S_i = S(t = iT/M)\,; \quad I_i = I(t = iT/M)\,; \quad \mu_i = \mu(t = iT/M)\,; \tag{2.18}$$

$$i \in \{0, \dots, M\}\,. \tag{2.19}$$

The state dynamics are included as constraints to the minimisation. For simplicity we here use an Euler integration step between grid points:

$$S_{k+1} = S_k + \frac{T}{M}\left(-\beta S_k I_k\right) \tag{2.20a}$$

$$I_{k+1} = I_k + \frac{T}{M}\left(\beta S_k I_k - \mu_k I_k\right)\,. \tag{2.20b}$$

These equality constraints, as well as the initial conditions, are included in the minimisation of the objective, which is also discretised:

$$J = I_M + \sum_{j=0}^{M-1} c\mu_j I_j \frac{T}{M}\,. \tag{2.21}$$

The Python minimize function from the SciPy library (Jones et al., 2001–) is used to perform the optimisation, and the results are shown in Figure 2.2. Both methods converge to the same optimal solution, but there are numerical inaccuracies in the direct method because of the simple discretisation scheme chosen here.

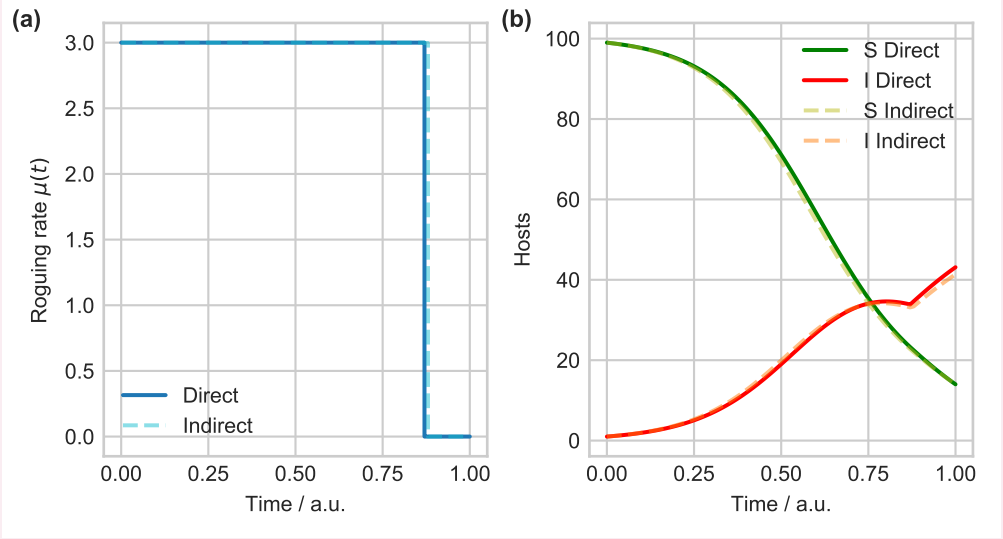

Figure 2.2: Solution to the optimal roguing problem using the direct transcription approach. Dashed lines show the indirect solution found previously. Numerical inaccuracies in the Euler approximation used lead to small differences between the methods, but both approaches converge on the same solution. The direct method loses some accuracy, but does not require computation of the adjoint system or the Hamiltonian.

The numerical inaccuracies are due to the simple Euler discretisation, here used for illustration. In later chapters we use more accurate integration schemes, and as we will show in the next chapter, these significantly reduce the numerical differences between the methods. Also note that this scenario is meant purely to illustrate the methods; control starts unrealistically fast at $t = 0$ with a short time horizon that may not be useful for a real epidemic. In practice, the objective would also take into account the number of removed individuals. Due to the switching function for the optimal control (Equation 2.16), the optimal strategy will be bang-bang. However, switches may not be seen when using different timescales for control.

2.5 Conclusions

In conclusion, OCT has allowed epidemiology to optimise control strategies whilst taking economic factors into account. Strategies often involve bang-bang solutions, where control is either maximal or minimal, leading to policies with precise switching times when priorities change. For finding these optimal strategies, there are two distinct classes of numerical method: the indirect method that is more closely connected to the underlying mathematics but that requires more analysis, and the direct method that is more flexible and robust but provides less insight.

3 A simple case: protecting a high value region

3.1 Introduction

Whilst established diseases can be difficult or impossible to eradicate (e.g. Gottwald and Irey, 2007), local control to protect high value resources can still be effective (e.g. Hansen et al., 2019). The sudden oak death (SOD) epidemic in California cannot be controlled at the landscape scale (Cunniffe et al., 2016), but there are still highly valuable local resources that could be protected from the disease. This includes commercially valuable timber stocks and areas important ecologically or for tourism, such as national parks. Optimally allocating limited control resources could significantly improve protection of valuable regions, and we seek here to answer how to characterise these optimal strategies. In this chapter we will investigate how the optimal control methods introduced in the previous chapter can be applied to a relatively simple metapopulation model of disease invasion into a high value region. We will use the direct and indirect optimisation methods described in the previous chapter to find time-dependent roguing strategies that best protect the high value region.

In this chapter we use OCT to find an optimal time varying allocation of limited control resources, balancing control in a buffer region and a high value region. We will show how these controls can be optimally deployed to minimise disease in the high value region. We show that two main strategies arise: one prioritising control in the high value region, and the other switching prioritisations during the course of the epidemic. The sensitivity of the control strategy to the model parameterisation is tested, and we show that OCT results do not always remain effective when parameters are not known accurately. We use the optimal strategies to compare the direct and indirect numerical optimisation approaches, finding that the direct method is more robust and reliable.

3.2 Methods

In this section we will develop a relatively simple deterministic model that captures the dynamics of a pathogen threatening to invade a high value region. The model will incorporate control through removal of infected hosts, and we will set up an optimal control problem to determine where limited resources should be allocated to this control over time.

The optimisation problem will be solved using both the direct and indirect formulations described in the previous chapter.

3.2.1 Invasion model

The model abstracts the situation of an invading pathogen by splitting the host landscape into three regions: a generally infested area where the disease is well established, a buffer region in which the disease has not yet established, and a high value region that is to be protected. To reduce the number of state variables and parameters the infested region is modelled as a source of external inoculum, generating a constant force of infection on the buffer region. The buffer and high value regions are modelled as well-mixed patches, meaning the only spatial component is the between-patch coupling. An SIR model, in which hosts can be susceptible to the disease (S), infected by the disease and infectious (I), or removed by the disease or control (R), is used for the epidemic dynamics (Keeling and Rohani, 2008). As the model is based on SOD dynamics, removal represents both death caused by the pathogen, and roguing as part of the disease management. This roguing is the control method that will be optimised. Note that we do not model any spread caused by the implementation of roguing, for example through use of infected tools or movement of infected material. A schematic of the model is shown in Figure 3.1.

Each patch has a fixed population size (N_B and N_V for the buffer and high value regions respectively including susceptible, infected and removed hosts) that is constant over time:

$$N_B = S_B(t) + I_B(t) + R_B(t) \tag{3.1a}$$

$$N_V = S_V(t) + I_V(t) + R_V(t) \tag{3.1b}$$

where the subscripts B and V refer to the buffer and high value patches respectively. The two patches are linked by a coupling constant ϵ. The within-patch transmission rate is given by β, leading to the following system of ODEs:

$$\dot{S}_B = -\mathcal{F}S_B - \beta I_B S_B - \epsilon\beta I_V S_B \tag{3.2a}$$

$$\dot{I}_B = \mathcal{F}S_B + \beta I_B S_B + \epsilon\beta I_V S_B - \mu I_B - f_B(t)\eta I_B \tag{3.2b}$$

$$\dot{S}_V = -\beta I_V S_V - \epsilon\beta I_B S_V \tag{3.2c}$$

$$\dot{I}_V = \beta I_V S_V + \epsilon\beta I_B S_V - \mu I_V - f_V(t)\eta I_V \tag{3.2d}$$

The external force of infection from the infested region is given by $\mathcal{F}$. The infectious period of the disease in the absence of control is given by $1/\mu$, and hosts are removed by roguing at a rate η. The time dependent control inputs $f_B(t)$ and $f_V(t)$ are the proportion of infected hosts that are being controlled at that time in the buffer and high value regions respectively.

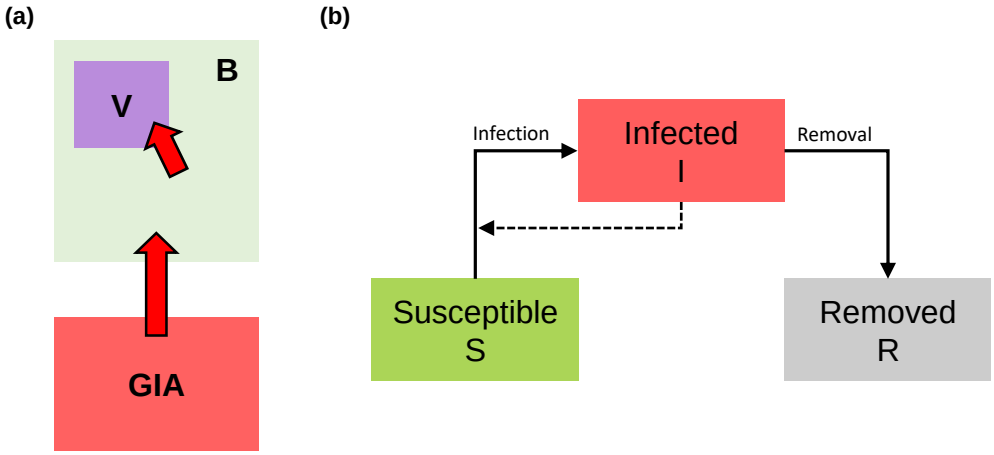

Figure 3.1: Schematic of the invasion model used in this chapter. **(a)** shows the spatial structure. A generally infested area (GIA) provides a source of inoculum, generating a constant force of infection on an uninvaded buffer region (B). The infection can then in turn invade the high value region (V). The goal of control is to protect this high value region. Within each region the epidemic follows SIR dynamics as shown in **(b)**, where hosts move from susceptible to infected and are removed by the disease or through control. The dashed line indicates that the number of infected hosts influences the infection rate.

Throughout, we scale time by the infectious period such that the disease induced removal rate μ is equal to one. The meanings and default values of all parameters and variables are given in Table 3.1.

3.2.2 Optimal control problem

The objective of control in this system is to protect the high value region from infection. Optimisation of the roguing strategy should minimise the impact of the pathogen in the high value region. We choose an objective of minimising the number of infected and removed hosts at the terminal time T. The population size is constant in each region, so this is equivalent to maximising the number of healthy and susceptible hosts that are retained in the high value region. We impose a restriction on the number of hosts that can be rogued per unit time to capture the economic and logistical limitations of disease management. The maximum expenditure rate, or number of hosts that can be rogued at any time across both patches, is given by M. This gives the following optimal control problem:

$$\min_{f_i(t)} \quad J = N_V - S_V(T) \quad i = B, V \tag{3.3a}$$

$$\text{subject to} \quad f_B(t)I_B(t) + f_V(t)I_V(t) \leq M \quad \forall t \tag{3.3b}$$

$$0 \leq f_i(t) \leq 1 \quad i = B, V \tag{3.3c}$$

where the state dynamics are given by Equations 3.2.

Table 3.1: Table of parameter and variable meanings and default values. Ranges are given where appropriate. The parameters were chosen to cover a range of epidemic sizes, with the default parameters giving a reasonably large epidemic under control as shown in Figure 3.2 on p. 48.

Symbol	Meaning	Default value [value range]
β	Infection rate	[0.001–0.004] host^{-1} t^{-1}
$\mathcal{F}$	External force of infection	0.0 t^{-1} [0.0–0.01] t^{-1}
ϵ	Coupling between regions	0.3
μ	Pathogen-induced death rate	1.0 t^{-1}
η	Roguing rate	0.2 t^{-1} [0.01–3] t^{-1}
N_B	Number of hosts in the buffer region	500
N_V	Number of hosts in the high value region	100
S_B	Number of susceptible hosts in the buffer region	$[0, N_B]$
S_V	Number of susceptible hosts in the high value region	$[0, N_V]$
I_B	Number of infected hosts in the buffer region	$[0, N_B]$
I_V	Number of infected hosts in the high value region	$[0, N_V]$
$f_B(t)$	Proportion of infected hosts in buffer region rogued at time t	$[0, 1]$
$f_V(t)$	Proportion of infected hosts in high value region rogued at time t	$[0, 1]$
M	Maximum expenditure rate	10 hosts [10^{-5}–50] hosts
J	Objective function	Equation 3.3a

In the absence of the budget constraint it might be expected that control would be maximal at all times, treating all infected hosts in both regions throughout the epidemic. We verified that this is indeed the case for the default parameters. We might therefore expect that, with the constraint, the optimal strategy would prioritise control in the more important region—the high value region—and allocate any remaining resources to the buffer region. We will test this hypothesis using the OCT framework introduced in the previous chapter, and use this simple system to compare the direct and indirect OCT frameworks. The following sections formulate the two versions of the optimal control problem.

3.2.3 Indirect formulation

In the indirect formulation, the Pontryagin maximum principle is used to set up an adjoint system and a Hamiltonian that is minimised along the optimal control path. The adjoint system ($\boldsymbol{\lambda}$) dynamics and Hamiltonian H are defined by the following equations:

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \mathbf{x}} \tag{3.4a}$$

$$H = \boldsymbol{\lambda} \cdot \dot{\mathbf{x}} \tag{3.4b}$$

where $\mathbf{x}$ is the state system, defined for this system in Equations 3.2. The terminal conditions of the adjoint system are defined from the payoff function $\Psi(T)$:

$$\boldsymbol{\lambda}(T) = \frac{\partial \Psi(T)}{\partial \mathbf{x}} \,. \tag{3.5}$$

For this system the payoff function is equal to the objective function J, and so all adjoints terminate at zero, apart from λ_{S_V} which has a final value of -1. This completes the two point boundary value problem, with the state fixed at the initial time, and the adjoint fixed at the final time, allowing us to use the FBSM algorithm introduced in the previous chapter (Algorithm 1, p. 34). As explained before, this solves the state equations forward in time, then the adjoint equations backwards in time, before updating the control function by minimising the Hamiltonian pointwise. This is repeated until the control converges.

Defining the state and adjoint as follows:

$$\mathbf{x} = \begin{pmatrix} S_B & I_B & S_V & I_V \end{pmatrix}^{\mathrm{T}} \tag{3.6a}$$

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_{S_B} & \lambda_{I_B} & \lambda_{S_V} & \lambda_{I_V} \end{pmatrix}^{\mathrm{T}} \tag{3.6b}$$

the Hamiltonian is then given by:

$$\begin{aligned} H = {} & (\lambda_{I_B} - \lambda_{S_B})\,(\mathcal{F} + \beta I_B + \beta\epsilon I_V)\, S_B - \lambda_{I_B} I_B\,(\mu + f_B\eta) \\ & + (\lambda_{I_V} - \lambda_{S_V})\,(\beta I_V + \beta\epsilon I_B)\, S_V - \lambda_{I_V} I_V\,(\mu + f_V\eta) \end{aligned} \tag{3.7a}$$

$$= g(\mathbf{x}, \boldsymbol{\lambda}) - \lambda_{I_B}\eta I_B f_B - \lambda_{I_V}\eta I_V f_V \,, \tag{3.7b}$$

where:

$$\begin{aligned} g(\mathbf{x}, \boldsymbol{\lambda}) = {} & (\lambda_{I_B} - \lambda_{S_B})\,(\mathcal{F} + \beta I_B + \beta\epsilon I_V)\, S_B - \lambda_{I_B} I_B \mu \\ & + (\lambda_{I_V} - \lambda_{S_V})\,(\beta I_V + \beta\epsilon I_B)\, S_V - \lambda_{I_V} I_V \mu \,. \end{aligned} \tag{3.7c}$$

The function $g(\mathbf{x}, \boldsymbol{\lambda})$ is therefore independent of the control, showing that the Hamiltonian is linear in each control function. In the control update step the control inputs are chosen so as to minimise the Hamiltonian at each time. Since the Hamiltonian is linear in the control this leads to bang-bang control, where control is either maximal or minimal. To

take into account the budget constraint, the Hamiltonian is minimised by solving a linear programming problem at each time point subject to the constraint. This specifies the new control at each time point which is then used in the next iteration of the FBSM algorithm.

The adjoint dynamics are derived from this Hamiltonian using Equation 3.4a. These are found to be:

$$\dot{\lambda}_{S_B} = (\lambda_{S_B} - \lambda_{I_B})(\mathcal{F} + \beta I_B + \beta\epsilon I_V) \tag{3.8a}$$

$$\dot{\lambda}_{I_B} = (\lambda_{S_B} - \lambda_{I_B})\beta S_B + (\lambda_{S_V} - \lambda_{I_V})\beta\epsilon S_V + \lambda_{I_B}(\mu + f_B\eta) \tag{3.8b}$$

$$\dot{\lambda}_{S_V} = (\lambda_{S_V} - \lambda_{I_V})(\beta I_V + \beta\epsilon I_B) \tag{3.8c}$$

$$\dot{\lambda}_{I_V} = (\lambda_{S_V} - \lambda_{I_V})\beta S_V + (\lambda_{S_B} - \lambda_{I_B})\beta\epsilon S_B + \lambda_{I_V}(\mu + f_V\eta)\ . \tag{3.8d}$$

This completes the necessary equations for using the FBSM algorithm.

Mixed constraint

In the above analysis the budget constraint was only used when minimising the Hamiltonian; we did not consider the constraint when constructing the adjoint equations. Mixed state/control constraints such as this can be difficult to handle in optimal control problems (Hartl et al., 1995). To verify that not including the constraint is valid, we here show that by introducing a penalty function for exceeding the maximum budget, the constrained problem reduces to the form given above, provided that the Hamiltonian minimisation is subject to the constraint. The approach taken here is similar to one used by Sage (1968, Chapter 4).

The budget constraint is given by h:

$$h(\mathbf{x}, \mathbf{f}, t) = f_B I_B + f_V I_V - M \leq 0\ . \tag{3.9}$$

Let us define a new state variable y with the following dynamics:

$$\dot{y} = h^2\theta(h) \tag{3.10a}$$

$$y(0) = 0 \tag{3.10b}$$

where $\theta(\cdot)$ is the Heaviside step function (with $\theta = 1$ for $h \geq 0$). This term is zero when the constraint is adhered to, and is positive when there is constraint violation. We then add to the objective function a penalty term based on this state variable:

$$J = (N_V(T) - S_V(T)) + \frac{1}{2}Cy(T)^2 \tag{3.11}$$

where C is a positive constant. The minimal value of the second term is zero, corresponding to no constraint violation. Defining a new adjoint variable λ_y for y, the Hamiltonian is now given by:

$$H = g(\mathbf{x}, \boldsymbol{\lambda}) - \lambda_{I_B} \eta I_B f_B - \lambda_{I_V} \eta I_V f_V + \lambda_y h^2 \theta(h) \, . \tag{3.12}$$

The adjoints for the susceptible classes are unchanged, but the infected class adjoints are affected. These, and the dynamics of λ_y are now given by:

$$\dot{\lambda}_y = 0 \tag{3.13a}$$

$$\dot{\lambda_{I_B}} = (\lambda_{S_B} - \lambda_{I_B}) \beta S_B + (\lambda_{S_V} - \lambda_{I_V}) \beta \epsilon S_V + \lambda_{I_B} (\mu + f_B \eta) - 2\lambda_y h \theta(h) f_B \tag{3.13b}$$

$$\dot{\lambda}_{I_V} = (\lambda_{S_V} - \lambda_{I_V}) \beta S_V + (\lambda_{S_B} - \lambda_{I_B}) \beta \epsilon S_B + \lambda_{I_V} (\mu + f_V \eta) - 2\lambda_y h \theta(h) f_V \, . \tag{3.13c}$$

Whilst this system could be solved without imposing the constraint when minimising the Hamiltonian, all the magenta terms in the above equations are zero when the chosen controls adhere to the budget constraint. This means that when imposing this constraint on Hamiltonian minimisation, the adjoint equations previously used are correct without considering the constraint. The simpler solution is therefore used to find the optimal control satisfying the constraint.

3.2.4 Direct formulation

The setup for the direct formulation is considerably simpler than for the indirect case, since no derivation of the Hamiltonian or adjoint system is required. The handling of constraints is also much simpler. We use the BOCOP package (v.2.0.5) to generate and solve the NLP problem (Team Commands, Inria Saclay, 2017). The state dynamics (Equations 3.2) are coded in C++, and the package discretises the system using a fourth order Runge-Kutta method (direct transcription approach, as described in the previous chapter). The NLP problem is solved using the software Ipopt (Wächter and Biegler, 2006), which implements an interior point optimisation method.

3.3 Results

3.3.1 Optimal strategies

For most parameter sets, the optimal strategy found using either the direct or indirect approach prioritises control in the high value region (Figure 3.2). The maximum control is allocated to the high value region, up to the maximum expenditure rate M, and any remaining resources are allocated to the buffer region. For some parameter sets however, the optimal strategy is a switching strategy (Figure 3.3). Early in the epidemic it is optimal to prioritise the buffer region, but it becomes optimal to switch back to prioritising the

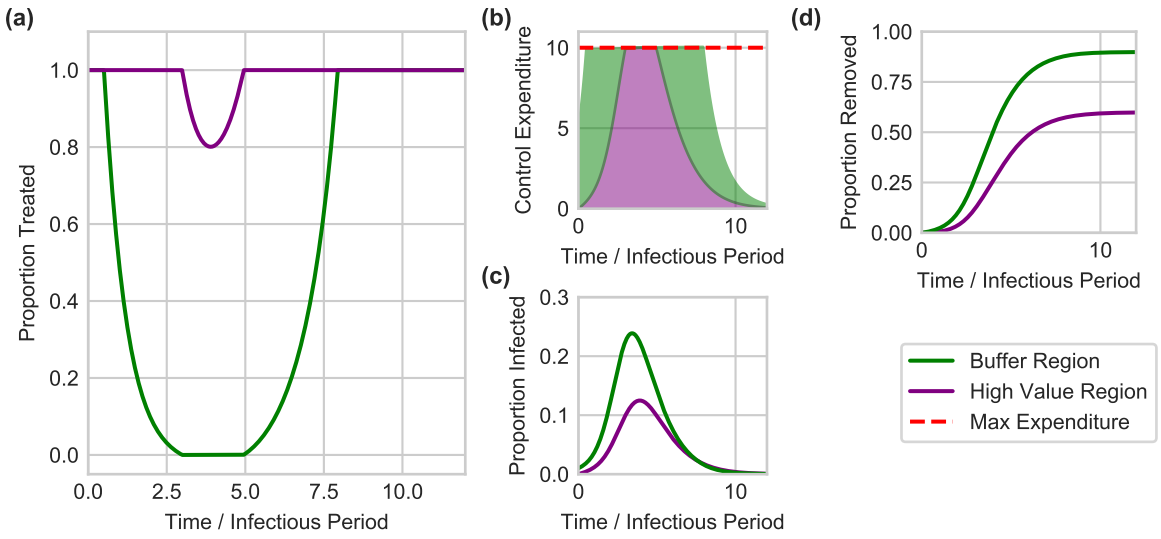

Figure 3.2: The general strategy that is found in the majority of cases prioritises control in the high value region. Here the default parameters are used, with 5 hosts initially infected in the buffer region and an infection rate of 0.005 $\text{host}^{-1}\,\text{t}^{-1}$. **(a)** shows the proportion of hosts being treated in each region over time, with the number of hosts being treated shown in **(b)**. The strategy allocates as many resources as possible to the high value region, before spending the remainder in the buffer region. After around 3 time units the number of infected hosts in the high value region exceeds the maximum expenditure M, and not all can be treated. **(c)** shows the disease progress curve and **(d)** the proportion removed for each region.

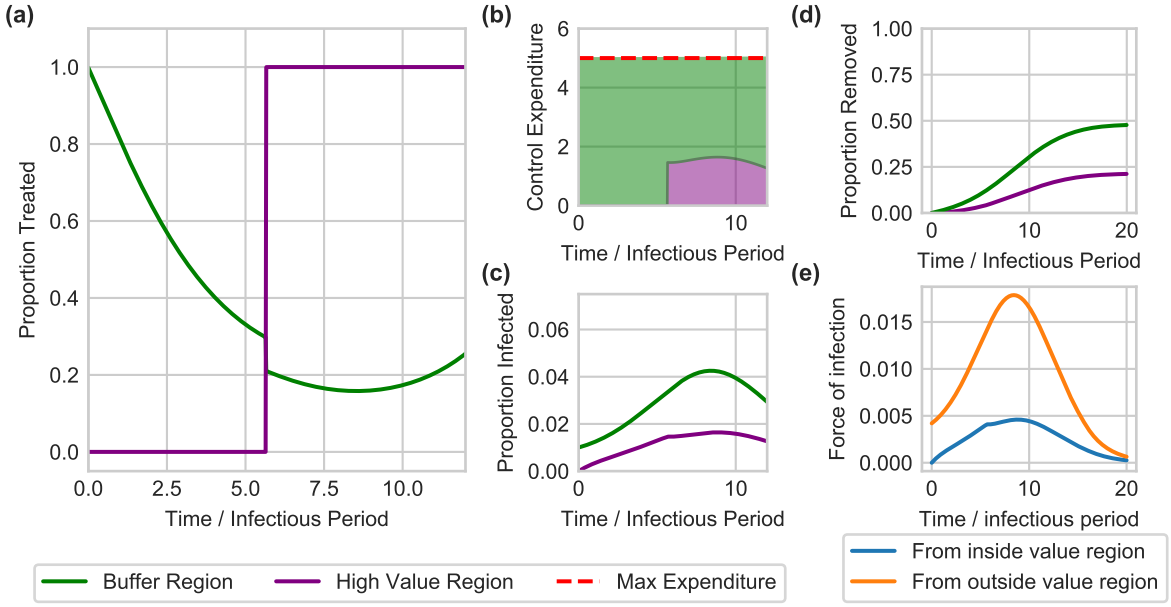

Figure 3.3: An alternative switching strategy is found for some parameters. The default parameters are used here with 5 hosts initially infected in the buffer region, an infection rate of 0.0028 $\text{host}^{-1}\,\text{t}^{-1}$, and a maximum expenditure rate of 5 hosts. **(a)** and **(b)** show the proportion and number treated over time in each region. For approximately the first 6 units of time, the buffer region is prioritised, before switching to prioritising the high value region. The disease progress curves and proportion removed are shown in **(c)** and **(d)**. **(e)** shows the force of infection in the high value region from inside and outside the same region under the optimal control strategy. The switch time is not simply when the force of infection from within becomes greater than that from outside the high value region.

high value region later. This strategy is a bang-bang control since the maximum budget is allocated to the buffer region, before switching to the maximum budget being allocated to the value region. However, the control expenditure and proportion of hosts treated vary smoothly to maximise this budget. For the ranges of parameters we considered here, we do not see strategies where priorities switch multiple times when budgets are limiting.

A naive expectation might be that the strategy switch is caused by the region generating the larger force of infection on the high value region, i.e. the region dominating the infectious pressure is prioritised for control. However, the driver for switching priorities is not simply the force of infection on the high value region. As shown in Figure 3.3(d), the force of infection from within the high value region can be smaller than that from the buffer region at all times, but a switch is still optimal. The switching strategy is found for intermediate values of the infection rate (Figure 3.4). Low values of the infection rate are easy to manage, and so can be controlled simply by treating in the high value region. High infection rates give epidemics that spread rapidly, and so the more important value region must be prioritised to keep the epidemic under control there. At intermediate levels though, the disease spreads slowly enough to allow reduction of infectious pressure by treating in the buffer region, but quickly enough that this reduction of pressure from the buffer is necessary.

3.3.2 Formulation comparison

The switch time, if there is one, can easily be extracted from an optimised solution, and so can be used to compare the direct and indirect approaches. Figure 3.4(a) shows that in general the two methods agree, but for some parameters the indirect FBSM does not find an optimal switching strategy where the direct method does. As shown in Figure 3.4(b), the direct method is more accurate as the switching strategy is optimal. Because of the difficulties with convergence using the indirect approach, the FBSM can in practice converge on suboptimal solutions more frequently than the direct method. Note here that the differences in epidemic cost are small for this simple system, but do show numerical differences in the control strategies between the methods. In reality here the additional costs associated with strategy switching would likely outweigh the benefits in epidemic cost.

Other convergence issues are also seen with the FBSM, causing significantly slower convergence than using the direct method. For example, the optimisation underlying the strategy in Figure 3.2 takes 11.7 s using the direct approach, whereas for the same time resolution the indirect approach takes 98 s. Both measurements were made using a 4.0 GHz Intel Core i7-4790K with 32 GB of RAM. In many cases the FBSM also fails to converge to a solution.

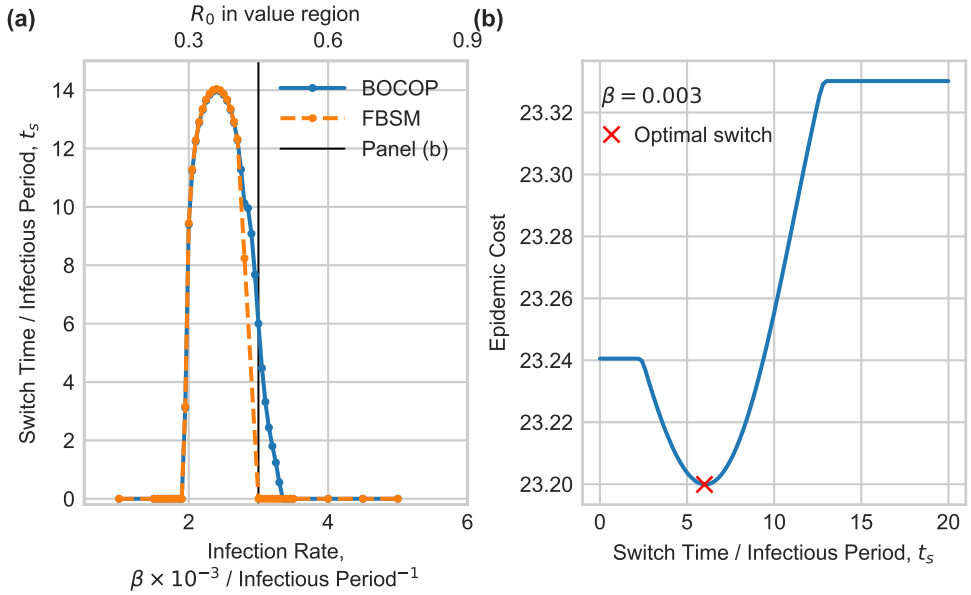

Figure 3.4: Using the optimal switch time to compare the direct and indirect approaches. The default parameters are used with 5 hosts initially infected in the buffer region. **(a)** shows the switch time found by the direct (BOCOP) and indirect (FBSM) methods. Switching strategies are found at intermediate infection rates. The values of R_0 in the value region corresponding to the values of β are also shown. Whilst R_0 is here always below 1, transmission will be increased because of spread in the buffer region. In most cases the switch times found using the two methods agree, although in some cases the direct method finds an optimal switch where the indirect method does not. **(b)** shows a scan over the switch time for one of these cases. The optimal switch time is at 6.0 time units, with an associated epidemic cost of 23.20. For this value of β, the FBSM does not find a switch whereas BOCOP identifies the correct switch time of 6.0 units.

3.3.3 Parameter sensitivity

In this section we test the sensitivity of the switching strategies to parameterisation of the model. The effects of the maximum expenditure rate M and the control rate η are shown in Figure 3.5. In these cases, as control becomes more effective either through an increased budget or faster treatment, the intermediate range over which a switching strategy is optimal shifts to higher values of the infection rate. The control is more effective and so the epidemic is easier to control, shifting the switching strategies to faster spreading epidemics with higher infection rates.

Figure 3.6 shows the response of the switch time from the direct optimisation to the infection rate, the number of infected hosts initially in the buffer region, and the external force of infection. At low levels of initial infection and external force of infection, the switch strategy is optimal at intermediate infection rates, as found previously. As the rate of invasion in the buffer is increased, through more initial infection or a higher external force of infection, the range of infection rates over which a switching strategy is optimal decreases. For high rates of invasion in the buffer region a switching strategy is not optimal because

the disease spreads faster in the buffer, and so control is less effective there than in the high value region. Note that for some values of the external force of infection, optimal strategies with additional switches are found.

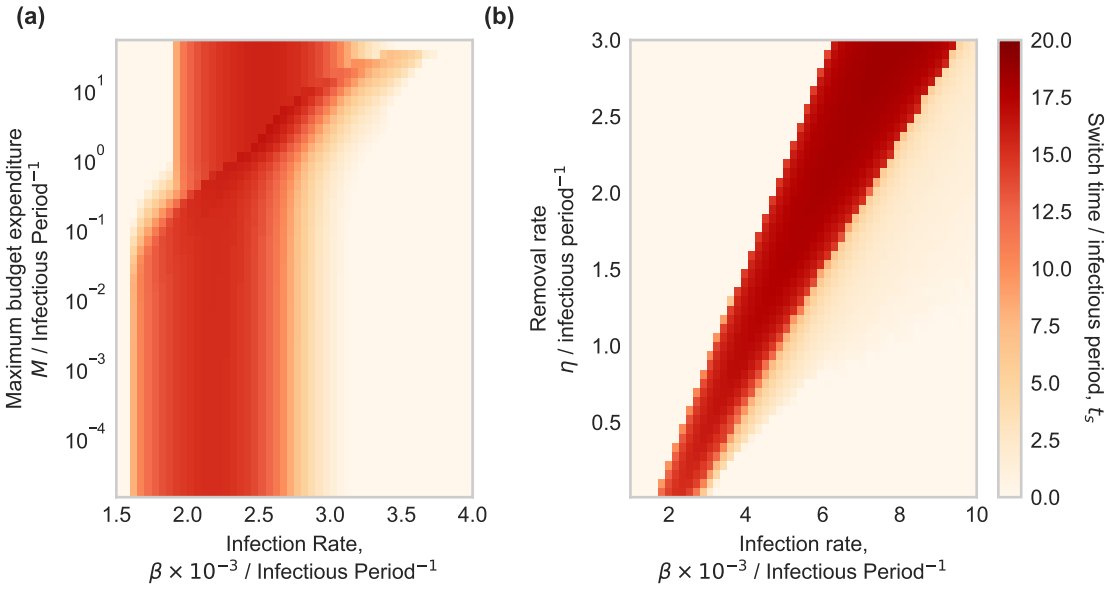

Figure 3.5: Effect of the control budget and control rate on the optimal switch time, using default parameters and one initially infected host in the buffer region. **(a)** and **(b)** show the effect of infection rate combined with the maximum budget and removal rate respectively. As control becomes more effective the range of infection rates over which a switch is optimal is shifted to higher values. The shape in **(a)** is caused by a change in regime between parameters where maximum control is possible and where the maximum expenditure rate is limiting.

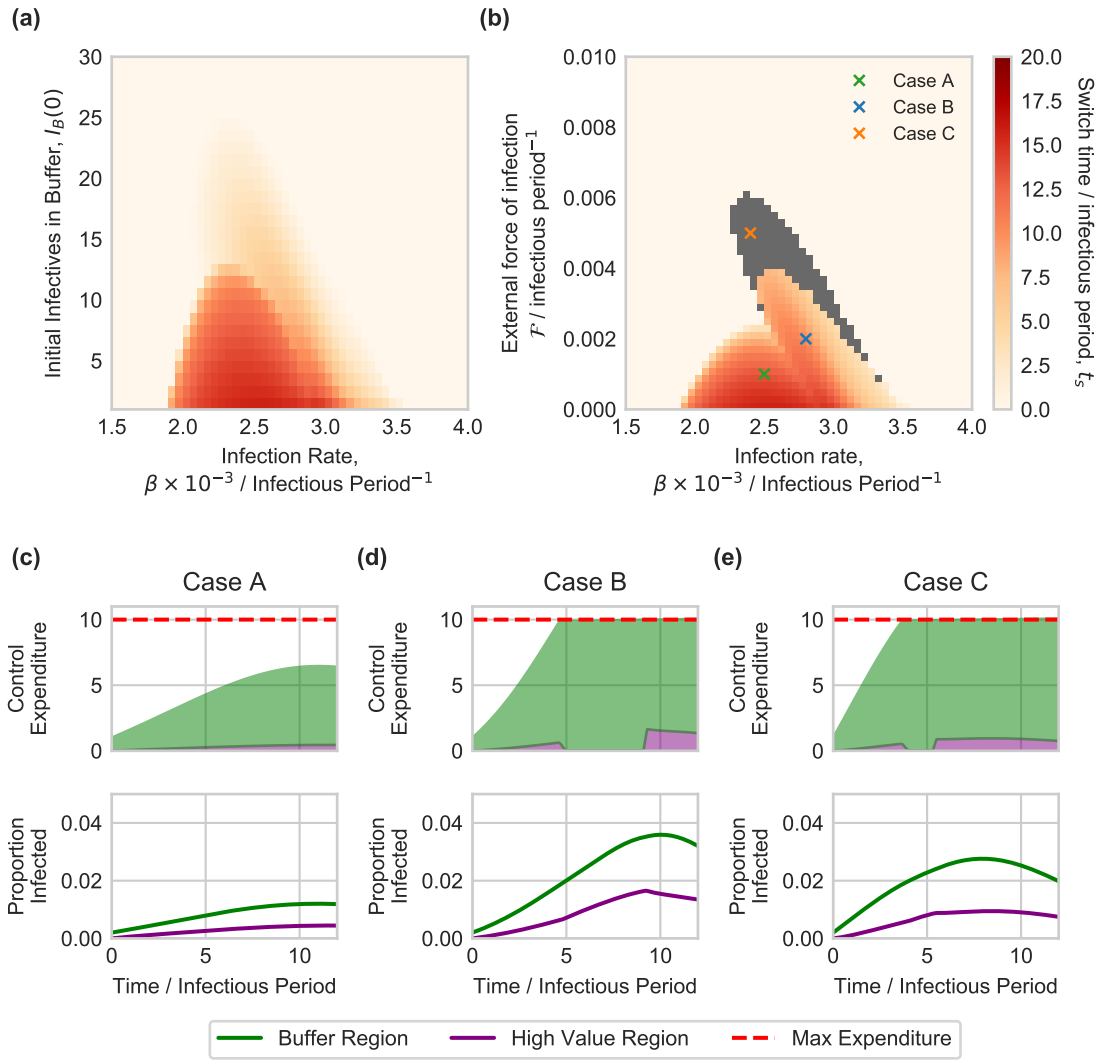

Figure 3.6: Sensitivity of the optimal strategy to the rate of infection spread, using the direct method and default parameters. **(a)** shows the switch time as a function of the infection rate and the initial number of infected hosts in the buffer region. **(b)** shows the effect of the infection rate and the external force of infection on the buffer region, with a single initially infected host in the buffer region. The grey region shows cases where multiple switches are identified. For both **(a)** and **(b)**, as the epidemic invades faster in the buffer region (moving up the y axes) the switching strategy becomes less beneficial. The shape of the response is affected by where the maximum expenditure rate becomes limiting, with the budget becoming more limiting towards the upper right corners. A diagonal line separates regimes where maximum control is possible from cases where the maximum expenditure rate is limiting. Cases A, B and C are highlighted in **(b)**, and control expenditure and disease progress curves are shown for each of these in **(c)**, **(d)** and **(e)**. Case A finds a switch but as the budget is not limiting, it has no effect. In case B the budget is limiting so the single switch at approximately $t = 8$ has an effect. In case C there is an additional switch early in the epidemic, but since the budget is not limiting at that point it has no effect. Only the switch from buffer to value prioritisation at around $t = 5$ has an effect.

3.3.4 Testing robustness

Finally, we test how these optimal strategies perform when the underlying model is not known accurately. We introduce a systematic error into the infection rate of the model used to optimise the control strategy. The resulting optimal control specifies an expenditure over time ($f_i(t)I_i(t)$ for each region) which is applied to a model with the correct infection rate. The results are shown in Figure 3.7. Larger errors in the infection rate lead to worse control of the epidemic. For large underestimates of the infection rate, the optimised control is worse than simply allocating the full budget to the high value region, with no treatment in the buffer region. The optimised control strategies lead to wasted resources that could be better allocated to the other region (Figure 3.7(c)).

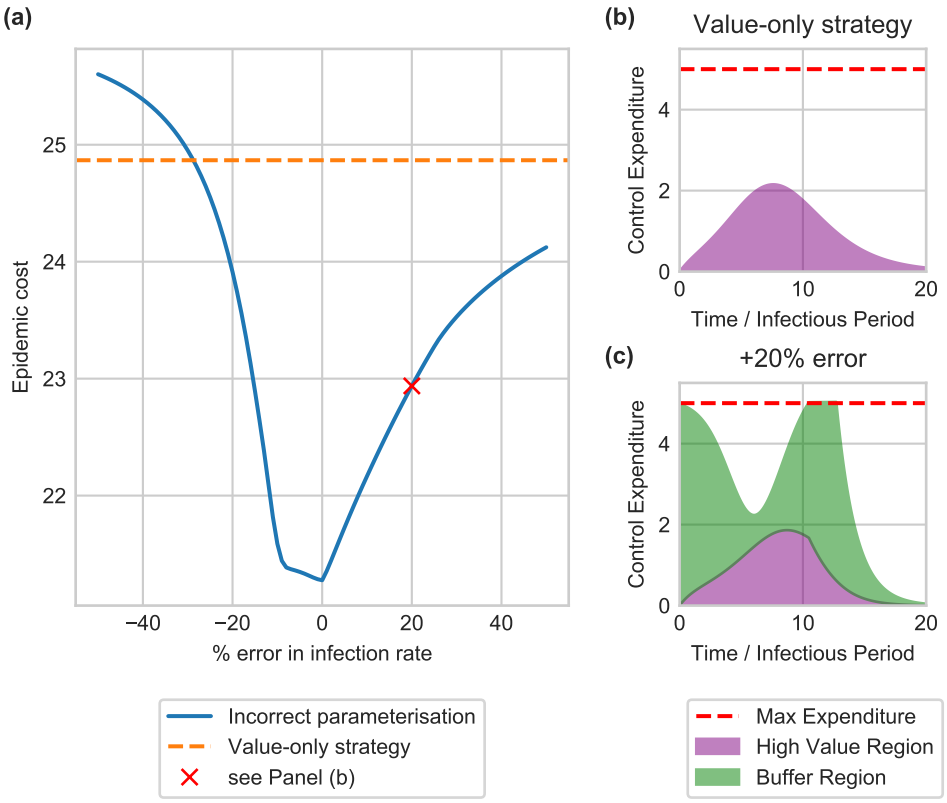

Figure 3.7: Effect of incorrect parameterisation on the optimal control performance. Control is optimised using an incorrect infection rate, and the budget allocations applied to the correct model. The baseline parameters use the default values, with an infection rate of $0.0028\,\text{host}^{-1}\,\text{t}^{-1}$ and a maximum expenditure rate of 5 hosts, the same values as used for the switching strategy in Figure 3.3. **(a)** shows the epidemic cost as the percentage error in the infection rate is varied. Larger errors lead to worse performance, and can be worse than the simple strategy of full allocation to the high value region (shown in **(b)**). **(c)** shows the wasted control resources using an infection rate overestimated by 20 %. Here too much control is allocated to the high value region that cannot be spent, and so is wasted. Similarly, not enough resources are allocated when the infection rate is underestimated, also leading to wasted control resources.

3.4 Discussion

In this chapter we have shown how OCT can be used to find time-dependent control strategies. We have shown how OCT identifies less intuitive strategies, for example the switching strategy here. For the simple model considered here these optimal strategies could have been identified by an exhaustive scan over switch times, but for more complex models with more complex switching strategies this will not generally be true. Whilst the setting in this chapter was highly simplified, we can already see some of the potential limitations of these more complex strategies for practical use. Although not included in our analysis, there are additional costs associated with changing policy during an epidemic, i.e. switching region priorities in this chapter, and these costs must be balanced with the benefit to epidemic control. These costs could be included in the OCT objective function, for example by adding a term penalising rapidly changing controls.

We showed how the optimal control strategy changed as various model parameters were adjusted, in particular investigating the sensitivity of the control switching time to the infection rate, control efficacy (through the maximum budget and removal rate parameters), and the rate of invasion in the buffer region (through the initial level of infection and the external force of infection). This showed how more effective control results in switching strategies being optimal for faster spreading epidemics, and if the epidemic spreads too quickly in the buffer then all resources must be focussed on the value region. We did not test how the connectivity between the two regions (ϵ) impacts the control but the results are likely to follow the same pattern. With increased connectivity the disease will spread faster into the value region, and so resources should switch to prioritising the value region faster. With decreased connectivity control can focus for longer on reducing inoculum from the buffer region. Similar to the dependence on infection rate, there is likely to be an intermediate range of connectivity values where switching strategies are optimal.

We have also shown how incorrect parameterisation can lead to less effective disease management, and in such cases it may even be better to use the simplest possible strategy rather than use OCT at all. Consideration of these uncertainties when determining the optimal strategy is important (Epanchin-Niell and Hastings, 2010). A study by Carrasco et al. (2009) showed that when controlling invasive species, the optimal strategy when parameters are known precisely is not always optimal when parameter uncertainty is introduced. In these cases it is important to find a robust strategy that can handle any uncertainty present.

The strategies we have identified are similar to those found in other plant disease studies, despite our unusual objective of protecting a single region. For example, Ndeffo Mbah and Gilligan (2011) found that disease management across sub-populations is most efficient when a switching strategy is employed, first allocating resources to the more infected group and

later switching to treat the less infected group. A similar strategy is found by Ndeffo Mbah and Gilligan (2010b) for management of sudden oak death. Similar switching strategies have also been found more broadly for: management of invasive species (Carrasco et al., 2009), distribution of prophylactic vaccines in a spatially structured population (Keeling and Shattock, 2012), and switching between immunisation and palliative care during an epidemic when resources are limited (Klepac et al., 2012).

The type of control strategy identified here, a bang-bang switching control, can be sensitive to precise knowledge about the optimal switch time. Forster and Gilligan (2007) showed that when applying control optimised using a mean-field model of an epidemic to a spatially-explicit model, errors in the switch time can lead to performance that is worse than a simple constant strategy. Similarly, we showed here that incorrect parameterisation can lead to ineffective disease management. The accuracy of the underlying model used for optimisation is limiting the practical applicability of the resulting strategies.

In this chapter we also compared the direct and indirect formulations for solving the optimal control problem. We found that the direct method in some cases finds superior switching strategies, and is also more robust and reliable than the indirect method. This means that the direct method is more likely to converge, whereas convergence using the indirect approach can be highly sensitive to parameters. The direct method is also simpler to implement, particularly when considering constraints that mix state and control variables. Other studies have also found that the direct method is more robust and reliable (Betts, 2010), but results from the indirect method can be more accurate when they do converge (Rao, 2009). In this chapter the accuracy of the direct and indirect approaches has been comparable when both methods converge, but if additional accuracy is required the results from the direct method could be used to initialise an indirect method (Von Stryk and Bulirsch, 1992). The purpose of this thesis is to find robust and practical management strategies, and so the direct approach is most appropriate. Overall, we will therefore use the direct method going forward in this thesis.

3.5 Conclusions

We have shown that switching strategies can be important when considering management strategies to protect a specific high value region. Direct approaches to numerically optimising the strategy perform better than indirect approaches, giving faster and more reliable convergence and more accurate results. However, the optimal strategies do not always perform well when there is inaccuracy in the parameters underlying the model. New frameworks are needed to make use of the OCT results when accounting for uncertainties and complexities in the real world.

4 Applying optimal control theory to complex models

4.1 Introduction

In the previous chapters we have introduced OCT and shown how it can be used to find the best disease management strategies to protect a high value region. However, because of the mathematical complexity of finding these optimal strategies, major simplifications to the system as modelled are required to allow progress to be made using OCT. There is no standard for putting the results from these mathematically motivated simplifications into practice, nor for assessing the impact of the model simplifications on predictive accuracy. As a result, it is often unclear how these strategies would perform if adopted by policy makers. In this chapter we will investigate frameworks for translating OCT results into practical and realistic management policies[1].

Robust policy decisions clearly require accurate predictions of future disease dynamics. Increasingly, complex simulation models incorporating detailed representations of disease transmission processes are used to assess the potential impact of a given intervention strategy (Lofgren et al., 2014). To ensure accurate epidemic predictions, simulation models designed to aid decision making must often capture highly complex dynamics (Savary and Willocquet, 2014). As we discussed in Chapter 1, this often makes optimisation of control strategies infeasible—particularly when control measures can vary over time, in space or according to disease risk—because of Bellman's 'curse of dimensionality' (Bellman, 1957). For most simulation models the only viable option is then to evaluate a small subset of 'user-defined' plausible strategies that remain fixed during the epidemic, potentially scanning over a single parameter such as a culling radius. We refer to this approach as 'strategy testing'. Using this approach makes it difficult to have high confidence in the best-performing strategy, since with no framework for choosing it, the set of strategies under test is very unlikely to span the entire space of control options. This makes it unlikely that the true optimum will be found.

[1] This chapter is based on work published in Bussell et al. (2019). All code is available at `https://github.com/ehbussell/Bussell2018Model`

Simulation models accurately capture the dynamics of the real system but are often impossible to optimise. In this chapter we ask whether and how the optimisation capabilities of OCT might be combined with the accurate predictions made by simulation models, to give improved management strategies. What framework should be applied to make practical use of OCT? This chapter will set the foundations for the methodology to be applied to real-world systems in the following chapters. In Section 4.2 we describe two methods from control systems engineering for applying OCT results to simulation models: open-loop and model predictive control. A network epidemic model is described in Section 4.3 which allows us to showcase these frameworks, and we illustrate the potential benefit of using OCT alongside simulation models in Section 4.4. We seek to answer how, under current computational constraints, results from OCT can be applied whilst maintaining the realism required for practical application.

4.2 Frameworks for practical optimal control

In Chapter 2 we reviewed the history and use of OCT in epidemiology, but OCT has been applied widely beyond this, particularly in engineering and economics (Bertsekas, 2001). In mathematical biology, OCT has been used to optimise control of vector-borne diseases (Blayneh et al., 2009) and sudden oak death (Ndeffo Mbah and Gilligan, 2010a) for example, and more recently to control complex, agent-based models (ABMs; An et al., 2017), a type of model that simulates the individual behaviour of autonomous agents. An et al. (2017) suggest the use of a model that approximates the dynamics of the ABM, designed to be simple enough to allow mathematical analysis of the optimal control. A suitable approximate model is chosen and fitted either to real data, or to synthetic data from the ABM. The OCT results from the approximating model are then mapped onto the ABM to be tested in a process referred to as 'lifting'. This process of applying OCT to an approximate model and lifting the results to a complex ABM could also apply to the optimisation of detailed epidemic simulation models. The topic of this chapter is essentially to test the performance of this process. We now describe two possible frameworks from control systems engineering based on this control lifting approach.

4.2.1 Open-loop control

The first method is the simplest application of control lifting, and the framework implicitly suggested by An et al. (2017). Control is optimised on the approximate model once using the initial conditions of the simulation model. The resulting optimal control strategy is lifted to the simulator and applied for the full simulation run time (Figure 4.1). Other potential control strategies can be assessed against the OCT strategy by applying all strategies to the simulation model, possibly repeatedly to account for stochasticity in the simulation model.

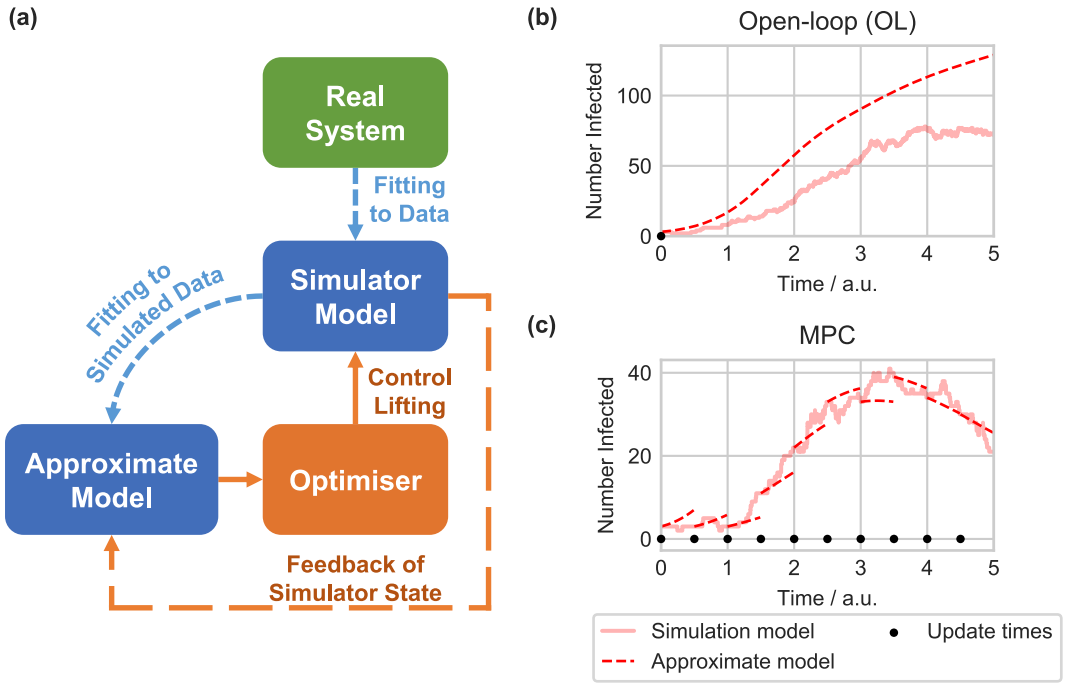

Figure 4.1: Open-loop and model predictive control (MPC) frameworks. In **(a)** the model hierarchy is shown, with optimised controls from the approximate model directly lifted to the simulation model. The real system is in green, the models and fitting processes are in blue, and the control framework is in orange. Without the orange dashed feedback loop, this is open-loop control. MPC resets the state of the approximate model at regular update times, before re-optimising and lifting controls to the simulation model until the next update time. **(b)** shows how this works for open-loop control. The simulation model and approximate model match at the initial time, and control is based on the predicted spread. The predicted spread deviates from the simulation over time. In MPC, as shown in **(c)**, at regular update times the approximate model is reset to match the simulation, ensuring the deviation remains small.

The optimisation gives a single, time dependent strategy for all simulation realisations, and so does not incorporate any feedback. It is therefore referred to as 'open-loop' control, as it is fully specified by the simulation initial conditions and the trajectory predicted by the approximate model. Use in epidemiology is uncommon, although Clarke et al. (2013) use OCT in an approximate model to find optimal levels of Chlamydia screening and contact tracing which are then mapped onto a network simulation.

4.2.2 Model predictive control

Open-loop control requires the approximate model to remain accurate over the timescale of the entire epidemic. However, for tractability the approximate model must necessarily omit many heterogeneities present in the simulation model, such as spatial effects and risk structure. When strategies resulting from OCT are then applied to the simulation model or to the real system, the disease progress curve is likely to deviate systematically from the trajectory predicted by the approximate model. Model predictive control (MPC) is an

optimisation technique incorporating system feedback that can take such systematic errors in the approximate model into account (Camacho and Bordons, 1995; Lee, 2011).

With MPC, both the approximate and simulation models are run concurrently. Optimal strategies are still lifted from the approximate model to the simulation, but at regular update times the values of the state variables in the approximate model are reset to match those in the simulation. The control is then re-optimised using the new initial conditions in the approximate model, and the new control strategy is applied to the simulation until the next update time. Running the approximate and simulation models concurrently, with multiple time-limited optimisations per realisation, ensures that the approximate model and control strategy closely match each individual simulation realisation (Figure 4.1(c)). These multiple optimisations are computationally costly but remain tractable, unlike performing optimisation on the full simulation model.

MPC has had some use within the epidemiological literature, the majority being for control of drug applications for single individuals rather than control of epidemics at the population level. Examples include finding management strategies for HIV that are robust to measurement noise and modelling errors (Zurakowski and Teel, 2006; David et al., 2011), and control of insulin delivery in patients with diabetes (Hovorka et al., 2004). These studies highlight the benefits of MPC for robust control, i.e. control that remains effective despite system perturbations. However, only one study concentrates on epidemic management (Sélley et al., 2015). In that study, Sélley et al. develop a pairwise ODE model as an approximation of an epidemic on a network. Time-dependent control is optimised on this model, and applied using an MPC framework. Crucially though, the control is applied to the approximate model rather than to an individual-based network simulation. This means the practical results of the strategy are not tested. In the bioinvasion literature, Simas et al. (2019) apply MPC to the control of invasive golden mussels in rivers in Brazil. The underlying model used by Simas et al. (2019) is a fitted polynomial that tracks the numbers of 10 species of fish, and the effect of controlling the golden mussel on the fish population. Again, the control is not applied to a more complex model to test how the strategy would perform in practice.

4.3 Controlling an epidemic spreading on a network

As an illustrative example which demonstrates open-loop and MPC for epidemic management, we use a stochastic SIR network model including host demography and risk structure, and vaccination as a control method (Figure 4.2). The model is deliberately kept simple to show how the underpinning idea is broadly applicable across plant, animal and human diseases. Whilst the model and its parameters are arbitrary and do not represent a specific

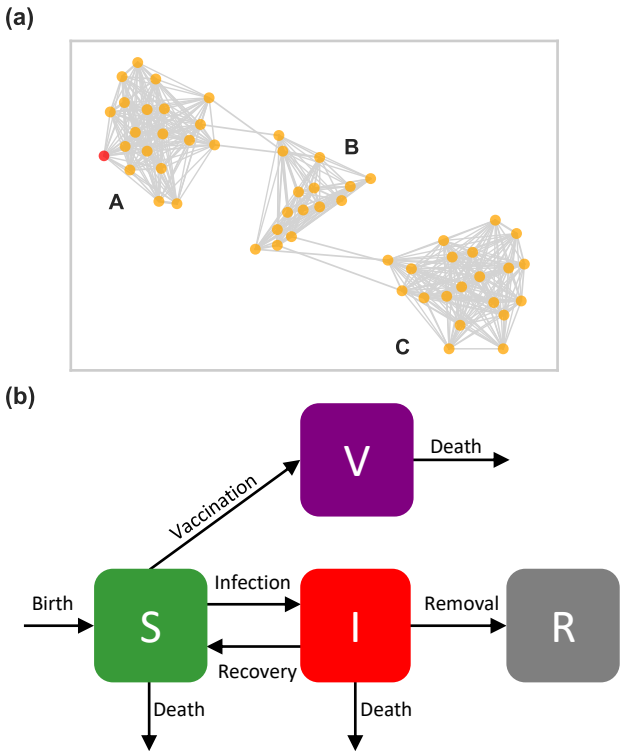

Figure 4.2: **(a)** The network used for the simulation model, including region labels. The epidemic is seeded in the red node in region A, and can spread between connected nodes (grey lines). **(b)** Diagram of host states and transitions. For transition rates see Table 4.1 on p. 64. Hosts are classed as susceptible (S), infected (I), vaccinated (V), or removed (R).

disease, we use it to represent a scenario in which a simulation model has already been fitted to a real disease system; the network model is therefore used here as a proxy for a potentially very detailed simulation model. The model itself includes features abstracted from metapopulation or network epidemic models as routinely used for plant, animal and human diseases (e.g. Rowthorn et al., 2009; Keeling et al., 2001; Keeling and Gilligan, 2000; Margosian et al., 2009; Bansal et al., 2007).

Both the open-loop and MPC frameworks require an approximate model with which optimisation using OCT is possible. The size of the system has a significant impact on the tractability of solving the optimal control problem, but complex dynamics can also make finding numerical solutions difficult and there is no clear way to know a priori whether a solution will be possible. In this chapter we will use two different levels of approximation to assess how this approximate model should be chosen. The resulting strategies using both open-loop and MPC will also be compared with a strategy testing approach as a baseline, in which a limited number of plausible interventions that do not vary during the epidemic are tested on the simulation model.

4.3.1 Simulation model

In the simulation model, infection spreads stochastically across a network of nodes. Nodes are positioned randomly in three separate regions, with 20 nodes in regions A and C, and 15 nodes in region B (Figure 4.2(a)). The nodes are positioned such that no two nodes are closer than a predetermined threshold, here 0.2 spatial units. This threshold avoids very high connectivity between any pair of nodes. The connectivity between nodes within a region is calculated using an exponential kernel with a scale parameter α of 0.2 units ($\sigma_{ij} = \mathrm{e}^{-d_{ij}/\alpha}$) where d_{ij} is the distance between nodes i and j. Between regions, three connections are made between pairs of nodes in regions A and B, and B and C, with a coupling value of 0.1β for each. This is approximately equal to the average coupling between pairs of nodes within either of the two larger regions, A and C. There are no direct connections between regions A and C. Note that this network is arbitrary, and is used simply as an illustrative scenario.

Each node initially contains a total of 30 hosts, stratified into high and low risk groups with different infection rates. High risk hosts are both more susceptible to the disease, as well as more infectious once infected. On average 10% of hosts within each node start in the high risk group. To give heterogeneity across the network the exact number of high risk hosts for each node is chosen using a binomial trial. Each individual host can be in one of three active states: susceptible (S), infected (I) or vaccinated (V), or removed (R) by the disease (Figure 4.2(b)). The infection can spread between individuals within nodes and between connected nodes. The net rate of infection of risk group r in node i is given by:

$$S_i^r \sum_j \beta \sigma_{ij} \left(\rho^{rH} I_j^H + \rho^{rL} I_j^L \right) , \tag{4.1}$$

where S and I are numbers of susceptible and infected hosts respectively, and β is the transmission rate. Throughout superscripts refer to the risk group (high: H, low: L), and subscripts identify the node. For example, S_3^{H} represents the total number of high risk susceptibles in node 3. In addition, we use N to refer to the total number of active hosts that could be controlled, such that N_3^{H} is the number of susceptible, infected and vaccinated high risk hosts in node 3. The sum in Equation 4.1 is over all connected nodes including the focal node itself, with the relative transmission rate into node i from node j given by σ_{ij}, and risk structure given by the 2×2 matrix ρ, where $\rho = \begin{pmatrix} \rho^{\mathrm{HH}} & \rho^{\mathrm{HL}} \\ \rho^{\mathrm{LH}} & \rho^{\mathrm{LL}} \end{pmatrix}$.

Mass vaccination is the only intervention we consider, with the potential to target based on both risk group and region, but randomised across host infection status (e.g. the vaccine is given to all hosts but is only effective on susceptibles). This could represent, for example, vaccine-laden baits distributed to wild animals, or chemicals sprayed onto plants. For simplicity in the example model considered here, vaccinated hosts are not susceptible

to infection, i.e. the vaccine is totally effective, and the vaccine does not wane over the timescales considered. Logistical and economic constraints are included through a maximum total vaccination rate ($\eta_{\max}$) that can be divided between risk groups and regions. Since only susceptible hosts are affected by the vaccine, within each group susceptibles are vaccinated at rate $f\eta_{\max}S/N$, where f is the proportion of control allocated to that group, and N is the total group population.

Trajectories are simulated using the Gillespie direct method (Gillespie, 1977). The possible events are host birth and death, infection, vaccination, and removal and recovery of infected hosts (Figure 4.2(b)). These events and the associated rates are given in Table 4.1. Parameters, as specified in Table 4.2, were chosen to give a large epidemic under no control intervention, spreading across all three regions. Typical simulation trajectories with no control are shown in Figure 4.3.

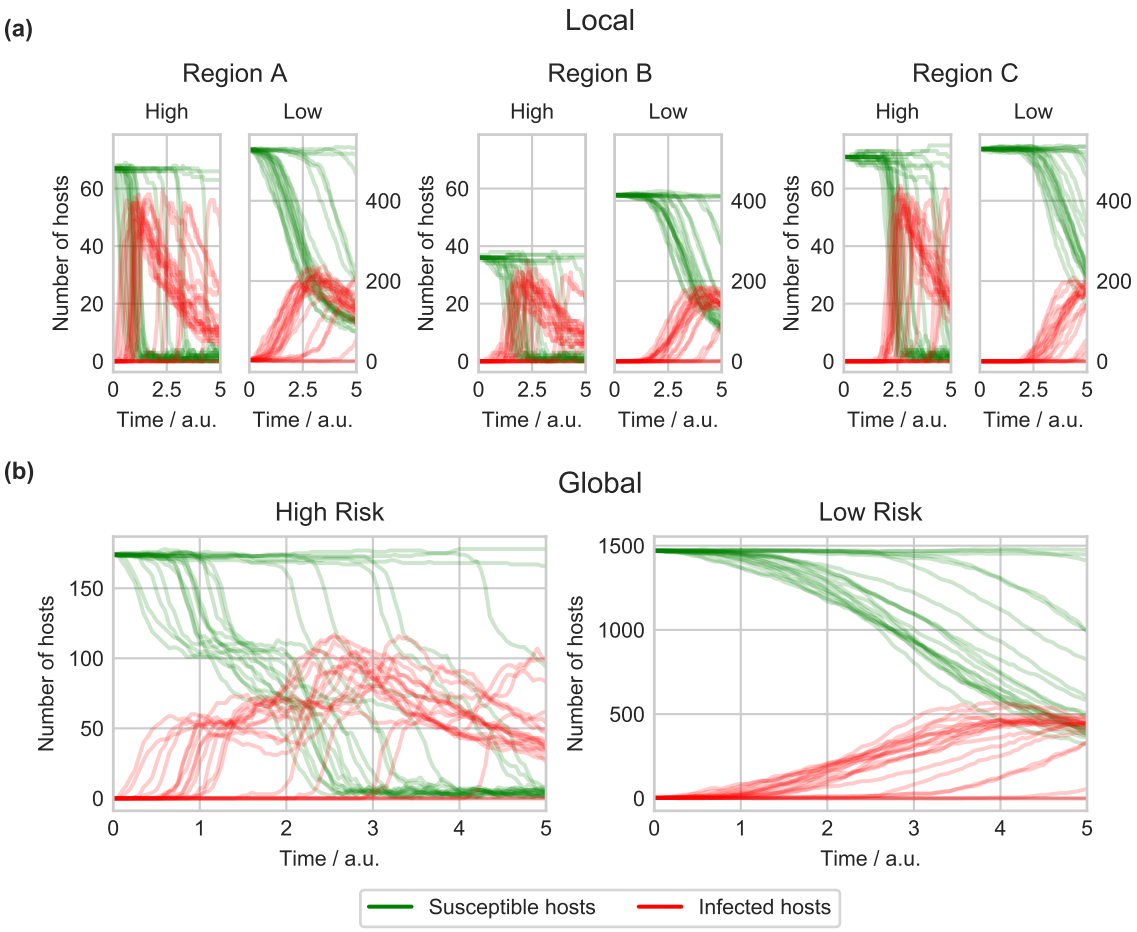

Figure 4.3: Twenty typical simulation model trajectories with no control. Host dynamics are shown for each risk group (high and low) and for each region in **(a)**. In **(b)** the dynamics are shown across the whole landscape for each risk group. Green and red lines represent susceptible and infected hosts respectively.

Table 4.1: Possible events in the simulation model with associated rates.

Event	State Change	Rate
Birth	$\emptyset \to S_i^r$	bN_i^r
Death	$\{S_i^r, I_i^r, V_i^r\} \to \emptyset$	dN_i^r
Infection	$S_i^r \to I_i^r$	$S_i^r \sum_j \sum_{r'} \beta \sigma_{ij} \rho^{rr'} I_j^{r'}$
Vaccination	$S_i^r \to V_i^r$	$f_i^r \eta_{\max} S_i^r / N_i^r$
Removal	$I_i^r \to \emptyset$	μI_i^r
Recovery	$I_i^r \to S_i^r$	νI_i^r

Table 4.2: Parameter values used for simulation model trajectories. Note that the units of time are arbitrary.

Meaning	Parameter	Default Value
Birth Rate	b	$0.01\,\mathrm{t}^{-1}$
Death Rate	d	$0.01\,\mathrm{t}^{-1}$
Maximum Vaccination Rate	$\eta_{\max}$	$200\,\text{hosts}\,\mathrm{t}^{-1}$
Removal Rate	μ	$0.5\,\mathrm{t}^{-1}$
Recovery Rate	ν	$0.25\,\mathrm{t}^{-1}$
Risk Coupling	$\begin{pmatrix} \rho^{\mathrm{HH}} & \rho^{\mathrm{HL}} \\ \rho^{\mathrm{LH}} & \rho^{\mathrm{LL}} \end{pmatrix}$	$\begin{pmatrix} 1.0 & 0.008 \\ 0.008 & 0.016 \end{pmatrix}$
Spatial Coupling	σ	*see text*
Transmission Rate	β	$2.5\,\text{host}^{-1}\,\mathrm{t}^{-1}$

Optimal allocation of vaccination resources minimises the epidemic cost J, defined here as representing the disease burden of the epidemic across all infected hosts over the simulation time (T):

$$J = \int_{t=0}^{T} I(t)\mathrm{d}t \,, \tag{4.2}$$

where I is the total number of infected hosts across the network. In common with the particular control we consider and the risk and spatial structures, this simple choice of objective function was made merely to illustrate our methods, but as we later show in Chapters 6 and 7, the framework generalises immediately to more complex settings.

4.3.2 Approximate models

Exhaustive optimisation of control using the simulation model across space, risk group and time would clearly be impossibly computationally expensive, since optimisation by region involves optimising 6 vaccination rates at every time point. To make optimisation feasible we use an approximate model, as described in Section 4.2. To assess how the choice of approximation affects the control, we consider two different deterministic approximate models of the simulator. The first model is purely risk structured, factoring out all spatial

information and leaving one high risk and one low risk population group. This model is therefore based on the assumption that all nodes are spatially well-mixed with each other, and so we refer to it as the risk-based model.

The second approximate model is more complex, as it is still deterministic and risk structured, but also groups hosts by region. The model is therefore split into three regions, with two risk classes but no spatial structure within each region. Spatial dynamics are included between but not within the three regions, essentially assuming that nodes are spatially well-mixed within each region. This maintains enough simplicity to obtain optimal control results. For example, this could represent optimising control across countries, but not at the regional level within countries. We refer to this model as the space-based approximate model.

Risk-based approximate model

This model factors out all spatial information from the simulations, and approximates the resulting risk structure using a set of ordinary differential equations. The ODE system has one equation for each of the six host states: susceptible, infected and vaccinated in the high risk ($S^{\mathrm{H}}, I^{\mathrm{H}}, V^{\mathrm{H}}$), and low risk groups ($S^{\mathrm{L}}, I^{\mathrm{L}}, V^{\mathrm{L}}$). The state values are mapped from the simulations simply by summing all hosts in that state across the whole network (e.g. $S^r = \sum_i S_i^r$, where i is the node). For risk group r, the full system of equations is given by:

$$\dot{S}^r = bN^r - dS^r - S^r\left(\hat{\rho}^{r\mathrm{H}} I^{\mathrm{H}} + \hat{\rho}^{r\mathrm{L}} I^{\mathrm{L}}\right) - \frac{f^r \eta_{\max} S^r}{N^r} + \nu I^r \; ; \quad r = L, H \tag{4.3a}$$

$$\dot{I}^r = - dI^r + S^r\left(\hat{\rho}^{r\mathrm{H}} I^{\mathrm{H}} + \hat{\rho}^{r\mathrm{L}} I^{\mathrm{L}}\right) - \mu I^r - \nu I^r \; ; \quad r = L, H \tag{4.3b}$$

$$\dot{V}^r = - dV^r + \frac{f^r \eta_{\max} S^r}{N^r} \; ; \quad r = L, H \tag{4.3c}$$

where $\hat{\rho}$ is a 2×2 matrix giving the approximated risk structure, i.e. the rate at which each risk group infects each other risk group. This must be calculated by a model fitting step (see Section 4.3.3), since the exact spatial structure is not modelled. Heterogeneity in the risk structure across the network, and averaging across nodes, means that the simulation risk structure does not capture the correct dynamics in the approximate model. All other parameters are specific to individual hosts and so can be used directly from the simulation model.

Space-based approximate model

This model includes regional spatial information as well as risk structure. This gives 18 possible states: susceptible, infected and vaccinated across risk groups (high and low),

and regions (A, B and C). The states here are summed across nodes within a region (e.g. $S_A^r = \sum_{i \in A} S_i^r$). For risk group $r \in \{\mathrm{H}, \mathrm{L}\}$, and region $X \in \{A, B, C\}$, the system of differential equations is given by:

$$\dot{S}_X^r = bN_X^r - dS_X^r - S_X^r \sum_{\substack{X' \in \\ \{A,B,C\}}} \tilde{\sigma}_{XX'} \left(\tilde{\rho}^{r\mathrm{H}} I_{X'}^{\mathrm{H}} + \tilde{\rho}^{r\mathrm{L}} I_{X'}^{\mathrm{L}} \right) - \frac{f_X^r \eta_{\max} S_X^r}{N_X^r} + \nu I_X^r \tag{4.4a}$$

$$\dot{I}_X^r = - dI_X^r + S_X^r \sum_{\substack{X' \in \\ \{A,B,C\}}} \tilde{\sigma}_{XX'} \left(\tilde{\rho}^{r\mathrm{H}} I_{X'}^{\mathrm{H}} + \tilde{\rho}^{r\mathrm{L}} I_{X'}^{\mathrm{L}} \right) - \mu I_X^r - \nu I_X^r \tag{4.4b}$$

$$\dot{V}_X^r = - dV_X^r + \frac{f_X^r \eta_{\max} S_X^r}{N_X^r} \,. \tag{4.4c}$$

The 3×3 matrix $\tilde{\sigma}$ approximates spatial coupling between regions, and $\tilde{\rho}$ is again a 2×2 matrix giving the approximated risk structure (but note that the parameters will be different from the risk-based model).

4.3.3 Model fitting

The approximate models both have parameters that must be fitted to the simulation model output. This ensures the dynamics of the simulation are captured as accurately as possible given the simplifications made in the approximate models. For both models we use maximum likelihood estimation (MLE) to fit the parameters (Aldrich, 1997). States must be mapped between the simulation model and the approximate models. As described previously, the simulation states are simply summed across all nodes in the network for each risk group in the risk-based model, and by region for the space-based approximate model.

To fit the risk-based model the 4 parameters in $\hat{\rho}$ must be chosen. The likelihood is defined as the probability of observing a set of simulation realisations, given a parameter set in the approximate model. Maximising the likelihood finds the best fitting set of parameters. The likelihood is computed as the product of contributions from 200 realisations of the simulation model with no interventions (i.e. no control is implemented). Since all events occur with exponentially distributed waiting times, each individual event within a realisation occurring after a time δt contributes a factor δL_i to the likelihood:

$$\delta L_i = r_k \mathrm{e}^{-\sum r_i \delta t} \,. \tag{4.5}$$

The r_i are the rates of all possible events, and r_k is the rate of the event that actually occurs. This δL_i is the probability of observing that event given the rates of all events, which are calculated from Equations 4.3. The full likelihood is the product of all δL_i across all realisations. This overall likelihood is then maximised by varying the elements of $\hat{\rho}$ using the

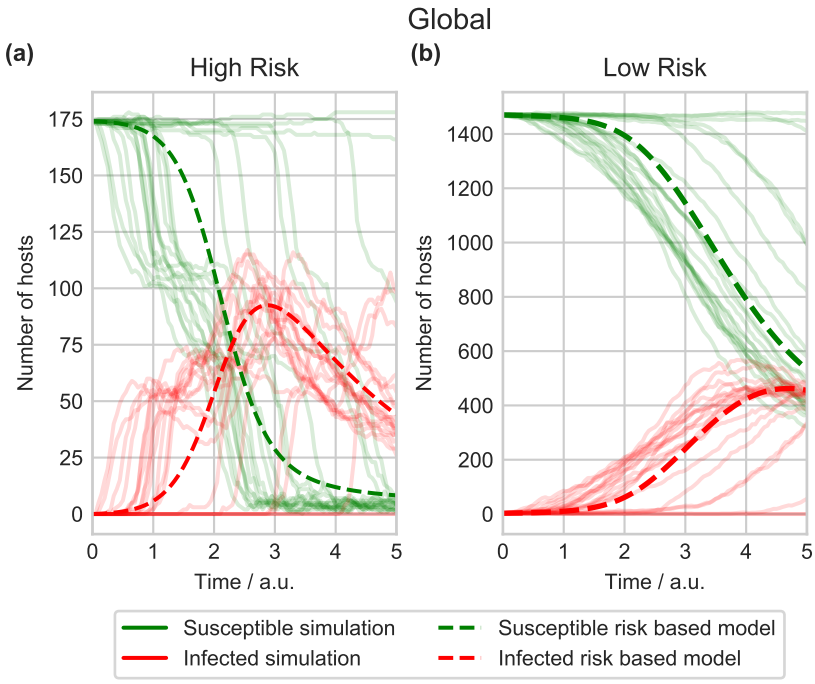

Figure 4.4: Fits to simulation data for the risk-based model. Dashed lines show the risk model, and faded lines show 20 of the simulation runs. Green and red lines represent susceptible and infected hosts respectively.

L-BFGS-B algorithm (Byrd et al., 1995), as implemented in the SciPy library (Jones et al., 2001–).

The same fitting procedure is used to fit $\tilde{\rho}$ and $\tilde{\sigma}$ in the space-based model. Here there are a total of 13 parameters, but we set $\tilde{\rho}^{\mathrm{HH}}$ to one to ensure identifiability (since the infection rate is proportional to the product of $\tilde{\rho}$ and $\tilde{\sigma}$). To simplify estimation, we also set the coupling between regions A and C ($\tilde{\sigma}_{AC}$ and $\tilde{\sigma}_{CA}$) to zero, as well as from B into A and from C into B. This is because when seeded in A, the epidemic spreads from A to B to C, and backward spread is negligible. This leaves a total of 8 parameters to fit.

Fits to simulation data are shown for the risk-based model in Figure 4.4, and in Figure 4.5 for the space-based model. The risk model captures the median risk dynamics well. The highly stochastic nature of spread between regions however, means the deterministic spatial model does not necessarily capture the timings of introductions accurately. This effect is due to stochastic fade outs after introduction events, as well as negative covariance between susceptible and infected hosts, leading to reduced infection rates in the stochastic simulations (Keeling and Rohani, 2008, pp. 227–229 and pp. 238–240). For now we continue to use MLE despite this limitation, but we show in Appendix A that alternative fitting methods do not change our results.

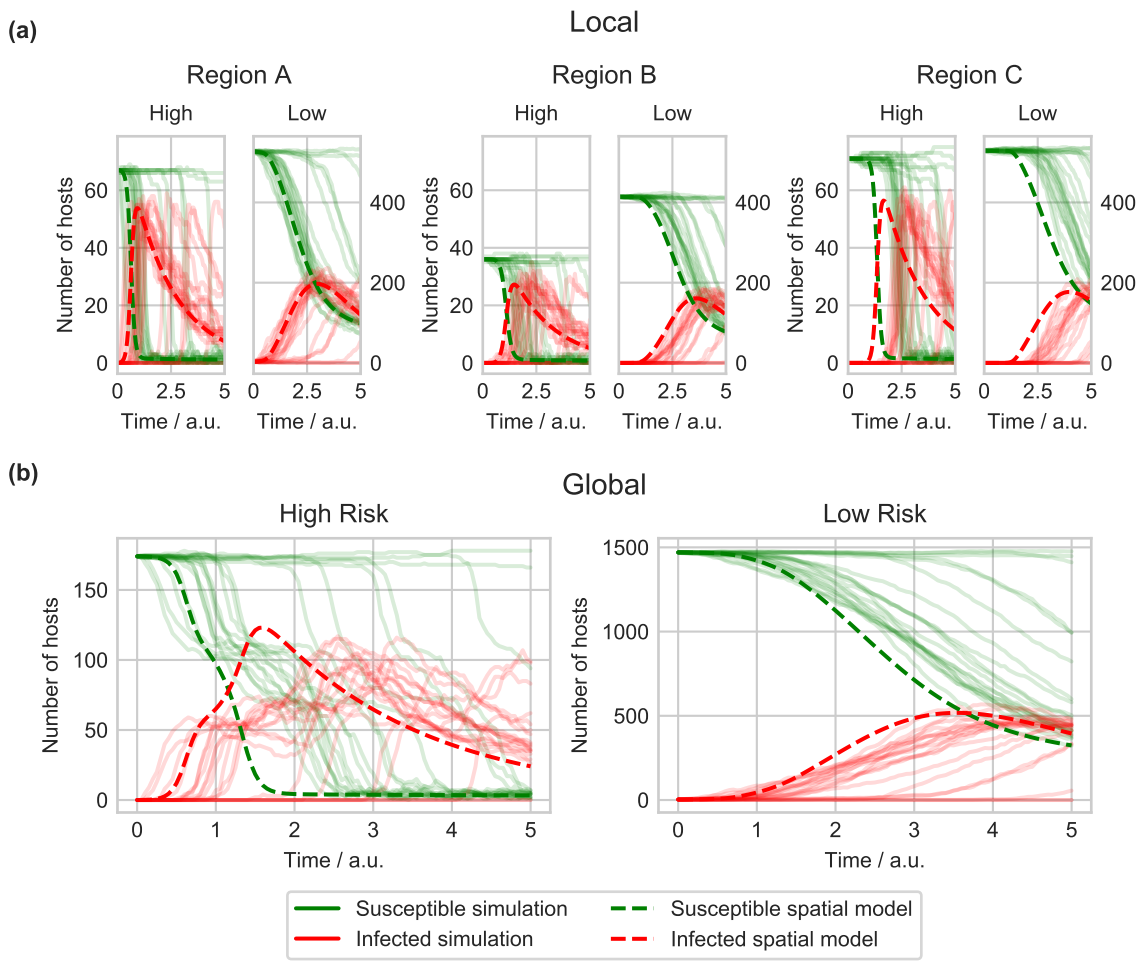

Figure 4.5: Fits to simulation data for the space-based model. Dashed lines show the space-based model, and faded lines show 20 of the simulation runs. Green and red lines represent susceptible and infected hosts respectively. We note that the approximate dynamics are faster than those in the simulations. This effect is explained in the text and does not affect our results.

To verify the optimisation we generated profile likelihood plots (Bolker, 2008), shown in Figure 4.6. The profile likelihood of a given parameter θ scans over values of θ. At each value θ is fixed, and all other parameters are re-fitted, maximising the likelihood subject to the fixed value of θ. This is carried out for each parameter in each model. This shows the maximum possible likelihood for each parameter value, increasing confidence that the fitting process has found the correct maximum. A confidence interval can be constructed from the log-likelihood ratio statistic, finding the likelihood at which there is a 95 % probability of difference between the models. This corresponds to:

$$2\log\left\{\frac{\hat{\pi}}{\pi}\right\} \leq \chi^2_{0.95}(1) \tag{4.6}$$

where $\hat{\pi}$ is the maximum likelihood, and π is the profile likelihood, and these are compared with the 95$^{\text{th}}$ percentile of the chi-squared distribution with one degree of freedom. All profile

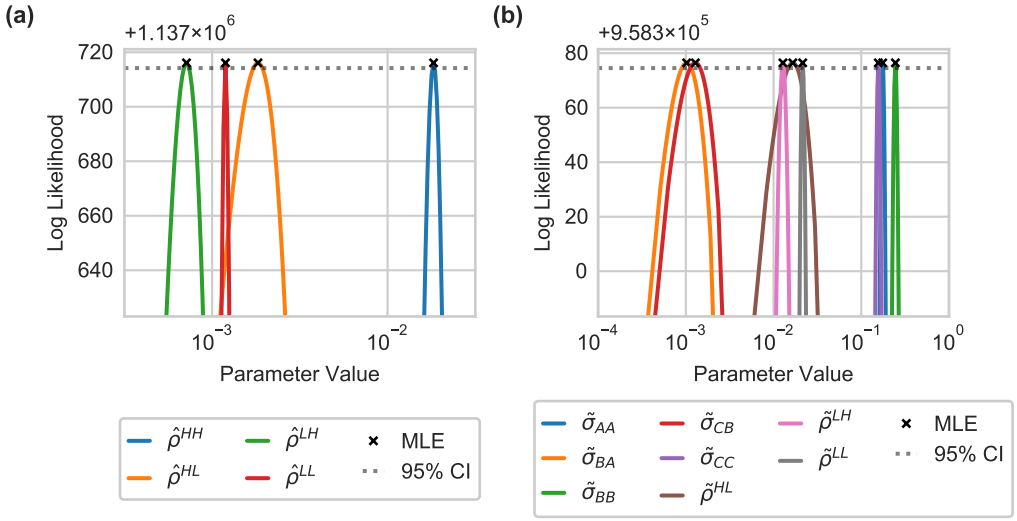

Figure 4.6: Profile likelihood analysis for: **(a)** the risk-based approximate model, and **(b)** the space-based approximate model. The offsets (1.137×10^6 for **(a)** and 9.583×10^5 for **(b)**) must be added to all values on the y axes. The profile for every parameter crosses the 95% confidence interval threshold, ensuring identifiability of each parameter. In all cases maximum likelihood estimates from the unconstrained optimisation coincide with the maximum along the profile.

likelihoods cross this threshold, meaning that the parameters are practically identifiable (Raue et al., 2009).

4.3.4 Control scenario testing

We test six different control scenarios, which compare strategy testing of controls based purely on the simulation model (scenarios 1 and 2, 'user-defined') with open-loop and MPC applied using both of our approximate models (scenarios 3 to 6):

1. **'High':** exclusively vaccinate high risk individuals
2. **'Split':** partition control resources between high and low risk groups based on an optimisation performed in advance using the simulation model only
3. **'Risk OL':** open-loop control using the risk-based approximate model
4. **'Risk MPC':** MPC using the risk-based approximate model
5. **'Space OL':** open-loop control using the space-based approximate model
6. **'Space MPC':** MPC using the space-based approximate model

The optimal constant allocation for the 'Split' strategy was found by running many simulation model realisations for each of a range of partition values, similar to the approach used by Cunniffe et al. (2015). The partition values specify the proportion of resources that are allocated to each risk group. We selected the value that gave the lowest mean epidemic

cost. The six strategies are assessed by repeatedly running the simulation model under each control scenario.

Optimisation in scenarios 3 to 6 is carried out using BOCOP v2.0.5 (Team Commands, Inria Saclay, 2017), as described in Chapter 3. We implement open-loop and MPC using Algorithm 2. When lifting the resulting controls to the simulation model, resources are allocated such that all non-removed individuals within the targeted group have an equal probability of being vaccinated. With the risk-based model, resources are allocated across all nodes and active individuals in the targeted risk group are selected randomly, although the vaccine is only effective on susceptible hosts. In the space-based model resources are spread over nodes in the targeted region. Again, the probability of selecting any single active host is constant. This amounts to weighting the resource allocated to a particular node by its total population.

To fully compare open-loop and MPC we must run the simulations repeatedly to account for stochasticity. We do this using both approximate models as well as with the naive 'High' and 'Split' strategies that do not use OCT. As described before, the 'High' strategy allocates all control resources to the high risk group. The 'Split' strategy uses an optimisation performed in advance using the simulation model. By running the simulation model repeatedly and allocating different resource proportions to the high and low risk groups, the optimal constant ratio can be found by minimising the mean objective value (Figure 4.7). The proportion that gives the lowest mean objective value is found to be 63% to the high risk group, with the rest used to vaccinate low risk individuals. As shown in Figure 4.7 the optimum occurs in a very broad minimum, so the precise value of the optimal split is uncertain. Since the precise value has little effect on the epidemic cost (hence the broad minimum), we simply use 63% as a representative value.

Algorithm 2: MPC and open-loop algorithms. Open-loop simulates for the full time (i.e. step 2–6), whereas MPC re-optimises the control at the update times (step 2–7 with repeated loops back to step 4).

1. Fit simulation model to real data
2. Set initial conditions for simulation model
3. Fit approximate model to simulation data
4. Initialise approximate model at current simulator state
5. Optimise control on approximate model
6. Lift control to simulation model and simulate forward
7. If MPC then at next update time go to step 4

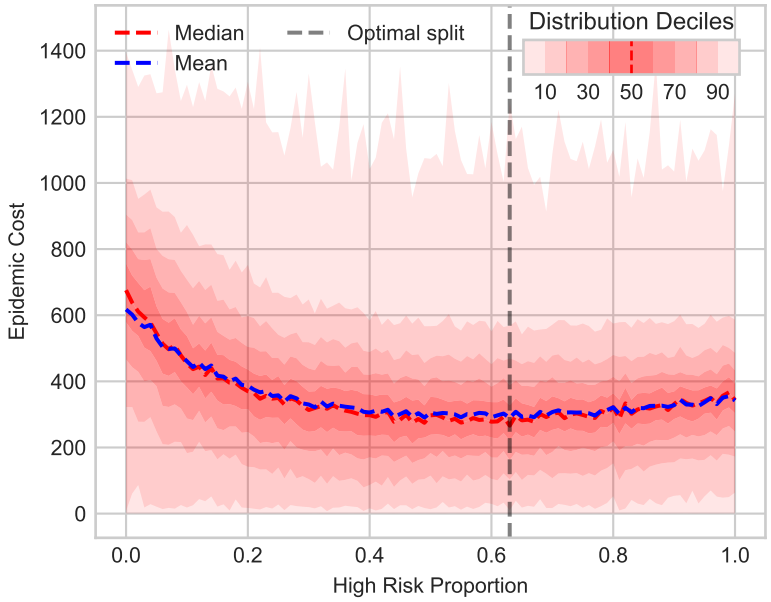

Figure 4.7: Scan of objective values for 1000 simulations, varying the proportion of control that is allocated to the high risk group. Remaining control is allocated to the low risk group. The optimal allocation to the high risk group is shown by the vertical dashed line, and is found to be 63%. Deciles in the objectives are shown by the gradient in colour.

4.4 Results

4.4.1 OCT strategies

We first look at the OCT results from optimisation on the two approximate models. The optimal vaccination strategy in the risk-based approximate model exclusively vaccinates high risk hosts early in the epidemic (Figure 4.8(a)). The strategy then switches to vaccinating both high and low risk hosts, with the majority of control resources allocated to the low risk group. There is then a further switch to vaccinating the low risk group exclusively. The optimal strategy in the space-based model shows a very similar allocation to risk groups across time (Figure 4.8(b)), but shifts these allocations across regions as the epidemic spreads through the network (Figure 4.8(c)). This allows the control to track the progress of the epidemic, and hence target control more effectively. The spatial strategies are therefore more complex than the risk-based controls.

4.4.2 Strategy testing

Next we compare the different control scenarios applied to the simulation model. This could give a more realistic indication of performance in the real world, since the simulation model is a more accurate representation of reality. First we compare the strategies resulting from the open-loop and MPC frameworks. The open-loop framework uses the same control

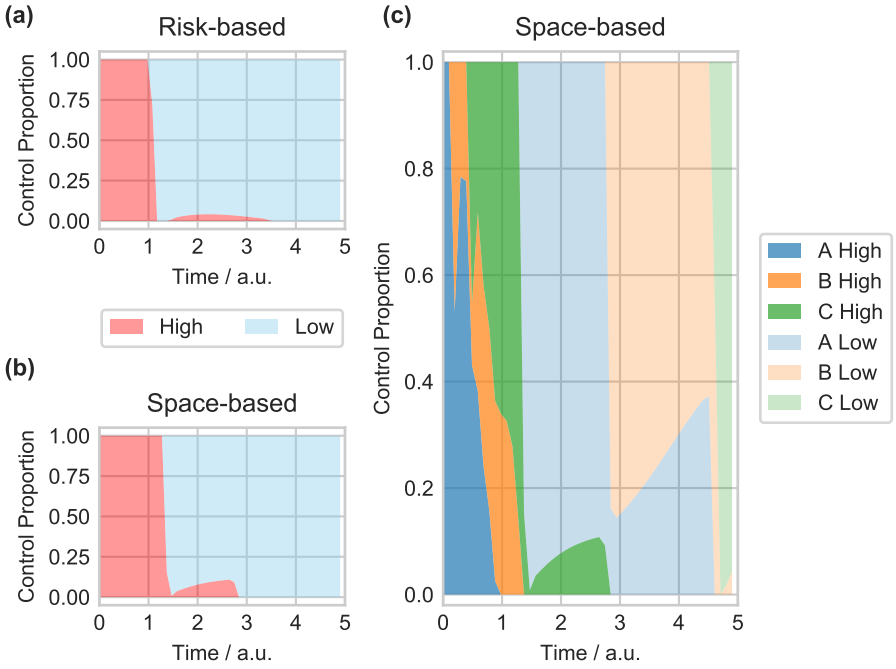

Figure 4.8: Comparison of optimal control strategies in the approximate models. **(a)** shows the optimal allocation of control resources to high and low risk groups in the risk-based approximate model. **(b)** shows the equivalent for the space-based approximate model. This allocation is broken down further in **(c)**, showing the distribution across regions and risk groups for the space-based model.

strategy (as described in the previous section) for every realisation of the simulation model. Since the simulations are stochastic, the epidemic will be different in each realisation but the time-dependent control is the same each time. The MPC framework can update the control dependent on how the individual realisation has progressed. Each realisation will therefore have a different control strategy. Figure 4.9(a) shows the control strategy for a single realisation using open-loop and MPC with the space-based model. Figures 4.9(b) and (c) show the corresponding disease progress curves in the simulation alongside the predictions made by the approximate model. We can see that the MPC framework changes the control allocations to account for deviations from the predicted behaviour. By resetting and re-optimising the approximate model the control strategy is tailored to the individual realisation. This allows the MPC framework to keep the number of infected hosts significantly lower than with the open-loop framework (note the different y axis scales on Figure 4.9(b) and (c)).

For each control scenario we run the simulation model 250 times to generate a distribution of objective values (Figure 4.10). The 'Split' strategy leads to lower objective values, and hence improved control, compared to the 'High' strategy. Using the risk-based approximate model improves the epidemic management further, but with little difference between the open-loop and MPC frameworks. Adding space into the approximate model improves control

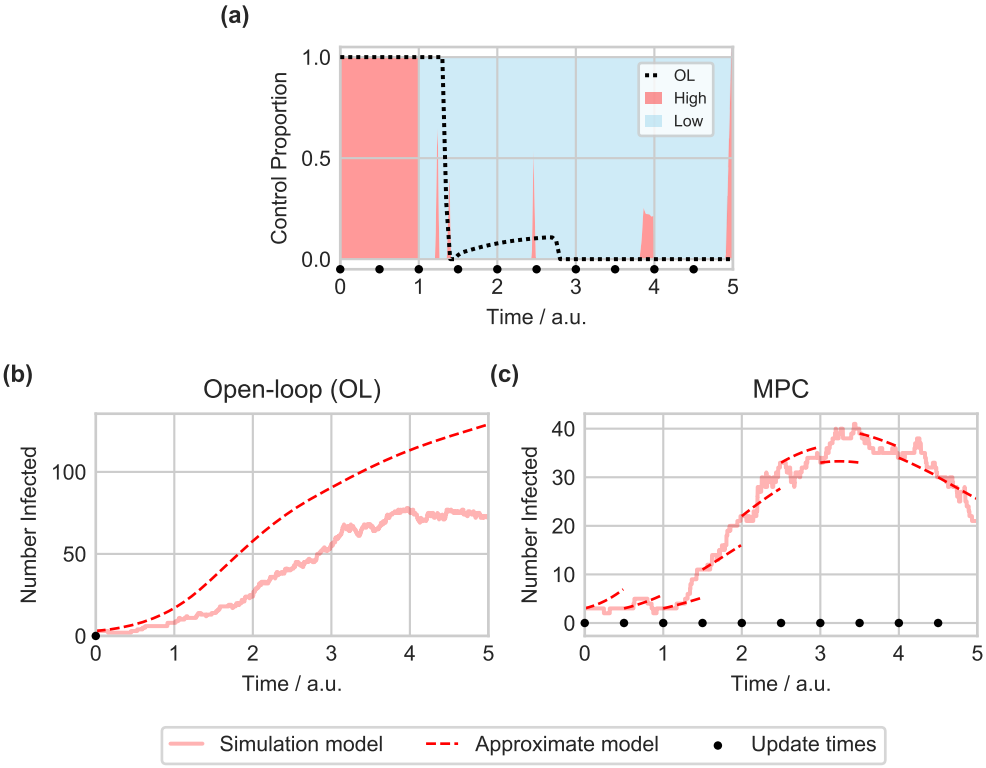

Figure 4.9: In **(a)** the control allocation is shown for a single space-based MPC run, with the corresponding open-loop allocation indicated by the black dotted line. The simulations chosen are the 60th percentiles of the epidemic cost distributions for the open-loop and MPC runs. Panels **(b)** and **(c)** are repeated from Figure 4.1. **(b)** shows the total number of infected individuals under a single run of space-based open-loop control. Control is based on the prediction of the approximate model starting from the initial conditions. **(c)** shows the number of infected individuals in the simulation and space-based approximate model corresponding to the MPC control carried out in **(a)**. Here the prediction is reset to match the simulation at every update time (0.5 time units) and the control is re-optimised. By repeatedly correcting for differences between short-term model predictions and realised numbers of infected individuals—rather than relying on a potentially increasingly inaccurate prediction made at the initial time—MPC gives better predictions of the simulation state as well as improved control when compared to open-loop (note different y axis scales).

further again, leading to the smallest epidemic costs when the spatial MPC framework is used.

The results here demonstrate the management improvements that can be achieved by combining OCT with both open-loop and MPC frameworks. The key results of the OCT analyses are the control switching times. Using the switching controls from either approximate model with open-loop control gives lower epidemic costs than the naively chosen 'user-defined' strategies. The feedback present in the MPC controllers allows further reductions to the epidemic cost. By re-evaluating the timing of the switches during the epidemic, and potentially including additional switches, the control can respond more closely to the exact trajectory of the current simulation realisation (Figure 4.9). This gives control

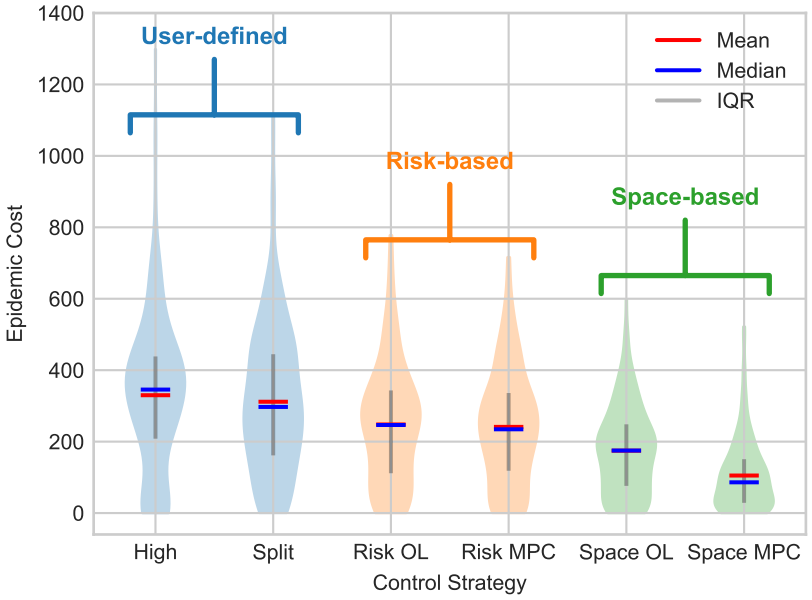

Figure 4.10: Results of different control optimisation schemes on the illustrative simulation model. Spatial MPC performs best, showing an improvement over both open-loop and user-defined strategies.

that is more robust to uncertainty and systematic errors in the approximate model, and hence performs better on the complex simulation model.

In the risk-based strategies there is little difference between open-loop and MPC. This is because the precise timing of the switch from high to low risk group vaccination does not significantly affect the epidemic cost. Figure 4.11 shows simulation objectives scanning over this switch time, where control vaccinates exclusively high risk hosts until the switch time, then only low risk hosts after the switch. These simulations lead to a broad minimum in the objective, showing that small changes to the switch time have little effect. The timings of disease introduction into regions B and C are highly variable between simulation runs (Figure 4.3). The potential for additional switches in the space-based approximate model gives more flexibility for the MPC controller to respond to this variability, and so spatial MPC shows a significant improvement over open-loop which cannot adapt to perturbations. The performance of the control is closely linked to the accuracy of the approximate model. In our example, spatial dynamics are clearly important because of the timing of spread between regions, and so the more informed controls of the spatial model outperform the risk-based strategies.

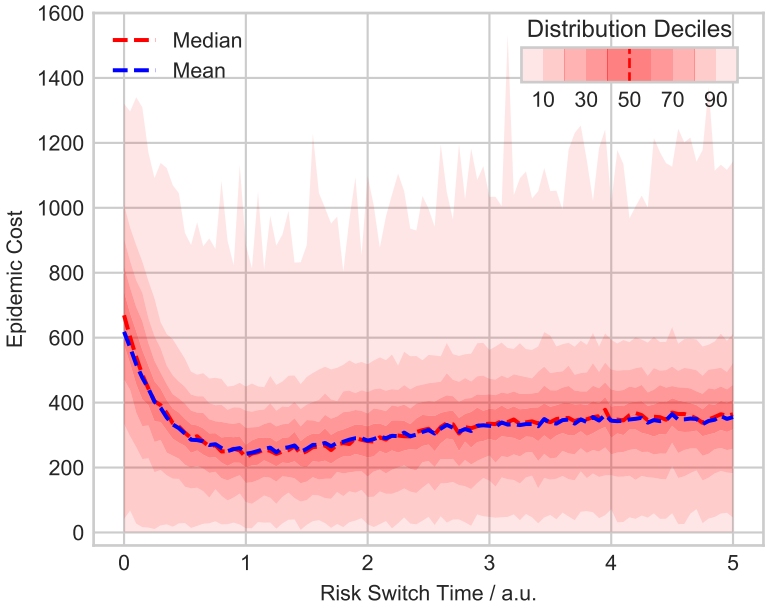

Figure 4.11: Scan over switch time showing distribution of objective values for 1000 simulations at each time. The switch time specifies when the control stops vaccinating only high risk hosts, and starts vaccinating only low risk hosts. Deciles in the objectives are shown by the gradient in colour. Whilst the switch time does affect performance, the precise timing does not vary the epidemic cost significantly.

4.4.3 Parameter robustness

The MPC framework is expected to provide improved control regardless of the exact form taken by the model or objective function. However, this does rely on using an appropriately accurate approximate model so that the OCT results are based on realistic dynamics. We tested a range of arbitrary but reasonable parameter sets, and in all cases the spatial MPC framework performs best in this network model. To illustrate this in a concrete setting we explore systematic adjustments to the risk structure used in the simulation model.

The risk structure is defined by the matrix ρ, which in the standard simulation model is given by:

$$\rho = \begin{pmatrix} \rho^{\mathrm{HH}} & \rho^{\mathrm{HL}} \\ \rho^{\mathrm{LH}} & \rho^{\mathrm{LL}} \end{pmatrix} = \begin{pmatrix} 1.0 & 0.008 \\ 0.008 & 0.016 \end{pmatrix} \tag{4.7}$$

We make the system more homogeneous or more heterogeneous by respectively doubling or halving ρ^{HL}, ρ^{LH} and ρ^{LL}. That is, the two alternative matrices used are:

$$\rho_{\mathrm{hom}} = \begin{pmatrix} 1.0 & 0.016 \\ 0.016 & 0.032 \end{pmatrix} \tag{4.8a}$$

$$\rho_{\mathrm{het}} = \begin{pmatrix} 1.0 & 0.004 \\ 0.004 & 0.008 \end{pmatrix}. \tag{4.8b}$$

Using these values we then scale the transmission rate parameter β, such that the mean epidemic cost under no control is within 1% of that using the standard risk structure. The values were found to be $\beta = 1.49$ and $\beta = 4.30$ for the more homogeneous and heterogeneous cases respectively, compared with $\beta = 2.5$ for the default risk structure. The value of ρ^{HH} is left equal to one without loss of generality since the whole matrix is scaled by β. The means we only vary the relative transmission rates.

For each new ρ matrix, we rerun the full analysis described above, assessing the same six control scenarios with refitted approximate models. Two main effects can be investigated using this analysis. Firstly, the ordering in performance of the six control scenarios can be compared with the ordering using the default risk structure, ρ. Secondly the performance of each strategy can be compared with the same strategy using ρ. We now look at each of these in turn.

For ρ_{hom}, the more homogeneous case, we find that the order of the control scenarios is unchanged (Figure 4.12). As before we find that spatial MPC leads to the best performance, as found for the default ρ. The more heterogeneous case results in a different ordering of control strategy performance (Figure 4.13). The ordering of the 'user-defined' and risk-based strategies is as before, but when the risk structure is more heterogeneous (i.e. using ρ_{het}) the spatial open-loop scenario leads to worse performance than the risk-based strategies. This is because for this parameter set and approximate model, the resulting strategies cannot respond to the variability in timing of the epidemic invading a new region, meaning that control is then targeted at the wrong regions. Importantly though, the spatial MPC strategy reduces average epidemic costs below those of all other strategies. Here the feedback strategy can greatly improve disease management, despite the limitations of the standard open-loop approach. This is similar to the effect seen by Forster and Gilligan (2007), where inaccuracies in the switch times lead to ineffective control, but here the feedback loop has mitigated this issue, ensuring that the OCT results are still beneficial.

We now consider the second effect, namely relative performance of each strategy under the different risk structures (Figure 4.14). For ρ_{hom}, the 'user-defined' and risk-based strategies have higher epidemic costs than were found with the default risk structure, ρ. This is because with a more homogeneous system, risk is less important and so the risk-based strategies are less powerful. Using ρ_{het} the risk-based strategies perform relatively better because of the increased importance of risk structure in the simulations. There is little change in epidemic cost for the spatial open-loop strategy using all three risk structures.

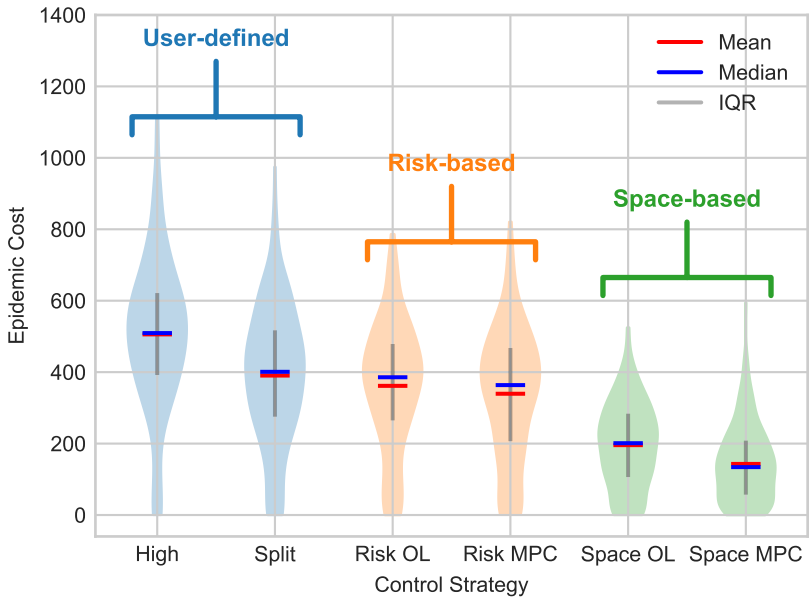

Figure 4.12: Results of control optimisation scenarios with alternative, more homogeneous risk structure ρ_{hom}.

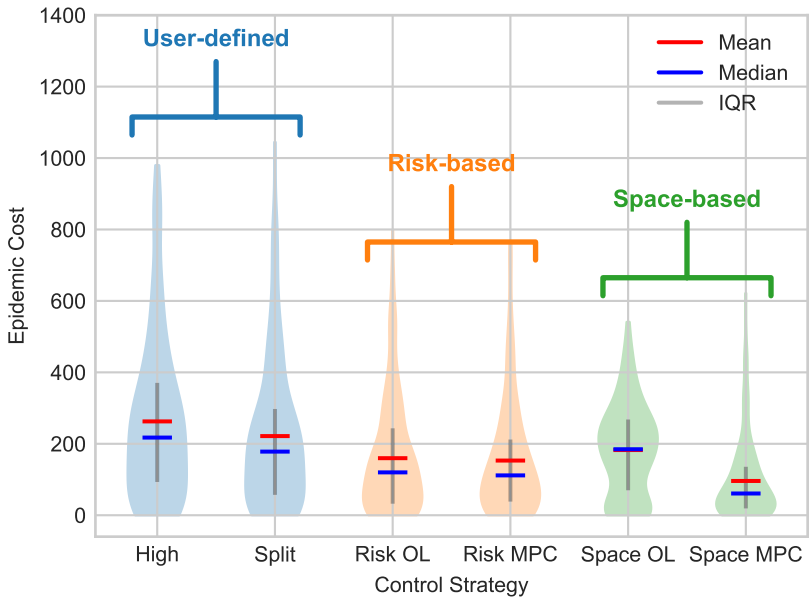

Figure 4.13: Results of control optimisation scenarios with alternative, more heterogeneous risk structure ρ_{het}.

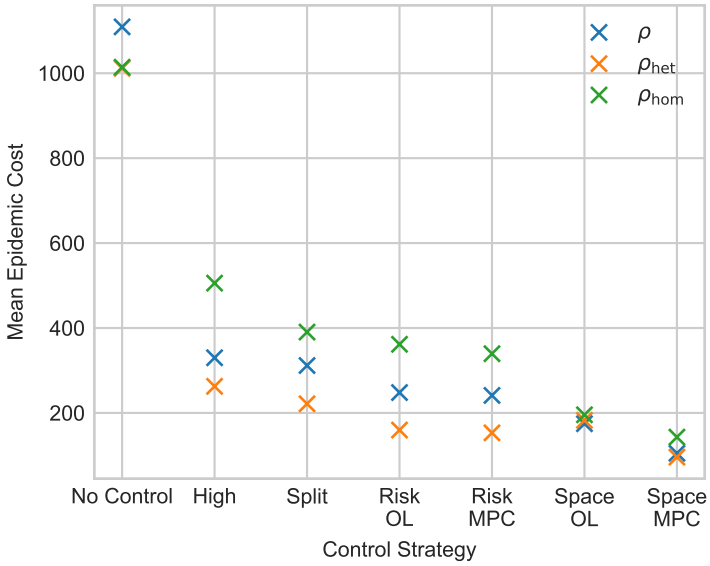

Figure 4.14: Mean performance of each control strategy using all three risk structures.

4.5 Discussion

In this chapter we have described the open-loop and MPC frameworks for applying results from OCT to complex, simulation models. Using an illustrative network model we have demonstrated how both frameworks can find strategies that are an improvement over 'user-defined' policies. The feedback loop in the MPC framework improves control further, by ensuring the simulation and approximate models are always closely aligned. In this section we will discuss some of the limitations and outstanding questions about the frameworks, and how the resulting controls might be applied in practice.

4.5.1 Applying and testing OCT

In this chapter we aimed to find methods for putting OCT results into practice. We focussed on how OCT can be used—when coupled with a simulation model using the open-loop or MPC frameworks—to optimise control in a complex system. The lifting of results to a simulation model that captures more realism allows the control based on the assumptions in the approximate model to be tested. This testing of OCT results is vital to ensure that the mathematical results are at all applicable in the real world. We here tested a single simulation model, and also tested robustness to the risk structure. Other similar spatial networks showed comparable results, but we have not systematically tested sensitivity to spatial structure or connectivity between regions. As discussed in Chapter 3, this connectivity is likely to affect the switching times—in this chapter switches between risk groups and

between regions. Further investigation into this dependence could provide insight on how enforcing changes to connectivity can impact the efficacy of control interventions.

Coupling OCT with simulation models allows the design of strategies which make the most out of both systems: OCT brings mathematical rigour and optimisation of time-dependent controls, whilst simulation models allow for highly accurate predictions of future spread. The simulation model we used here was a stochastic network model, but in principle any model that can accurately evaluate control strategies for the system in question could be used. The example in this chapter showed that the OCT strategies from the approximate model performed well on the simulation. However, in general it can be difficult to know when an OCT approach is likely to be effective, and whether to apply the open-loop or MPC framework. When the simulation model accurately captures real-world dynamics, the ultimate test to answer these questions is to compare them when applied to the simulation model, as we have done in this chapter.

4.5.2 Robust framework

As well as knowing whether open-loop or MPC will be more effective, choosing an appropriate approximate model can be challenging. Our results show that the choice of approximate model affects the performance of both open-loop and MPC strategies. Here we have found a suitable approximate model in an ad hoc manner, by testing two potential models. Clearly other choices could have been made though; for example the patch model from Chapter 3 would also have been appropriate. However, exhaustively testing all possible options is unlikely to be feasible, so modelling decisions must be made to determine what approximations are likely to be valid and how much complexity is necessary. In some cases, for example when parameters are very uncertain or the system is highly stochastic, simpler models may actually fit better than more complex models, and result in better control.

A more accurate model may give better predictions, and hence control that is closer to the true optimum, but simpler models are often sufficient (Thompson and Hart, 2018). Accuracy must be balanced with added complexity and optimisation constraints. It may not always be clear whether the complexity of a model will be prohibitive in solving the optimal control problem. The number of state and control variables is the main factor affecting tractability, but complex dynamics can also influence convergence. An advantage of the direct optimisation approach is that adjoint and Hamiltonian equations do not have to be derived for each model. This can speed up the testing of a plausible approximate model. As well as this, the open-loop and MPC frameworks can be applied to any simulation and approximate models. The only requirement is that there is a mapping of states and controls between the two models.

The frameworks are also agnostic to the form of the objective function. The choice of objective function can significantly impact optimal control strategies (Probert et al., 2016), and in this chapter we chose a very simple objective. However, the frameworks immediately extend to other choices of objective function. An advantage of OCT is the ability to balance multiple costs as shown, for example, by Brown and White (2011) and Bokil et al. (2019). In Chapter 6 we show how objectives can be used that balance economic, cultural, and ecological management goals.

The feedback in the MPC framework ensures the approximate model and optimised control match the simulation realisation, making the system robust to differences between the models. Whilst exhaustive testing of alternative simulation model parameterisations is beyond the scope of this study, we have shown that the performance of MPC is robust to one type of re-parameterisation in Section 4.4.3. We have assumed throughout that an accurate simulation model of the real system in question can be built, and that a single set of parameters can be fitted for the chosen deterministic approximate model. In reality, parameters for plant diseases models are often fitted to limited data using Bayesian methods, leading to parameter distributions (e.g. Kleczkowski and Gilligan, 2007; Parry et al., 2014). This could mean that fitting a single deterministic model may be challenging. The lack of data and the huge potential losses from plant disease (Savary et al., 2019) mean that robust decision making in the face of uncertainty is vital. The continued surveillance in the update steps of MPC could allow robustness to this type of parameter uncertainty: an important question we will address in Chapter 6. The continued surveillance could also incorporate improved knowledge of parameters as the simulation proceeds (Thompson et al., 2018).

4.5.3 Practical implementation

The feedback in MPC requires continued surveillance of the system to assess the state of the epidemic. Both open-loop and MPC are known as feed-forward controllers, since control is optimised using predictions of the future dynamics. Accurate predictions can avoid continuous or very frequent surveys which may be expensive or logistically challenging. However, the repeated updates in the feedback loop of MPC improve these predictions and hence the performance of the control. Whilst each update improves control, the associated surveillance could be expensive, and so surveillance costs must be balanced with the benefits of improved disease management. This could be done by varying the update frequency or the intensity of surveillance, both of which we analyse in more detail in Chapter 6.

We have shown in this chapter that open-loop and MPC frameworks are able to transfer optimal control results to more realistic simulations and so to practical application, but the complexity of the resulting strategies does raise the issue of communicability of results. With complex feedback strategies between two models, one complex in structure and the other

mathematically complex, the overall result is no longer simple to explain. Also, adoption by stakeholders could be prevented in practice by the plausibility of changing control strategy so frequently. However, the ultimate proof of performance is in the strategy testing technique using the simulation model. If potential strategies can be evaluated by stakeholders for real-world implementation using a complex model, as was the case for HPV (Choi et al., 2010) and citrus diseases (Cunniffe et al., 2015) for example, then the open-loop and MPC frameworks can simply be added as an alternative strategy for evaluation. Future research must therefore focus on improving the accuracy of simulation models, and analysing their reliability, so that simulations can be used to establish conclusively the benefit of these complex OCT based strategies.

4.6 Conclusions

OCT can be used to find effective control strategies for complex systems by applying OCT to an approximate model of the simulation. Results can then be lifted to the simulation for evaluation. In open-loop the optimised controls are used over the full simulation time, whereas with MPC a feedback loop updates the approximate model, and so the optimal control, at regular intervals. At these update times the initial conditions of the approximate model are set to the current state of the simulation model, and control is re-optimised and lifted to the simulation. Feedback allows control to closely match individual simulation realisations, improving control over open-loop. MPC provides an effective framework for identifying optimal control strategies.

5 Modelling tanoak decline in mixed species forest stands

5.1 Introduction

Throughout this work we have developed methods for evaluating and optimising disease management. In this chapter we will begin to apply these methods to a real-world case study: the control of the *Phytophthora ramorum* outbreak in the United States. We introduced *P. ramorum* and the current status of sudden oak death (SOD) epidemics in Chapter 1, and in this chapter we will focus on building a model of SOD within individual 20 ha mixed species forest stands. In the next chapter we will use this model to optimise strategies for stand-level disease management. Whilst widespread eradication of SOD in California is now impossible (Cunniffe et al., 2016), localised control within a stand for example, can still be effective at slowing disease spread (Hansen et al., 2019). Extensive control in Oregon is containing the spread within Curry county, with 2028 and 2038 being the estimated years of arrival into Coos county, with and without control respectively (Oregon Department of Forestry, 2019). Local controls can be particularly effective when management objectives are focussed on protection of high value resources with cultural, ecological or economic importance: the type of optimisation problem considered in Chapter 3. In these cases control does not necessarily require global eradication, allowing management effort to be highly targeted.

In this chapter we will look at what management goals are appropriate for SOD spread in mixed species stands, in particular for protecting a culturally important species: tanoak. Informed by these goals, we will develop a model that captures the dynamics of tanoak decline and potential control measures against SOD. This model will be based on work by Cobb et al. (2012), but adapted and reparameterised to realistically model host dynamics in a mixed species stand[1].

5.1.1 The tanoak tree

The tanoak, *Notholithocarpus densiflorus* syn. *Lithocarpus densiflorus*, is a medium sized Californian tree related to the American chestnut. Tanoak is present from along the western

[1] All code for this chapter is available at `https://github.com/ehbussell/MixedStand`

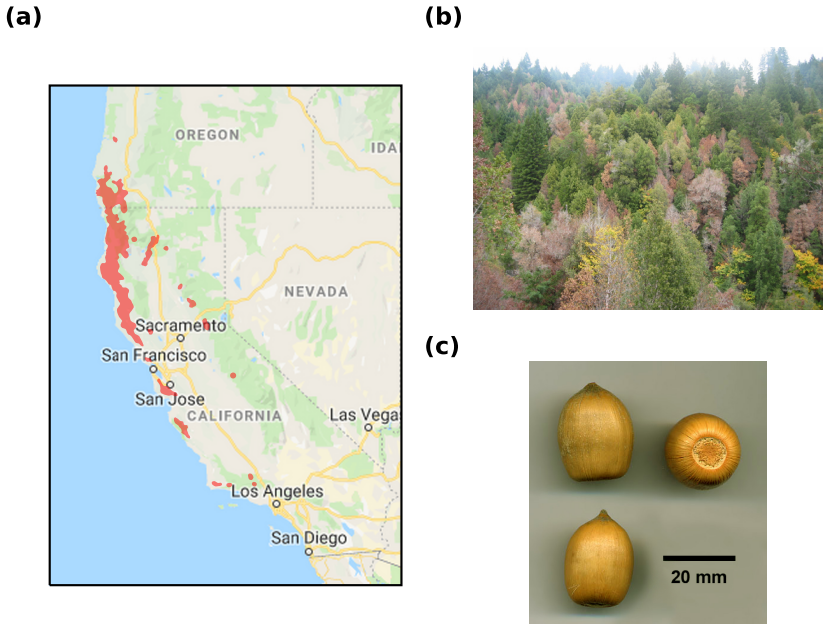

Figure 5.1: **(a)** Map showing the spatial range of tanoak in California and Oregon. Tanoak is found in the red areas. Data from: *PlantMaps* (2019). **(b)** Photograph of tanoak decline due to sudden oak death in Humboldt county, California from 2006. Source: Valachovic (2006). **(c)** Photograph of tanoak acorns, an important source of food for northern Californian Native Americans. Source: Wikimedia Commons (2004).

coast of California near Santa Barbara, up into southwestern Oregon (see Figure 5.1) (Tappeiner et al., 1990). Tanoak grows most effectively in humid conditions with seasonal precipitation. It grows from sea level up to elevations of 1,500 m and can thrive on soils less suitable for conifers. However, its higher moisture requirements mean it is found more often on northern slopes than southern slopes (Tappeiner et al., 1990). Bowcutt (2013) writes about the importance of the tanoak tree to Native American tribes in California. In this study Bowcutt explains that mature trees are heavy producers of acorns, which are highly valued by indigenous tribal people. The acorns are an important source of food for northern Californian Native American tribes, as much now as historically, with only salmon eaten in greater quantities. The thick shell and high tannin content mean the acorns can be stored for years, and hence for thousands of years they have been used as the basis for trade between tribes. Bowcutt emphasises that the tree is used by Native American tribes for much more than just nutrition though, with fishing nets, baskets and medicines made using tanoak bark and wood.

Tanoak is found alongside coast redwood (*Sequoia sempervirens*) across coastal Californian forests, and is believed to be highly important to these ecosystems (Noss, 2000). As

well as their cultural importance, tanoaks provide habitat and a vital winter food source for a variety of wildlife in these forests, including fishers and owls (Long et al., 2018). The winter food source is particularly important given the more unpredictable and less nutritious crops from redwood trees. Damaged—and even sometimes healthy—tanoak trees resprout prolifically (Tappeiner et al., 1990). This, and abundant seed production, could explain why tanoak is so ubiquitous alongside redwood, as it competitively excludes other species that could live in redwood forests (Ramage et al., 2011). This resprouting also ensures that the population regenerates after the forest fires that are common in this region (McDonald and Tappeiner, 2002; Ramage et al., 2010).

5.1.2 Effects of sudden oak death

The spread of *P. ramorum* is a significant threat to tanoak, having caused drastic declines in populations that, if continued, could lead to the extinction of this important species (McPherson et al., 2010). Tanoak trees of all ages are highly susceptible to SOD and have a very high mortality from the disease (Davis et al., 2010). In most other species susceptible to SOD such as coast live oak, bole infection—where the infection spreads to the trunk—requires a secondary foliar host (Rizzo et al., 2005). As tanoak is the only species for which stems, twigs and foliage can all be infected, bole infection does not require such a secondary host and can therefore occur more rapidly and more frequently (Rizzo and Garbelotto, 2003). Tanoak infected with *P. ramorum* has a mortality rate of 6 % per year, with infection leading to eventual tree death in at least 50 % of cases (McPherson et al., 2010). Some report net mortality is likely to approach 100 % (Ramage and O'Hara, 2010). Field studies have also found that mortality increases with tree size, meaning the larger trees that produce more acorns are disproportionately affected (Cobb et al., 2012).

Maloney et al. (2005) use 120 study sites to track the establishment of SOD in coastal redwood forests dominated by redwood, tanoak and California bay laurel (*Umbellularia californica*). Whilst redwood trees are not affected by SOD, they find that the presence of bay laurel is a key factor in the decline of tanoak due to the disease. *P. ramorum* can sporulate prolifically on bay trees but the host is not killed by the disease (Davidson et al., 2008). Maloney et al. (2005) state that the differing host mortalities due to SOD could lead to dramatic shifts in forest composition. These compositional changes can have far-reaching consequences and knock-on effects. For example, tree deaths due to SOD increase levels of dry wood in forests, and so there are greater fuel loads leading to greater risk of forest fire (Forrestel et al., 2015). SOD reduces the numbers of large trees, which reduces the amount of CO_2 captured by the forest. Management of SOD to retain larger tanoak can help manage carbon emissions (Twieg et al., 2017), and these wider reaching implications make effective disease interventions even more important.

5.1.3 Predicting disease progression

Mathematical models can be used to predict the future spread of disease, and hence inform control strategies. In Chapter 1 we reviewed models of SOD (Section 1.3.2, p. 23), with models grouped largely into risk maps, landscape scale models, and smaller-scale models of forest stands. Our aim here is to optimise dynamic controls within a forest stand, with the specific goal of protecting tanoak populations. We therefore require a stand-level dynamic model, capturing the drivers of disease invasion rather than just risk factors, that can model disease impacts on tanoak. At this scale the landscape models of Meentemeyer et al. (2011) and Tonini et al. (2018) are inappropriate, since they do not explicitly model effects on different hosts. The differential equation model developed by Cobb et al. (2012) captures invasion dynamics and differing mortality and infection rates by species and tanoak age class within a stand. Whilst in its current form this model does not include controls, this is the only dynamic model at the stand level which explicitly models differences between hosts species, so we use this model to investigate the optimisation of local control strategies.

5.1.4 Key questions

In this and the following chapter, we will seek to answer the following key questions:

1. How can the model from Cobb et al. (2012) be adapted to allow optimisation of time-dependent disease management strategies?
2. How should time-dependent controls be deployed under resource constraints to best preserve the valuable tanoak population in coastal redwood forests?
3. How do these strategies compare with current recommended practice?
4. How robust and reliable are these control results? In particular, how do these strategies perform when information about the epidemiological parameters and the pathogen distribution is incomplete throughout the epidemic?

The first of these questions is the focus of the remainder of this chapter.

5.2 Mixed species forest dynamics

The epidemiological model in Cobb et al. (2012) was designed to investigate the medium- and long-term effects of sudden oak death on the host composition of mixed species, 20 ha forest stands. As we will use this model structure extensively, we will refer to this model as the Cobb model from now on. The trees present in the forest stands considered are tanoak, bay laurel and redwood. The authors were primarily interested in the decline of overstorey tanoak, here defined as trees over 10 cm d.b.h. (diameter at breast height), and the initial host compositions in the stand that lead to the eventual extinction of tanoak. The model was parameterised using data from longitudinal field studies conducted over 5 years and across 110 *P. ramorum* invaded plots and 95 uninvaded plots, all of area 500 m^2. The epidemiological model is then used to assess how the forest structure in similar plots will change over the next 100 years, under scenarios with varying initial host proportions. The model is used to find a threshold initial level of tanoak in the forest, above which disease progression leads to elimination of large tanoaks. This is found to be around 8 %, under a specific initial age structure. We will now describe the formulation of the model, as described in Cobb et al. (2012), in more detail.

5.2.1 Model description

The model tracks the stem density dynamics of three different host species or groups: redwood, bay laurel and tanoak. The redwood group is also used to represent all species that are not susceptible to *P. ramorum* infection, which for the stands considered is predominantly coast redwood. Bay laurel is a 'spreader' species that can be infected and is highly infectious, but does not itself suffer any significant effects from the disease. Tanoak however, is highly susceptible to *P. ramorum* infection and disease induced mortality is high, particularly in older and larger trees. This age dependence, and the importance of overstorey tanoak, drove the authors to divide the tanoak class into separate age groups in order to capture the effects of disease on the older trees. Four different age groups were created with the two oldest groups corresponding to the overstorey tanoak. This was deemed to capture the changes in susceptibility with age in enough detail, whilst also keeping the model as simple as possible. As differing effects on older trees are less important for the other hosts, and to reduce model complexity, the other host groups are not divided into age classes. The model tracks natural host demography, with natural mortality and seed recruitment rates for each host class. Recruitment depends on the amount of empty space available for seedling establishment, with each tanoak age class weighted to occupy differing amounts of space per stem. Over time tanoak hosts progress through the age classes. See Figure 5.2(a) for an overview of the different classes and possible transitions.

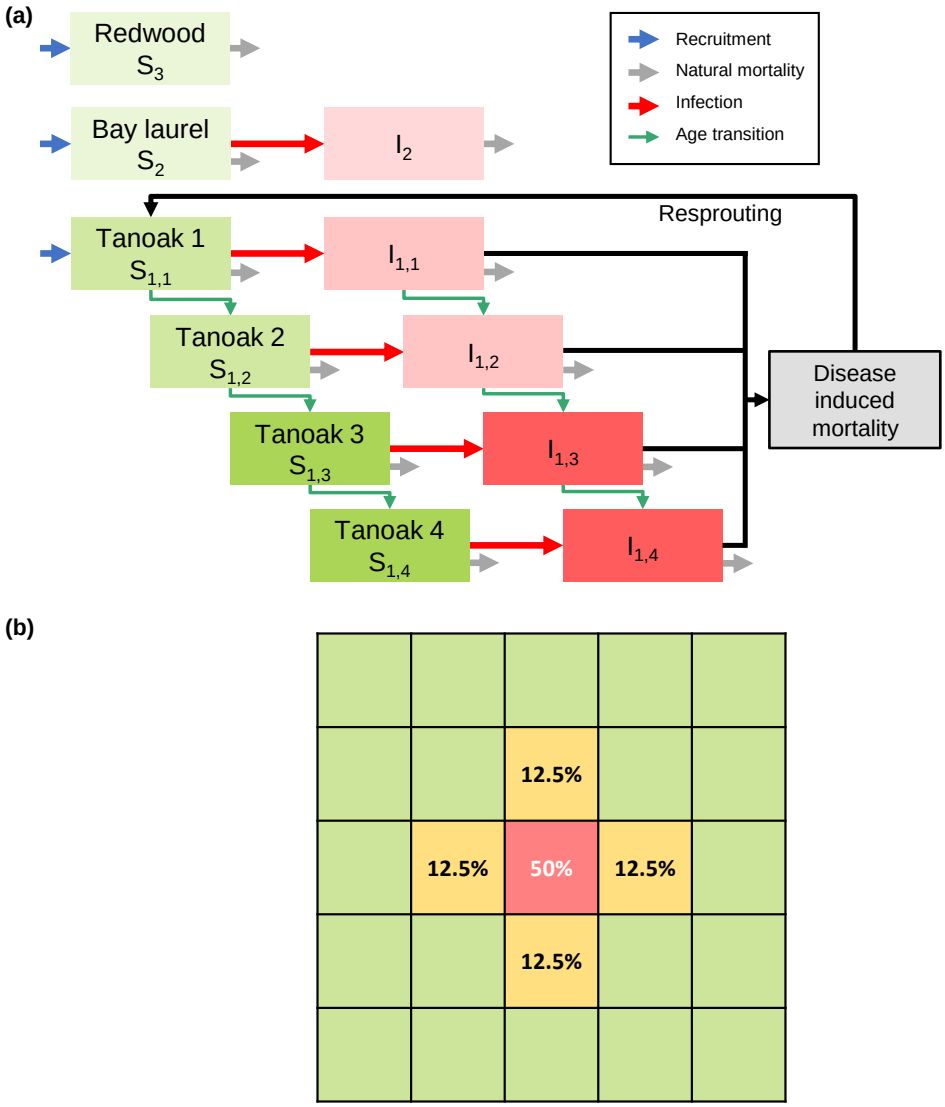

Figure 5.2: Diagram describing Cobb model host structure. **(a)** shows the possible host states and transitions. Only bay and tanoak are epidemiologically active, with the tanoak age classes grouped into small (tanoak 1 and 2) and large (tanoak 3 and 4) categories. **(b)** shows the spore deposition kernel described in Cobb et al. (2012), with 50 % of spores landing within the same cell and the other 50 % spread over the neighbouring 4 cells.

The model is spatial with hosts positioned on a grid in square cells each of area 500 m^2. Recruitment and age transitions occur within a single cell, with density-dependence in the recruitment rates based on the available space in the cell. This corresponds to an assumption that seed dispersal between cells is negligible. This is reasonable since the heavy acorns predominantly fall under the crown of the tree (Tappeiner et al., 1990). Infection dynamics are therefore the only interaction between cells. Infected hosts exert infectious pressure on susceptible hosts within the same cell and in the 4 adjacent cells. Infectious spores are distributed such that 50 % land within the same cell, and the other 50 % are distributed across the adjacent cells (Figure 5.2(b)). Bay and all age classes of tanoak are susceptible, and are infectious once infected without a latent period. There is little information available on latent *P. ramorum* infections, but evidence suggests progression to inoculum production happens quickly (within a few months; Davidson et al., 2005). This level of latency is insignificant over the time period we have considered (100 years).

The model is formulated as a system of ODEs resulting in 11 differential equations per cell. We use S to indicate healthy hosts, I to indicate infected hosts, and subscripts to indicate species (1: tanoak, 2: bay, 3: redwood), age class (1–4 where applicable), and cell location (x), following the notation used in Cobb et al. (2012) and as shown in Figure 5.2(a). The resulting equations for cell x and age class i are:

$$\begin{aligned}\dot{S}_{1,i,x} = {} & \delta_{1,i}\left[B_{1,x}E_x + r\alpha_{1,i}I_{1,i,x}\right] - d_{1,i}S_{1,i,x} - \Lambda_{1,i,x}S_{1,i,x} + \mu_1 I_{1,i,x} \\ & + \left[1-\delta_{1,i}\right]a_{i-1}S_{1,i-1,x} - \left[1-\delta_{4,i}\right]a_i S_{1,i,x}\end{aligned} \tag{5.1a}$$

$$\begin{aligned}\dot{I}_{1,i,x} = {} & -\alpha_{1,i}I_{1,i,x} - d_{1,i}I_{1,i,x} + \Lambda_{1,i,x}S_{1,i,x} - \mu_1 I_{1,i,x} \\ & + \left[1-\delta_{1,i}\right]a_{i-1}I_{1,i-1,x} - \left[1-\delta_{4,i}\right]a_i I_{1,i,x}\end{aligned} \tag{5.1b}$$

$$\dot{S}_{2,x} = b_2\left(S_{2,x} + I_{2,x}\right)E_x - d_2 S_{2,x} - \Lambda_{2,x}S_{2,x} + \mu_2 I_{2,x} \tag{5.1c}$$

$$\dot{I}_{2,x} = -d_2 I_{2,x} + \Lambda_{2,x}S_{2,x} - \mu_2 I_{2,x} \tag{5.1d}$$

$$\dot{S}_{3,x} = b_3 S_{3,x}E_x - d_3 S_{3,x} \tag{5.1e}$$

where tanoak dynamics are given by equations 5.1a and 5.1b, bay dynamics by 5.1c and 5.1d, and redwood dynamics by 5.1e (recall that all hosts in this class cannot become infected). All parameter meanings and symbols are given in Table 5.1. The delta function at the start of equation 5.1a, $\delta_{1,i}$, is equal to one for the smallest age class, and zero otherwise. This ensures that recruitment is always to the smallest age class. The other delta functions ensure only possible age transitions occur.

Table 5.1: Parameter values used in Cobb et al. (2012). Parameter values marked with an asterisk are set in model initialisation to impose dynamic equilibrium. See Section 5.2.2 for a full description. Tanoak age classes are grouped by diameter at breast height (d.b.h.).

Parameter		Symbol	Default Value
Infection rate	tanoak to tanoak (1 cm to 2 cm d.b.h.)	$\beta_{1,1}$	0.33 year^{-1}
	tanoak to tanoak (2 cm to 10 cm d.b.h.)	$\beta_{1,2}$	0.32 year^{-1}
	tanoak to tanoak (10 cm to 30 cm d.b.h.)	$\beta_{1,3}$	0.30 year^{-1}
	tanoak to tanoak (>30 cm d.b.h.)	$\beta_{1,4}$	0.24 year^{-1}
	bay to tanoak	β_{12}	1.46 year^{-1}
	bay to bay	β_2	1.33 year^{-1}
	tanoak to bay	β_{21}	0.30 year^{-1}
Natural mortality rate	tanoak (1 cm to 2 cm d.b.h.)	$d_{1,1}$	0.006 year^{-1}
	tanoak (2 cm to 10 cm d.b.h.)	$d_{1,2}$	0.003 year^{-1}
	tanoak (10 cm to 30 cm d.b.h.)	$d_{1,3}$	0.001 year^{-1}
	tanoak (>30 cm d.b.h.)	$d_{1,4}$	0.032 year^{-1}
	bay	d_2	0.02 year^{-1}
	redwood	d_3	0.02 year^{-1}
Recruitment rate	tanoak total cell x	$B_{1,x}$	Equation 5.3
	tanoak (1 cm to 2 cm d.b.h.)	$b_{1,1}$	0.0 year^{-1}
	tanoak (2 cm to 10 cm d.b.h.)	$b_{1,2}$	0.007 year^{-1}
	tanoak (10 cm to 30 cm d.b.h.)	$b_{1,3}$	0.02 year^{-1}
	tanoak (>30 cm d.b.h.)	$b_{1,4}$	0.073 year^{-1}
	bay	b_2	*
	redwood	b_3	*
Disease induced mortality rate	tanoak (1 cm to 2 cm d.b.h.)	$\alpha_{1,1}$	0.019 year^{-1}
	tanoak (2 cm to 10 cm d.b.h.)	$\alpha_{1,2}$	0.022 year^{-1}
	tanoak (10 cm to 30 cm d.b.h.)	$\alpha_{1,3}$	0.035 year^{-1}
	tanoak (>30 cm d.b.h.)	$\alpha_{1,4}$	0.14 year^{-1}
Tanoak age transition rate	(1 cm to 2 cm d.b.h.) to (2 cm to 10 cm d.b.h.)	a_1	0.142 year^{-1}
	(2 cm to 10 cm d.b.h.) to (10 cm to 30 cm d.b.h.)	a_2	0.2 year^{-1}
	(10 cm to 30 cm d.b.h.) to (>30 cm d.b.h.)	a_3	0.05 year^{-1}
Recovery rate	tanoak	μ_1	0.01 year^{-1}
	bay	μ_2	0.1 year^{-1}
Recruitment suppression weight	species i	W_i	1
	tanoak age class i	$w_{1,i}$	*
Resprouting probability	tanoak	r	0.5
Spore proportion	within cell	f_0	0.5
	between cell	f_1	0.125
Force of infection	tanoak age class i, cell x	$\Lambda_{1,i,x}$	Equation 5.4a
	bay, cell x	$\Lambda_{2,x}$	Equation 5.4b

The recruitment rates are density dependent as they depend on the space available for colonisation in each cell, E_x. The empty space in cell x is given by:

$$E_x = \max(\{0, 1 - W_1 \sum_{i=1}^{4} w_{1,i} \left(S_{1,i,x} + I_{1,i,x}\right) - W_2 \left(S_{2,x} + I_{2,x}\right) - W_3 S_{3,x}\}) \,, \tag{5.2}$$

with the maximum function ensuring that E_x cannot be negative. The species suppression weights W_j—which are assumed to all be equal to 1—give the relative area colonised, and hence unavailable for seedling recruitment, by each species per capita. The tanoak suppression weights $w_{1,i}$ give different relative space occupations for each age class. These suppression weights capture actual space occupied as well as seedling suppression by other means, for example by blocking sunlight. The tanoak recruitment rate is made up of each individual recruitment from each age class, but all seedlings enter at the smallest age class. The tanoak recruitment rate, $B_{1,x}$, in Equation 5.1a is given by:

$$B_{1,x} = \sum_{i=1}^{4} b_{1,i} \left(S_{1,i,x} + I_{1,i,x}\right) \tag{5.3}$$

so that recruitment is from all age classes, with seed production rates that are independent of infection status. Note that the tanoak natural mortality rates ($d_{1,i}$) are highest for the largest size class. Whilst many factors contribute to the age dependence of mortality, this is commonly seen for other tree species (e.g. Hurst et al., 2011) and a potential factor is the vigour of the tree. Larger and older trees grow slower because of nutrient constraints and mechanical abrasion due to wind, making the tree more vulnerable to stresses (Yang et al., 2003).

Finally we describe the force of infection terms Λ in Equations 5.1:

$$\Lambda_{1,i,x} = f_0 \left[\beta_{1,i} \sum_{j=1}^{4} I_{1,j,x} + \beta_{12} I_{2,x}\right] + f_1 \sum_{y \in N(x)} \left[\beta_{1,i} \sum_{j=1}^{4} I_{1,j,y} + \beta_{12} I_{2,y}\right] \tag{5.4a}$$

$$\Lambda_{2,x} = f_0 \left[\beta_{21} \sum_{j=1}^{4} I_{1,j,x} + \beta_2 I_{2,x}\right] + f_1 \sum_{y \in N(x)} \left[\beta_{21} \sum_{j=1}^{4} I_{1,j,y} + \beta_2 I_{2,y}\right] \tag{5.4b}$$

where β_{12} is the rate of infection from bay to tanoak, β_{21} from tanoak to bay, and β_2 within bay. The infection rate within tanoak is given by $\beta_{1,i}$, meaning each age class has a different susceptibility to infection from other tanoaks. Overall however, tanoak age classes do not vary in susceptibility to infection from bay, nor in the rate of infecting bay. Hence, β_{12} and β_{21} do not depend on age class. This is a simplification due to a lack of data. The parameters f_0 and f_1 give the proportion of spores deposited within and between cells respectively, where the sum over cells $N(x)$ is over the four cells adjacent to x.

In Cobb et al. (2012) and also throughout this chapter, the infection dynamics are initialised in the centre of a 20 by 20 grid (any one of the four central cells). Infection starts

in the bay population and the smallest tanoak age class, with 50 % of these hosts starting infected. When all 3 species are present (bay, tanoak and redwood) the proportions are taken from those used in Cobb et al. (2012) for mixed species forest stands. This corresponds to 40 % tanoak, 16 % bay, and 44 % redwood. The initial amount of empty space is found by initialising the model in dynamic equilibrium, see Section 5.2.2. The state variables track stem density in arbitrary units. The model is parameterised from the study plots in Cobb et al. (2012) which have an average tanoak density of 561 ha^{-1}. The stem densities from the model can therefore be scaled such that the initial stem density is 1,400 ha^{-1}, since tanoak makes up 40 % of the forest composition.

Conclusions using the model

Cobb et al. (2012) used this model to predict forest structure changes over the next 100 years. They found that in forests with the spreader species bay present, overstorey tanoak (defined as the two largest age classes in the model) declines to near extinction (Figure 5.3(a)). Further to this, the authors showed that in forests with no bay present, only when tanoak makes up less than 8 % of the initial stand composition is the pathogen unable to become established (Figure 5.3(b) and (c)). The Cobb study does not consider control strategies for managing tanoak decline but later studies do consider this using an updated version of the model (Ross, 2013). This R package, called SODDr, simulates the same dynamics but using a discrete time version of the model. The reason given for making the model discrete was to allow simpler inclusion of stochasticity and to improve model fitting, although both of these are possible with the continuous time version. The SODDr model has been used in conjunction with field studies to investigate the effectiveness of understorey vegetation thinning before and after an SOD disease outbreak (Cobb et al., 2017). Another study used SODDr to forecast the effect of bay removal and tanoak thinning on forest conditions (Valachovic et al., 2017a), and other recent work has looked at the impact of putative host resistance and tolerance using SODDr (Cobb et al., 2019). Whilst these controls are all either one-off interventions or natural resistance rather than the time-dependent control we will consider, this body of work indicates how the model is being used to ask questions about optimal disease management.

5.2.2 Model analysis

We aim to use the Cobb model to drive the open-loop and MPC optimisation frameworks developed in the previous chapter. This will help identify long-term, time-dependent control strategies that conserve tanoak populations in mixed species stands. More importantly, we use the Cobb model to test how reliable and robust the frameworks are to a realistic disease scenario, testing handling of uncertainties in both parameters and observations. To ensure

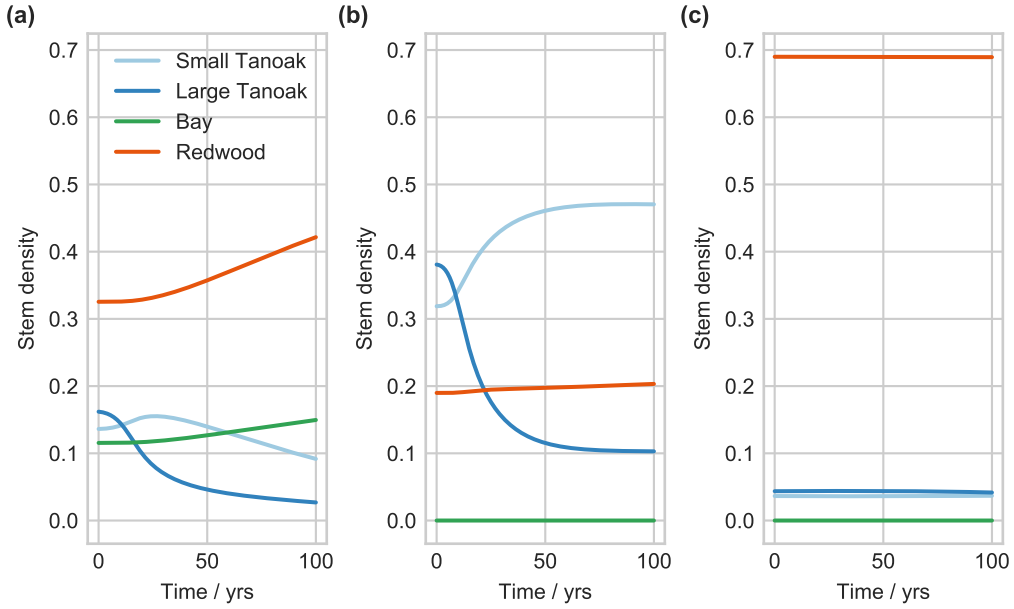

Figure 5.3: Dynamics of the Cobb model, parameterised as in Cobb et al. (2012). The three plots correspond to the three subplots in Figure 4 in Cobb et al. (2012). **(a)** shows dynamics for a stand with bay, tanoak and redwood present, where large tanoaks decline to near extinction over 100 years. **(b)** shows the dynamics with no bay present and tanoak making up 80 % of the initial composition. Here large tanoak declines but does not near extinction. In **(c)** however, with 8 % tanoak, there is no decline in the large tanoak population.

that our results are reliable we first analysed the Cobb model in more detail, to validate our own implementation and confirm results realistically capture the real-world spatial spread of SOD. To do this we tested our own Python implementation of the Cobb model versus SODDr as well as the original Berkeley Madonna (BM) code used to generate the figures in Cobb et al. (2012), which was provided to us directly by the lead author. In doing this we found some unphysical features of the Cobb model which required correction before we used the model, most importantly spore deposition patterns and implausible seedling parameters.

Spore deposition

We first analyse an error in the spore deposition pattern, resulting in spores deposited equally over the source cell and adjacent cells. This affects parameters f_0 and f_1 in Table 5.1, and means more spores are deposited than produced.

The paper explains that 50 % of spores fall within the same cell, and the other 50 % are distributed over the four nearest neighbour cells. However, in both the BM code and the SODDr code the infectious pressure between cells is not distributed in this ratio. In the BM code the source cell and each of the four nearest neighbours receive 100 % of the source cell's spores. In Equation 5.4 this corresponds to both f_0 and f_1 equal to 1.0, and hence a

clearly unrealistic 500 % of spores are deposited. In other words, an infection source cell provides the same force of infection to itself, and to each of its adjacent cells. Whilst these are just scaling coefficients, the effect of this error is to increase the spread rate, and also change the shape of the dispersal kernel. With the error the kernel is flat over the source cell and nearest neighbours rather than decreasing with distance as intended by the original authors.

The spore deposition error is partially fixed in the SODDr implementation where f_0 equals 0.5 and f_1 equals 0.125 however, the between cell spores are deposited over the 8 nearest neighbours rather than 4. This leads to 150 % of spores deposited, again physically impossible. The corrected deposition should use f_0 =0.5 and f_1 =0.125, over 4 nearest neighbours. The effect of this is to significantly slow the rate of disease spread, and in turn forest composition changes, when compared with the results from the paper (Figure 5.4).

In our implementation the realism of the spore deposition is improved further, by introducing an exponential dispersal kernel. This thin-tailed shape of kernel is more realistic for pathogens spread by small-scale splash dispersal (Skelsey and Garrett, 2013), as is the case for *P. ramorum*. The same proportion of spores (50 %) are deposited within the source cell as used in Cobb et al. (2012), corresponding to $f_0 = 0.5$. The other 50 % are distributed according to an exponential kernel with a scale parameter of 10 m. The kernel is normalised so that total spore deposition across all cells is 100 %. The spore deposition between cells becomes proportional to:

$$\exp\left(-d_{ij}/\sigma\right), \tag{5.5}$$

where d_{ij} is the distance from source cell to target cell, and σ is the scale parameter. The choice of 10 m as a scale parameter is somewhat arbitrary, although consistent with distances of splash dispersal found for *P. ramorum* (Davidson et al., 2005), and equal to the mean dispersal distance used by Cobb et al. (2012). This also falls within the 95 % credible interval found for short scale transmission by Meentemeyer et al. (2011), although they used a Cauchy type kernel and the best estimate of the dispersal scale was 20 m. We here use a smaller value to capture very local spread from splash dispersal, particularly given reports suggest that these larger scale models of SOD tend to overestimate disease spread (Valachovic et al., 2017b).

Model initialisation

A further problem with the Cobb model is the parameterisation and initialisation of the system, which leads to unrealistically high recruitment rates and suppression weights for the smallest tanoak age class. This affects parameters $b_{1,1}$ and $w_{1,1}$ in Table 5.1.

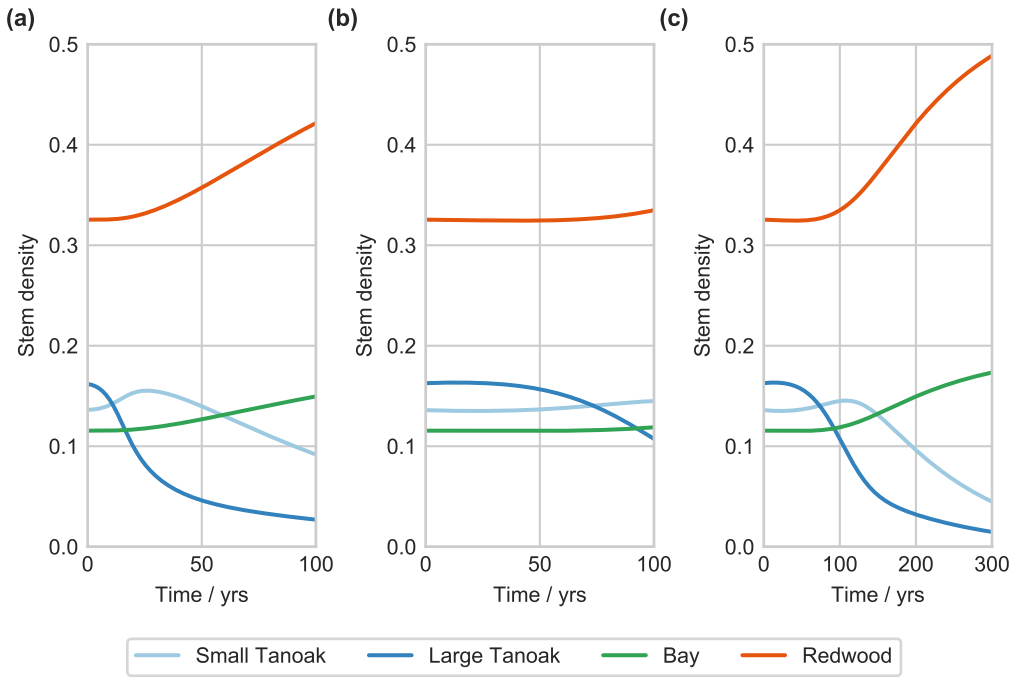

Figure 5.4: Change in dynamics when dispersal in the model is correctly parameterised, with realistic spore deposition. **(a)** shows the original dynamics using the implementation from Cobb et al. (2012). In **(b)** the dispersal kernel is corrected, using an exponential kernel to distribute spores across cells. This leads to significantly slower dynamics unless other parameters are rescaled. **(c)** shows the same corrected dynamics, but over a timescale of 300 years. This shows how the pattern of dynamics is unchanged by correcting the spore deposition, but the timescale is affected.

The age specific tanoak suppression weights, $w_{1,i}$, are fixed such that one quarter of the space occupied by tanoak is allocated to each age class. With the initial age distribution used in the BM code this corresponds to the youngest age class, the seedlings, suppressing recruitment the most. This is clearly unrealistic, since seedlings occupy the least space. We choose to make the more realistic assumption that suppression scales directly with basal area. Whilst this is relatively simplistic, it does at least ensure that smaller stems occupy less physical space. In the paper the initial age distribution, empty space and recruitment rates are set by running the model over 1000 years and finding parameters that give approximate dynamic equilibrium. In the BM code this leads to the highest recruitment rate in the smallest age class. In other words, the tanoak seedlings produce the most seeds. This is despite the paper claiming that recruitment rates are found under the condition that seed production increases with age.

To fix the initialisation, we solve for the dynamic equilibrium analytically. By setting infection rates to zero, we find the initial empty space and age distribution that gives dynamic equilibrium. We use the recruitment rates from Cobb et al. (2012), but set the

recruitment rate of the smallest age class to zero. More specifically, in the disease-free case Equations 5.1a for tanoak in a single cell become:

$$\dot{S}_{1,1} = \left(\sum_{i=1}^{4} b_{1,i} S_{1,i}\right) E - d_{1,1} S_{1,1} - a_1 S_{1,1} \tag{5.6a}$$

$$\dot{S}_{1,2} = a_1 S_{1,1} - d_{1,2} S_{1,2} - a_2 S_{1,2} \tag{5.6b}$$

$$\dot{S}_{1,3} = a_2 S_{1,2} - d_{1,3} S_{1,3} - a_3 S_{1,3} \tag{5.6c}$$

$$\dot{S}_{1,4} = a_3 S_{1,3} - d_{1,4} S_{1,4} \, . \tag{5.6d}$$

By setting the left hand side of these equations to zero we impose dynamic equilibrium within tanoak. With the condition that $b_{1,1}$ is zero, i.e. the recruitment rate of seedlings, these are solved to find the empty space and $S_{1,2}$, $S_{1,3}$, and $S_{1,3}$ in terms of $S_{1,1}$:

$$E = \frac{d_{1,1} + a_1}{b_{1,2} A_2 + b_{1,3} A_3 + b_{1,4} A_4} \tag{5.7a}$$

$$S_{1,2} = A_2 S_{1,1} \tag{5.7b}$$

$$S_{1,3} = A_3 S_{1,1} \tag{5.7c}$$

$$S_{1,4} = A_4 S_{1,1} \tag{5.7d}$$

where

$$A_2 = \frac{a_1}{a_2 + d_{1,2}} \tag{5.8a}$$

$$A_3 = \frac{a_2 A_2}{a_3 + d_{1,3}} \tag{5.8b}$$

$$A_4 = \frac{a_3 A_3}{d_{1,4}} \, . \tag{5.8c}$$

To find $S_{1,1}$ we fix the initial proportions of each host type, with p_1, p_2 and p_3 representing the proportion of hosts that are tanoak, bay or redwood respectively. This gives that:

$$S_{1,1} = \frac{p_1 (1 - E)}{1 + A_2 + A_3 + A_4} \, . \tag{5.9}$$

Finally, we fix the recruitment rates of bay and redwood such that those hosts are also in dynamic equilibrium. Namely:

$$b_2 = d_2 / E \tag{5.10a}$$

$$b_3 = d_3 / E \, . \tag{5.10b}$$

Reparameterisation

What effects do these errors have on the model dynamics and conclusions from the 2012 paper? The dominant effect is the change in rate of spatial spread due to the error in spore deposition. This significantly affects the timescale of SOD invasion, slowing it beyond what

is considered reasonable for SOD within stands (Figure 5.4). Whilst there is a lack of data covering spatial invasion rates at the scale of an individual forest stand, the timescales in Cobb et al. (2012) are broadly consistent with expectations. McPherson et al. (2010) found infection rates within stand of around 10 % per year in tanoak. Plots were 90 % infected after approximately 10 years, but the plots have an average size of 1,234 m^2 whereas the stand we are modelling is 200,000 m^2. In the original Cobb model the 500 m^2 cells adjacent to the source cell reach 90 % infection after approximately 8 years, consistent with the findings in McPherson et al. (2010). *P. ramorum* is known to spread locally via rain splash at distances of up to 20 m (Davidson et al., 2005), but can be carried much further in rare long distance dispersal events (Meentemeyer et al., 2011). For within stand spread we will consider only the short scale spread which corresponds to the invasion reaching the edge of the modelled 20 ha plot after ten to twenty years. This is again consistent with the timescale found in the original Cobb model, where infection reaches 5 % at the stand edge after 15 years.

With this confidence in the broad timescales in Cobb et al. (2012), we scale the infection rates in our reparameterised model with fixed dispersal and recruitment to match the original timescale of invasion found in Cobb et al. (2012). We use the time at which the population of small tanoak increases above the large tanoak population (crossover time) to define the rate of invasion. All infection rates in the corrected model are then scaled by the same factor so that relative rates are kept the same, but the rate of invasion matches that of the original Cobb model implementation (Figure 5.5). This ensures that the dynamics are correct, with a realistic kernel distribution, whilst matching the generally accepted spread rates for SOD. Our choice for defining the timescale using the crossover time was made for simplicity but is arbitrary. This structural change to forest composition is important ecologically. However, since the results show that the best fit is very close to the original at all times, other choices would not give very different results. We note that we also test parameter sensitivity in the next section.

A summary of all the parameter changes made based on the corrections described here is given in Table 5.2.

Table 5.2: Summary of changes made to Cobb model to correct unrealistic aspects of its parameterisation. The corrected spore deposition uses an exponential dispersal kernel with scale parameter, σ =10 m.

Issue	Parameters affected	Original	Corrected
Spore deposition	f_0	1.0	0.5
	f_1	1.0	$\propto \exp(-d_{ij}/\sigma)$ (also over other cells in the landscape)
Recruitment suppression weights	$w_{1,i}$	1/4 of space for each	Space proportional to basal area
Recruitment	$b_{1,1}$	Chosen to give dynamic equilibrium (≈ 0.055)	0.0
Initial conditions	$S_{1,i}(0)$	Chosen to give dynamic equilibrium	Equations 5.7–5.9

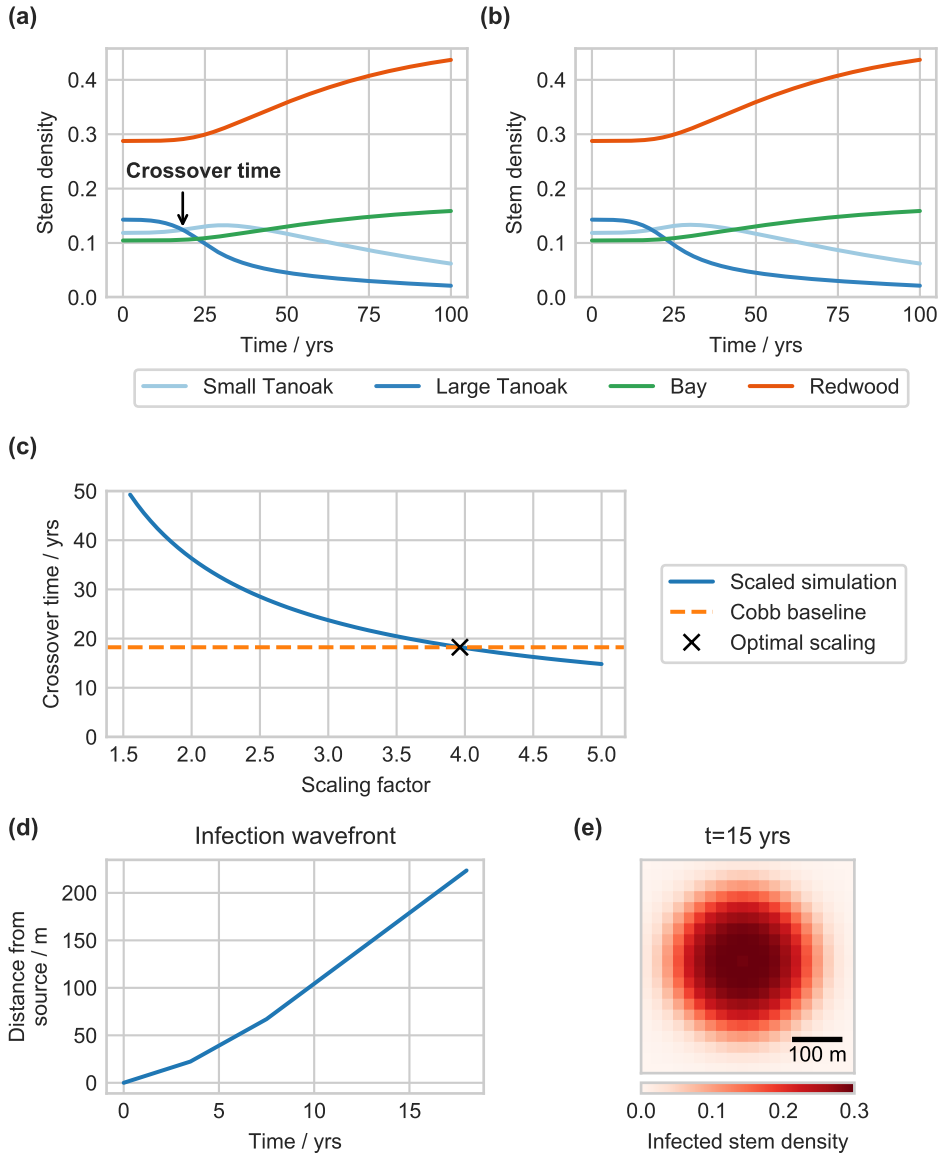

Figure 5.5: Effects of correcting parameterisation and spore deposition kernel, and matching timescales. **(a)** shows the original dynamics in Cobb et al. (2012), with the crossover time labelled. The crossover time is the time until the large tanoak population falls below the small tanoak population, and is the metric we have chosen to measure the timescale of the epidemic. In **(b)** we show the corrected and scaled dynamics, with the epidemic timescale matched to that in **(a)**. This is matched by scaling the infection rates, with the effect of this scaling factor on the crossover time shown in **(c)**. Using the scaled model and following disease introduction at $t = 0$ to the centre cell, **(d)** shows the distance from the source to the epidemic wavefront, here defined as the furthest cell with an infected stem density above 0.05. After a short transient (~5 years to 10 years), the epidemic wave travels at a constant speed. **(e)** shows the spatial distribution of infected hosts in the simulation model, 15 years after introduction.

5.2.3 Model sensitivity

Cobb et al. (2012) state that their model is 'designed to illustrate biotic factors affecting decline in tanoak populations invaded by *P. ramorum* and not to predict precise timescales for extinction'. To confirm that the Cobb model is still realistic after our reparameterisation, we test how sensitive the dynamics are to the exact parameters chosen. The aim here is to make sure that the general model dynamics, if not the exact numeric results, are unchanged for reasonable perturbations in the parameters. We will demonstrate that whilst the precise dynamics are complex, the overall trends in tanoak decline are robust to reparameterisation.

To test sensitivity we randomly perturb all parameters in Table 5.1 from the base parameter set. As described before, the empty space, tanoak suppression weights, initial conditions, and bay and redwood recruitment rates are calculated to give dynamic equilibrium for each reparameterisation. Using the perturbed parameter set, we run a forward simulation with no control to predict the future decline of tanoak. Each parameter is perturbed with a normally distributed error, with standard deviation equal to 25 % of the parameter value. A truncated normal distribution is used to ensure that parameters are not made negative, which would be biologically unrealistic. This process is carried out for 200 perturbed parameter sets. The results are shown in Figure 5.6.

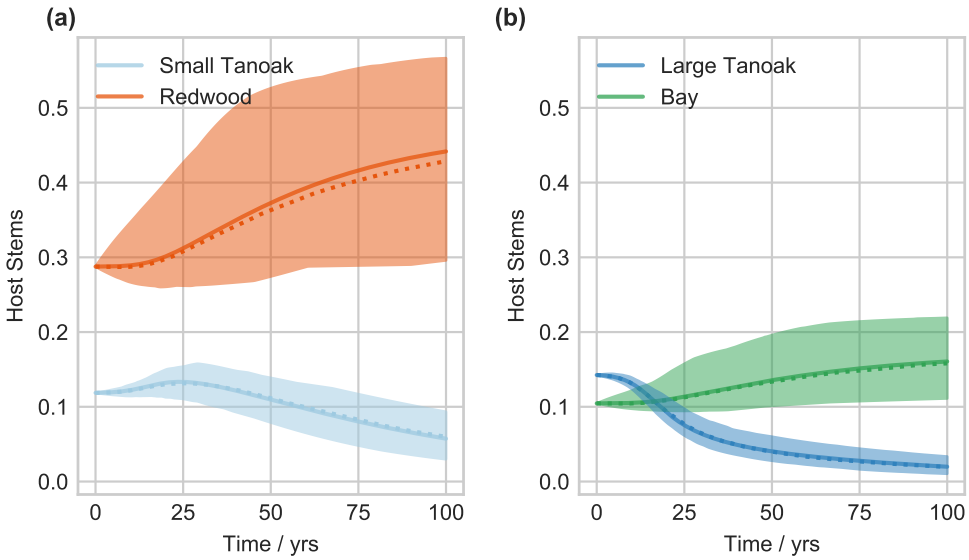

Figure 5.6: Sensitivity of the model dynamics to parameterisation. Solid lines show the dynamics with the baseline parameters. The dotted lines show the median from 200 random parameter perturbations, and the shaded areas show the 5th and 95th percentiles of the distribution. For clarity small tanoak and redwood are shown in **(a)**, and large tanoak and bay in **(b)**. In all cases there is significant decline of large tanoak. Note that whilst there is variation, the overall pattern of host compositional changes remains unchanged.

These results show that whilst there is variation in the predicted dynamics, the overall trends are highly consistent. Large tanoak is always predicted to decline significantly over the timescale considered, with smaller tanoak taking up a larger proportion of the tanoak age distribution. Bay and redwood are generally expected to increase in numbers, and certainly increase relative to the tanoak population. This sensitivity analysis confirms that the dynamics and main trend found in Cobb et al. (2012) are robust, despite the lack of data for accurate parameterisation.

5.3 Discussion

In this chapter we have corrected, developed and analysed a stand-level (20 ha) model of SOD-induced tanoak decline, based on a version originally formulated by Cobb et al. (2012). We have improved the realism of the model through changes in the transmission kernel, and reparameterisation of the initial conditions and seedling recruitment rates. *P. ramorum* spreads spatially by splash dispersal. In the corrected model, the spatial transmission kernel is implemented as a negative exponential, thin-tailed kernel, as is commonly used in the literature to model fine-scale processes such as splash dispersal (Skelsey and Garrett, 2013). Whilst other models use fat-tailed kernels to capture long-distance dispersal of *P. ramorum* (e.g. Meentemeyer et al., 2011; Harwood et al., 2009), these models make predictions at much larger scales than those considered here. In addition to this, as the stand-level model is deterministic, a fat-tailed kernel would lead to significantly higher—and unrealistic—rates of spatial spread unless the length scale of dispersal was highly restricted.

Our change to the kernel corrects the unrealistic spore deposition pattern in the original model. The corrections to the recruitment rates ensure that hosts in the seedling class do not produce seeds, as well as ensuring that seedlings correctly colonise less physical space than older trees. We have therefore fixed a pair of problems that were present in the original implementation. The necessity of these major corrections to the model could bring into question the validity of the rest of the model. However, the model is parameterised on data from a large and long-term network of plots—undoubtedly the most complete dataset for SOD impacts at the stand level. The parameter fits for the demographic processes are unchanged by the model corrections, and so the corrected model and parameterisation can be used with confidence. The model has also been used repeatedly in the literature (e.g. Cobb et al., 2017; Valachovic et al., 2017a; Cobb et al., 2019), and the timescales of tanoak decline are consistent with other studies (McPherson et al., 2010; Davidson et al., 2005), as discussed on page 96. When coupled with both our extensive testing and correcting of the original code, we can therefore be confident in the conclusions made and the underlying model dynamics. The particular parameterisation is of little importance, since we have demonstrated that the host composition dynamics are robust to parameter perturbations.

Since the dynamics of pathogen spread are largely insensitive to the exact parameterisation, and the improved transmission kernel only significantly affects the timescale of invasion, it could be asked why these corrections to the model are necessary at all. Firstly, the corrections are necessary for completeness. The model we have implemented has more realistic parameters than the original implementation, and by not imposing an artificially severe truncation in pathogen dispersal, results in more realistic patterns of spread. More importantly though, these corrections could have an effect on the dynamics under a disease management intervention. As we will explore in more detail in the following chapter, the optimal control strategy is likely to depend on the pattern of spread (Hyatt-Twynam et al., 2017) which is in turn affected by the choice of transmission kernel. The recruitment rate corrections are likely to affect the age distribution of tanoak after control strategies are carried out. It is therefore important that the underlying model is as accurate as possible before conclusions can be made about optimal control strategies.

Despite the large dataset underpinning the original publication of the model, there are data limitations that affect the model structure. There is very little data tracking the size progression of bay laurel and redwood trees. There is sufficient data to parameterise the four age classes within tanoak, but without additional data it is difficult to test whether the results are sensitive to the number of age classes implemented. The main difference between age classes is the higher pathogen-induced mortality for the oldest class. The change in this parameter through the four age classes suggests that fewer age classes would not be sufficient. It seems unlikely that additional classes would change this structure significantly, but there is little data to support this.

Other parameters in the model are somewhat arbitrary, for example the scale of dispersal. This is again due to a lack of data; here a lack of detailed infection timings at a fine spatial scale that would allow parameterisation of a small-scale transmission kernel. Other parameters from the original model could vary between plot locations, and the justification for the relative infection rates in Cobb et al. (2012) is minimal. We have demonstrated that these parameterisations have little impact on the overall dynamics, but they could have an impact on interpretation. Optimal control results must be interpreted taking these uncertainties into account, and tested for robustness, if the results are to be used in the real world. However, the current parameterisation does offer a plausible baseline for which the potential impact of OCT can be assessed, in time perhaps motivating the data collection necessary for further model parameterisation.

5.4 Conclusions

In this chapter we analysed and tested a model of SOD spread within a 20 ha forest stand, with a focus on the decline of overstorey tanoak. The model has been reparameterised to more realistically capture spatial spread, whilst maintaining the previously fitted rate of invasion. Having justified the underlying dynamics, the model will be used to answer questions about optimising disease control in the next chapter.

6 Optimising stand-level disease control to protect tanoak

6.1 Introduction

In the previous chapter we implemented a model of stand-level SOD dynamics. We adapted the Cobb model (Cobb et al., 2012) to incorporate realistic spatial dispersal properties, and reparameterised it to ensure the invasion timescales were still realistic. We now have a simulation model representative of SOD spread within a mixed species forest stand. In this chapter, we ask how control strategies can be optimised using this model, addressing the key question of how time-dependent controls should be deployed to best protect tanoak. As in Chapter 4, we will require an approximate model to optimise control since the simulation model is too complex. We will then use the open-loop and MPC frameworks we previously developed to integrate control strategies into the simulation model, and demonstrate the importance of continued surveillance for effective control.

Using our optimal control frameworks we will explore which control methods are most effective, and how deployment should vary over time. Most importantly, we will look at the effects of parameter and observational uncertainty on control efficacy, testing under which conditions MPC outperforms open-loop[1].

6.2 Stand-level disease control

In this section we describe the methodology for optimising control in the mixed stand model. We develop a non-spatial approximate model of the dynamics, incorporating thinning, roguing, and protectant controls. These are then optimised and applied using the open-loop and MPC frameworks. It is first necessary to quantify the effectiveness of a given strategy by defining the purpose of control in this system.

[1]All code for this chapter is available at `https://github.com/ehbussell/MixedStand`

6.2.1 Management objectives

As described in Section 5.1, it is important to retain tanoak in the forest stand due to its high cultural and ecological value. We therefore investigate strategies for specifically protecting tanoak, aiming to maintain a population of healthy overstorey tanoak trees in the stand over the medium- to long-term. Following the timescales in Cobb et al. (2012) we look at preventing decline of tanoak over a time horizon (T) of 100 years. Since the focus of the control is to ensure tanoak still exists in the future, we treat this as a terminal objective function. By this, we mean that the goal is to maximise the number of healthy overstorey tanoak trees at the end of the time horizon—i.e. after 100 years—rather than integrated over that time period.

This objective does not capture everything that is important to maintaining a healthy forest though. A control strategy that protects overstorey tanoak to the detriment of all other trees is clearly suboptimal, and this should be accounted for in the objective function. Trees in forests are important for wildlife habitats and food sources, recreational uses and carbon fixation (Swiecki and Bernhardt, 2013), and maintaining diversity ensures that these and other important ecosystem services are provided (Cadotte et al., 2011; Gamfeldt et al., 2013). Beyond this, diverse forests are more resilient to other disease threats (Keesing et al., 2010); there is little point to a control strategy that protects tanoak from SOD but makes the forest vulnerable to attack from another pathogen. The management goal must therefore capture a balance between protection of tanoak, and continued host diversity for provision of ecosystem services. The balance between these two objectives, however, will depend on the overall local management goals for the forest stand in question and economic valuation of the ecosystem services (Thompson et al., 2011).

There are many possible measures of diversity that could be used as part of the management objective. One measure that is very popular in the ecological literature is the Shannon index (Magurran, 2004, pp. 106–108). The Shannon index originates in information theory, based on the idea that diversity is a measure of the expected information content when observing the species of a random individual. The Shannon index H' is calculated using this equation:

$$H' = -\sum_i p_i \ln p_i \tag{6.1}$$

where p_i is the proportion of individuals in species i. It measures both species richness and species evenness, so is suited to our application in which the number of different species and evenness across species are both important. The Shannon index can be transformed into a value with more biological meaning: the effective number of species (ENS). This measures

how many equally common species would be required to achieve the same level of diversity, and for the Shannon index H' is given by:

$$\text{ENS} = \exp(H') \, . \tag{6.2}$$

We seek a control scheme that maintains ecosystem services. The relationship between biodiversity and ecosystem services is complex (Maes et al., 2012) but the two are most often positively correlated, particularly when the services provided are carbon sequestration and biomass production as in forestry (Paul et al., 2020). We therefore use the Shannon index measure of biodiversity as a proxy for ecosystem services. To maintain ecosystem services it is important that diversity is maintained throughout the simulation time. This ensures that further diversity is not lost in wildlife and other plant life during the epidemic. We therefore choose this to be an integrated metric, integrated over the full time horizon T. Overall then, our management objective is made up of a terminal term corresponding to preserving healthy large tanoak, and an integrated term to maintain maximum diversity across all times. The mathematical form of the objective is given by:

$$J = \gamma_1 \left(S_{1,3}(T) + S_{1,4}(T)\right) - \gamma_2 \left(\int_{t=0}^{T} \sum_i p_i \ln p_i \; dt \right) \tag{6.3}$$

where $S_{1,3}$ and $S_{1,4}$ are the densities of healthy tanoak in the third and fourth age classes respectively, and γ_1 and γ_2 are the weights associated with the tanoak retention and diversity conservation objectives respectively. There is an arbitrary choice in the balance between these two terms, which must be chosen by a policy maker or forest manager. In our case the weights are set such that in the disease free case, the contribution of the tanoak retention and biodiversity terms to the overall objective function are equal to 1 and 0.25 respectively. Whilst this is an arbitrary choice, we scan over the relative diversity benefit (γ_2) later in Section 6.3.2.

6.2.2 Control methods

Many different methods are recommended for controlling the spread of *P. ramorum* (Swiecki and Bernhardt, 2013). However, the methods can be grouped into three main classes: roguing, thinning and protecting. Roguing methods are based on finding and removing infected hosts, whereas thinning methods remove hosts regardless of infection status. Removal of hosts, either through thinning or roguing, is the only control that has been effective at the landscape scale (Hansen et al., 2008). Management recommendations made by the U.S. Forest Service highlight removal of the spreader species bay as very important for effective control (Swiecki and Bernhardt, 2013), but also recommend removal of infected hosts.

Whilst host removal is the only effective method at the landscape scale, protection methods could also be useful at the smaller scales of a single forest stand. These methods apply chemicals to uninfected trees to reduce their susceptibility to the disease. For SOD the main protectants used are phosphonates, that are approved for use on oak and tanoak species. The treatment only works as a preventative measure but it is recommended for protecting individual hosts (Lee et al., 2010). Reports of the effectiveness of phosphonate treatment vary, but most studies suggest it does slow infection (Swiecki and Bernhardt, 2013). Application by bark-spray or trunk injection is reported to be effective for up to around two years (Garbelotto and Schmidt, 2009), but evidence is lacking for the impact on host susceptibility. There are also conflicting reports about its efficacy, with some studies finding little effect of treatment (Kanaskie et al., 2011). Here we will assume a mild effect of 25 % reduction in susceptibility.

Roguing controls can be applied separately to infected small tanoak, large tanoak and bay laurel. The hosts are removed and do not resprout, consistent with application of a herbicide to the stump as is often recommended (Swiecki and Bernhardt, 2013). Thinning removes hosts of all infection status, and can be applied separately to small tanoak, large tanoak, bay and redwood. Protection can only be applied to small and large tanoaks, and only to susceptible hosts. These hosts are moved into new protected class with the same demographic dynamics (i.e. there is a protected class $P_{1,i}$ for each age class i of tanoak). The protected classes have reduced susceptibility (by 25 %) but return to the susceptible class at a rate of 0.5 year^{-1}. This corresponds to an average time of 2 years before protection wanes. Given the lack of evidence for the rate of loss of protection, we here use an exponentially distributed time until immunity wanes for simplicity. Table 6.1 summarises all the control methods and their effects.

Budget constraint

For each of the 9 controls in Table 6.1, we seek a time-varying control parameter $f_i(t)$ between zero and one, indicating the level of control i that minimises the management objective function. To model economic and logistic constraints we limit the total expenditure per unit time, where this is the product of the number of hosts controlled and the cost of that control method. The mathematical form of this constraint is given by:

$$\sum_i \left(f_i \eta_i X_i\right) c_i \leq B \tag{6.4}$$

where X_i is the stem density of the controlled hosts. For example, for roguing of small tanoak X_i would be $(I_{1,1} + I_{1,2})$. The term in brackets is therefore the rate of removal of hosts for each control. The relative cost of each control is given by c_i (measuring the cost

Table 6.1: Possible control methods implemented in the stand-level model. There are three main groups of control: roguing, thinning and protecting, and these can be targeted at different host groups. Approximate costs of each control are taken from Kovacs et al. (2011), but roguing costs are increased to account for additional costs of identification and removal of unstable diseased trees.

Control	**State changes**	**Rate** η_I / year^{-1}	**Cost** c_i / a.u.
Rogue small tanoak	$I_{1,1\text{–}2} \to \emptyset$	0.25	3000
Rogue large tanoak	$I_{1,3\text{–}4} \to \emptyset$	0.25	6000
Rogue bay	$I_2 \to \emptyset$	0.25	6000
Thin small tanoak	$\{S, I, P\}_{1,1\text{–}2} \to \emptyset$	1.0	250
Thin large tanoak	$\{S, I, P\}_{1,3\text{–}4} \to \emptyset$	1.0	500
Thin bay	$\{S, I\}_2 \to \emptyset$	1.0	500
Thin redwood	$S_3 \to \emptyset$	1.0	500
Protect small tanoak	$S_{1,1\text{–}2} \to P_{1,1\text{–}2}$	0.25	200
Protect large tanoak	$S_{1,3\text{–}4} \to P_{1,3\text{–}4}$	0.25	200

of treatment per unit host density in arbitrary units) and the maximum budget is given by B. Whilst the costs are chosen somewhat arbitrarily because of a lack of data, the scales are informed by the results of Kovacs et al. (2011). We include higher costs for roguing to capture the additional costs with identification and removal of unstable infected trees.

6.2.3 Approximate model

Optimisation of the chosen management objective using all nine time-dependent controls is computationally infeasible using the full spatial model. To allow progress we use an approximate model and lift control results from this simpler optimisation back to the simulation model, using the methods described in Chapter 4. Here we choose to make the approximate model non-spatial, so as to significantly reduce the state-space for optimisation. Further approximations could be possible, for example grouping together the age classes within the small and large tanoak groups. To ensure that we can lift demographic parameters directly from the simulation model however, we only change the spatial structure of the model. The approximate model therefore assumes that all hosts in the forest stand are well-mixed.

As all other features of the model are retained, the form of the equations is very similar to that of the simulation model (Equations 5.1a) so we will not repeat them here. The only difference is that in the approximate model the dependence on cell (x) of the states and empty space (E) has been dropped. The form of the force of infection terms is simpler since infection no longer spreads between cells. The infection rates in the non-spatial approximate

model cannot be lifted from the simulation model, since the approximate model now assumes infection comes from all infected hosts in the stand rather than just those in the immediate spatial vicinity. The force of infection terms in the approximate model into tanoak and bay laurel are therefore:

$$\tilde{\Lambda}_{1,i} = \tilde{\beta}_{1,i} \sum_{j=1}^{4} I_{1,j} + \tilde{\beta}_{12} I_2 \tag{6.5a}$$

$$\tilde{\Lambda}_{2} = \tilde{\beta}_{21} \sum_{j=1}^{4} I_{1,j} + \tilde{\beta}_{2} I_2 \tag{6.5b}$$

where the sum is over tanoak age classes, and $\tilde{\beta}$ indicates infection parameters that need to be fitted to the simulation model.

Fitting infection rates

The seven infection rates $\tilde{\beta}$ are the only parameters that must be fitted to the simulation model. The approximate model cannot capture the heterogeneous mixing present in the simulation model, however, the approximated dynamics may be accurate enough to give effective control strategies when optimised. We use the method of least squares to match the simulation and approximate models, as both models are deterministic. To fit the parameters, the simulation model is used to run a single trajectory with no control interventions. For the fitting process we use the same initial conditions as described in the previous chapter, with infection seeded in the centre of a 20 by 20 grid of cells. The disease progress curves of this simulation realisation are then used as the baseline for fitting the approximate model. For a trial set of $\tilde{\beta}$ parameters and an approximate model trajectory, we calculate the sum of squares as the sum of squared deviations between the simulation and approximate disease progress curves for each age class of tanoak, and for bay, at time points throughout the trajectory. The $\tilde{\beta}$ parameters are then optimised by minimising this total summed squared error (SSE). For a set of time points t_i, and where approximate model states are signified with a tilde, the equation for SSE is given by:

$$\text{SSE} = \sum_{i} \left[\sum_{j=1}^{4} \left(I_{1,j}(t_i) - \tilde{I}_{1,j}(t_i) \right)^2 + \left(I_2(t_i) - \tilde{I}_2(t_i) \right)^2 \right] \tag{6.6}$$

where the dependence on cell in the simulation terms has been dropped to indicate an average over all cells in the landscape, for example:

$$I_{1,j}(t) = \frac{\sum_x I_{1,j,x}(t)}{N_{\text{cells}}} \,. \tag{6.7}$$

An average is used so that the approximate model tracks stem density in the same units as the simulation model.

In the simulation model, infectious pressure is dominated by sporulation from bay laurel. This makes estimation of all 'within-tanoak' infection rates ($\beta_{1,i}$) difficult, as from the simulation data they are individually unidentifiable. We therefore use a two stage fitting process. For the first stage, all infection rates in the simulation model related to bay (β_2, β_{12}, and β_{21}) are set to zero. This makes bay epidemiologically inactive, but maintains the same demographic dynamics. The simulation model is run using these parameters, and the SSE is minimised to find the within tanoak-infection rates.

In the second fitting stage all infection rates are fitted, with bay epidemiologically active again. The within-tanoak rates relative to $\beta_{1,1}$ from the first stage are used as a constraint to ensure the identifiability of these rates. This means a single within-tanoak rate is fitted in stage 2, with all other within-tanoak rates fixed relative to this using the results from stage 1. This stage also fits the bay infection rate, and the cross-species infection rates.

As can be seen in Figure 6.1, despite lacking any spatial component, the approximate model can very closely capture the uncontrolled dynamics of the simulation model. However, the approximate model should also fit as accurately as possible when control strategies are introduced. In Figure 6.2, the fit of the approximate model is tested under constant control strategies using fixed control rates. It is clear that roguing at the same rate is more effective in the approximate model. This is because of the difference in mixing between the approximate and simulation models. The effect is small for thinning and protecting, but the same level of roguing in the approximate and simulation models gives very different dynamics.

Empirical parameterisation of control

Roguing is less effective in the simulation model because of an imposed spatial structure in the non-spatial control strategy. Roguing in the simulation model removes infected hosts from the core and edge of the spreading epidemic. Removal of hosts from the core has little effect on the rate of epidemic spread, since they are not near the wavefront. In the non-spatial model however, all hosts are well-mixed, so removal of infected hosts has a larger effect.

As a simple correction for this difference in roguing effectiveness, we investigate a simple scaling of the roguing rate in the approximate model. To test the plausibility of a single scaling rate for all approximate roguing controls, both models are run with constant roguing strategies. Roguing of small tanoak, large tanoak and bay are all set to occur at the same rate for the whole simulation, and this rate is varied between simulations. In the approximate model, the control rate is scaled by a single parameter which is also varied, and we analyse the difference in the final number of healthy large tanoak after 100 years between the

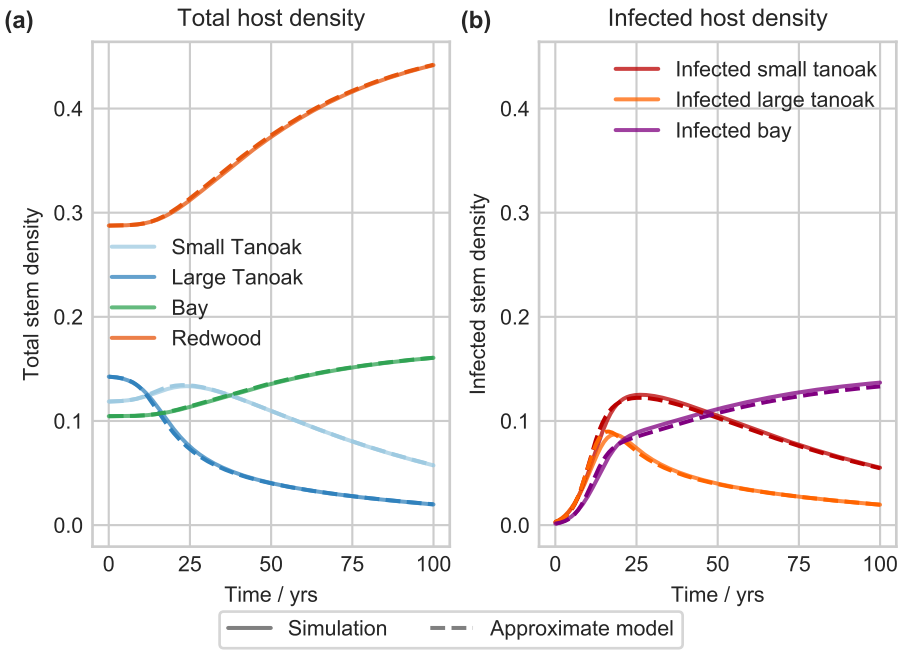

Figure 6.1: Fitting of the approximate model to match the output of the simulation model. **(a)** shows the overall stem density for each species class, with the dashed line showing the fitted approximate model. The fit is carried out by matching the disease progress curves as shown in **(b)**.

simulation and approximation (Figure 6.3). To minimise the deviation across a range of control rates, the value of the scaling parameter is optimised. We optimise the deviation in the number of large healthy tanoak, since this is the primary objective of the control and must therefore be captured as accurately as possible. The optimisation minimised the sum of squared errors (SSE) over the range of control rates.

The results in Figure 6.3 show that a single scaling factor can largely eliminate the deviation under constant roguing strategies. We do not expect this scaling to ensure that the approximate model is always closely aligned to the simulation model, particularly once control strategies are time-varying. The approximate model simply cannot capture the heterogeneities in host mixing. This scaling does however go some way to ensuring that control strategies from the approximate model perform well on the simulation model.

6.2.4 Control frameworks and lifting

In Chapter 4 we introduced the open-loop and MPC control frameworks. These same frameworks are tested here. To aid convergence, the controls inputs f_i are constrained in the optimisation to be held constant over 5 year stages. The main difference from Chapter 4 in applying the control frameworks is how controls are lifted from the approximate model to the simulation model. The budget constraint is a mixed constraint that couples the control

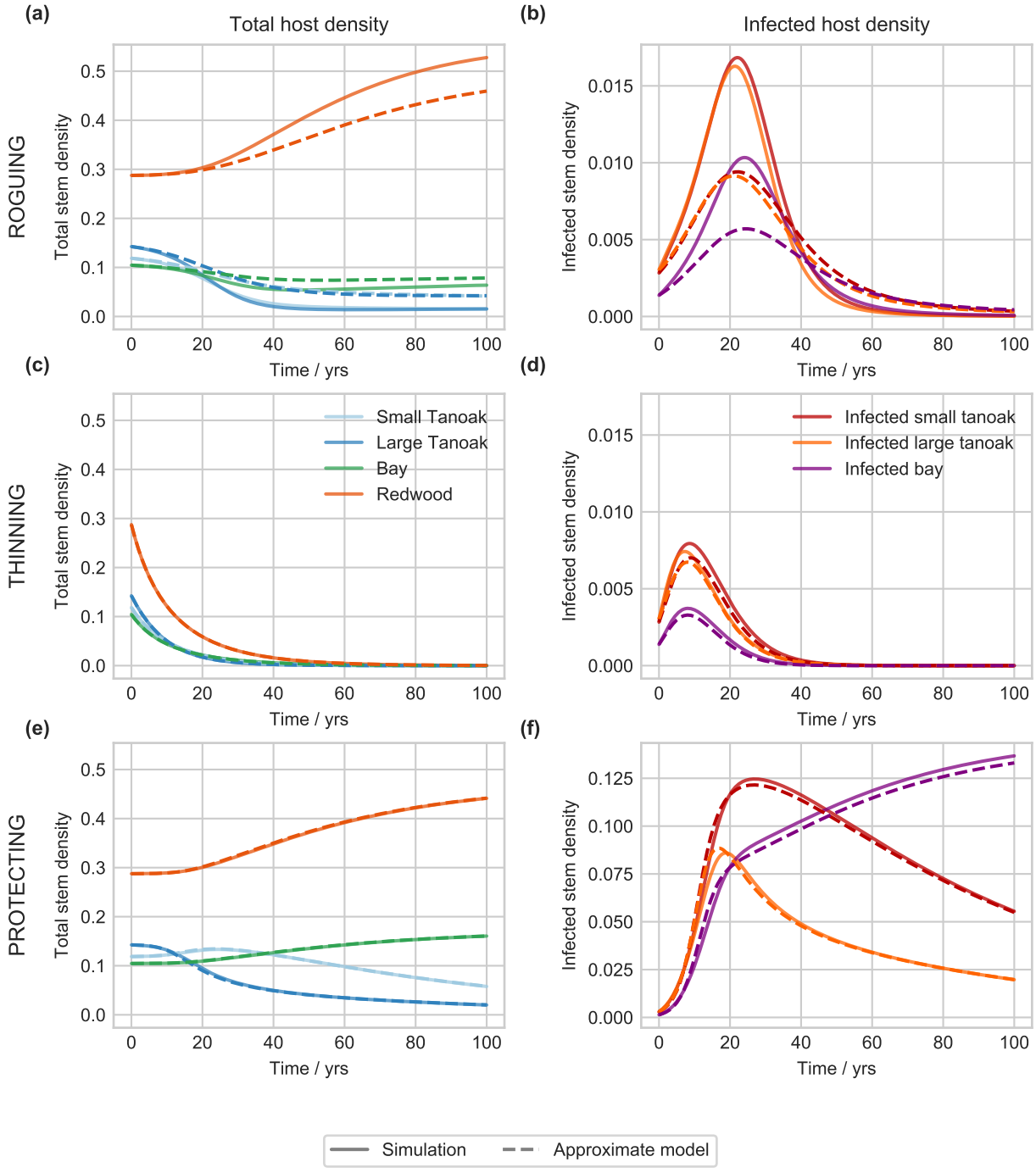

Figure 6.2: Testing of approximate model under constant control strategies. The rows from top to bottom show dynamics under constant roguing, thinning, and protecting strategies. The left plots show overall host dynamics, and the right plots show the infected host dynamics. The roguing and protecting strategies control at the maximum rates from Table 6.1, whereas the thinning strategy controls at 10 % of the maximum rate. The approximate model fits well under constant thinning and protecting strategies, but less well under a constant roguing strategy.

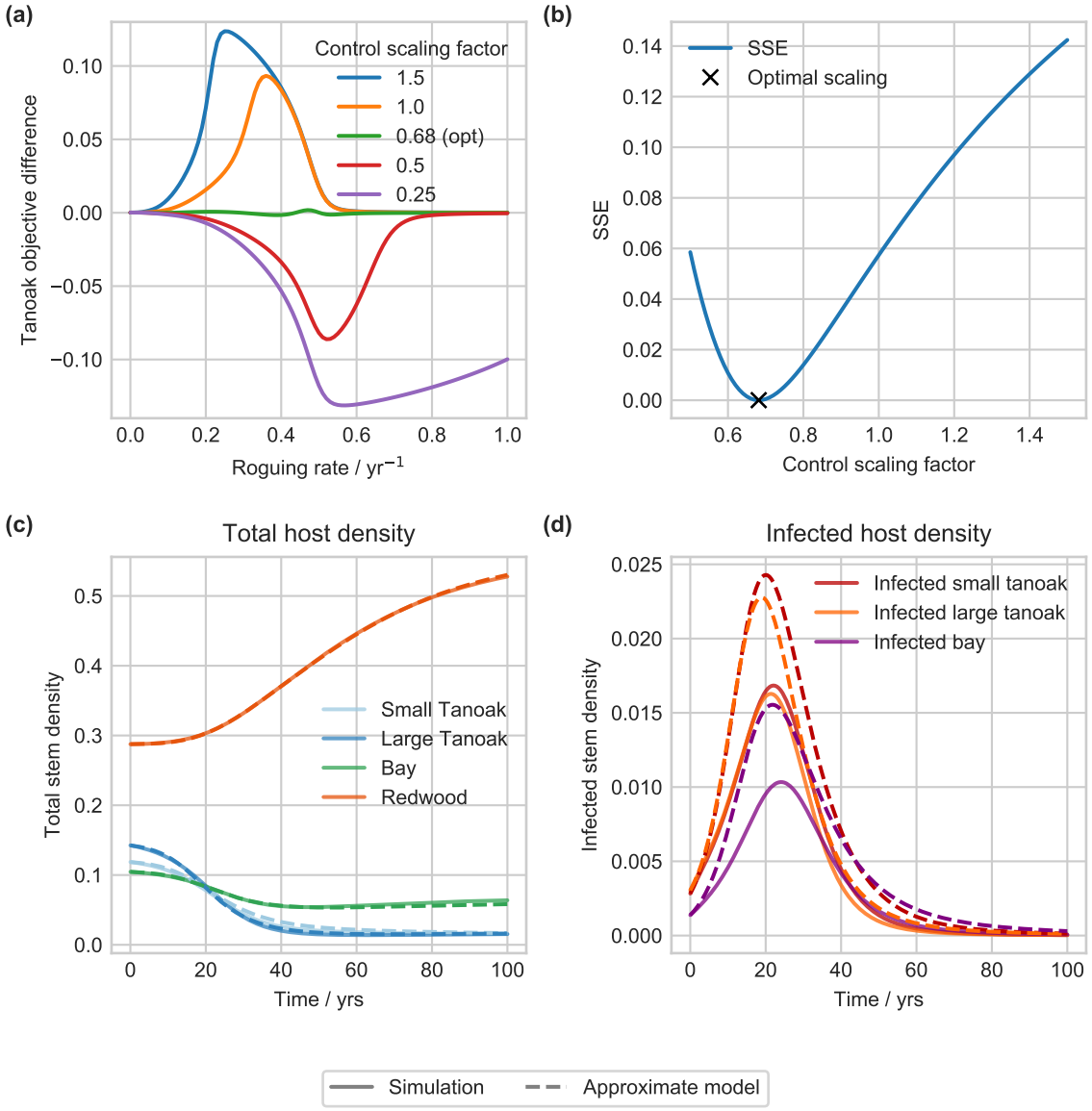

Figure 6.3: Empirical scaling of approximate model control rates to match simulation output. **(a)** shows the difference in the final number of large healthy tanoak as a function of the constant roguing rate, for a number of control rate scaling factors. The optimal value minimises the sum of squared errors (SSE) over all rates, as shown in **(b)**, and is found to be 0.682 (3 s.f.). **(c)** and **(d)** show the dynamics with the newly scaled approximate control rates, under a constant roguing strategy. The approximate model now fits the overall host dynamics well, but to do this slightly overestimates the level of infection.

inputs (f_i) with the state of the system (X_i). When the control inputs are lifted to the simulation model, there is no guarantee of the states being exactly the same, and so the expenditure by the control will not be the same. This can mean that direct lifting of the control inputs will lead to the budget being exceeded. When the budget is exceeded, the control inputs are multiplied by a factor to reduce the overall expenditure to meet the budget constraint. This avoids imposing a priority amongst control methods which would have

to be chosen arbitrarily. In mathematical terms, the corrected control inputs (f_i') from the approximate model depend on the state in the simulation (X_i) in the following way:

$$f_i' = f_i \frac{B}{\sum_j (f_j \eta_j X_j c_j)} \,. \tag{6.8}$$

This is only the case when the budget is exceeded, otherwise the control inputs f_i are used directly. We do not correct for under-allocation of resources, i.e. control inputs not spending the entire budget, since this could lead to extra control not accounted for in the approximate model. In practice correcting for this would lead to extra resources allocated to thinning, and hence removal of more healthy trees than is necessary.

6.3 Results

6.3.1 Optimal strategies

We first show the optimal strategies found using OCT on the approximate model. The strategies are lifted to the simulation model using the methods described above for both open-loop and MPC frameworks. Note that here control starts 6 years after the initial conditions used for fitting. This is to allow time for the disease to become established, making detection of the epidemic more likely. The initial conditions are therefore found by running the simulation model with no control for 6 years.

Open-loop strategies

The open-loop framework carries out control optimisation on the approximate model, and this strategy is then lifted to the simulation for the full duration of the epidemic. The strategy found using OCT when applied to the simulation is shown in Figure 6.4(a). The strategy focusses on thinning of bay, followed by thinning of redwood, early in the epidemic. Roguing is carried out throughout the epidemic but at a rate which increases towards the end. This is because late in the simulation infection re-emerges, so roguing uses more of the budget than anticipated by the approximate model. There are more infected hosts in the simulation, so it costs more to remove them. This can also be seen in Figure 6.4(b), where towards the end of the epidemic in the simulation, there is a decline in the tanoak numbers. This is not captured by the approximate model that anticipates tanoak numbers continuing to increase.

The open-loop strategy carries out a large amount of thinning early in the epidemic. As can be seen in Figure 6.5(a), this severely impacts the stand diversity. Over the course of the epidemic though, the diversity in the simulation returns to close to its initial value. In the approximate model diversity is not expected to recover as much. This reduction in diversity

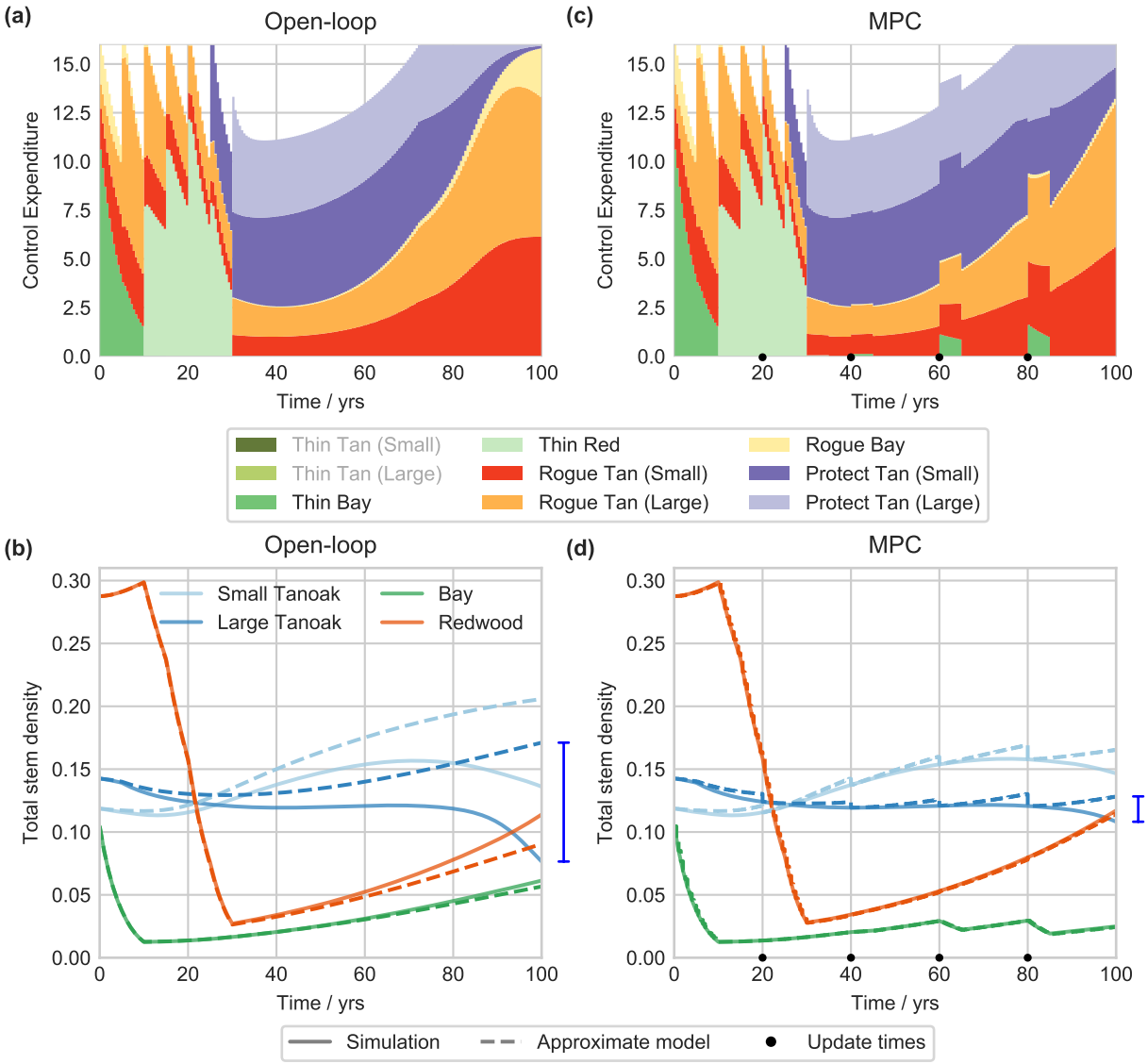

Figure 6.4: Optimal allocation of control resources using the approximate model. **(a)** shows the allocation over time to each control method using the open-loop framework. Control proportions (f_i) are fixed over 5 year intervals, but as the number of hosts changes during each interval, this leads to variable expenditure over the 5 years. Greyed out control methods in the legend are not used in either strategy. Heavy thinning of bay and then redwood is carried out early in the epidemic, with protectant controls only used once resource intensive thinning has been completed. In **(b)** the corresponding host dynamics are shown for both the simulation and approximate models. The approximation degrades towards the end of the epidemic, leading to unanticipated tanoak decline in the simulation. The blue bar highlights the difference between the approximate and simulation models in the density of large tanoak. **(c)** shows the MPC resource allocation, updated every 20 years, with a similar pattern to the open-loop strategy. However, additional thinning is carried out after each update later in the epidemic to manage the bay population. The corresponding host dynamics in **(d)** show that MPC repeatedly resets the approximate model trajectory, allowing more informed control decisions. This minimises the tanoak decline seen using open-loop, and gives a much lower error in the estimate of large tanoak stem density (blue bar).

is balanced though, by the retention of tanoak in the stand. As shown in Figure 6.5(b), the approximate model expects the control strategy to restore the full tanoak population, and increase it above its initial value. It is able to do this in the approximate model as disease is eliminated early, and healthy tanoak can grow into the space made available by the thinning of bay and redwood. This does not happen in the simulation model however, where the late re-emergence results in rapid decline of large healthy tanoak over the final 20 years. Despite this, the open-loop strategy shows a significant improvement over the dynamics under no control intervention. The strategy slows the spread of disease, keeping tanoak in the forest population for an additional 80 years.

MPC strategies

In the MPC framework the approximate and simulation models are run concurrently, with the approximate model reset to match the simulation and re-optimised at regular update steps. These updated controls are then lifted to the simulation model going forward until the next update time. We first test the MPC framework with updates every 20 years. For the first 20 years the control is exactly the same as the open-loop strategy (Figure 6.4(c)), since it is lifted from the same optimisation. After this, though, updates result in increased levels of thinning. In particular, after the updates at 60 and 80 years there is significant thinning of bay. This is to stem increased infection in bay, and keep the late disease re-emergence under control. As can be seen in Figure 6.4(d), the updates ensure the approximate model dynamics more closely match the simulation, compared to under open-loop control. The extra thinning of bay slows the disease re-emergence and there is less decline in tanoak numbers.

Figure 6.5(a) and (b) show the diversity and tanoak retention under MPC. MPC is more damaging to diversity than open-loop. Although the approximate model incorrectly still expects tanoak numbers to increase after each update, keeping the model close to the simulation dynamics does improve control. The MPC framework retains approximately 60 % of the original tanoak population after 100 years. Open-loop only retains 15 %. Figure 6.5(c) shows that both control frameworks significantly reduce the amount of empty space in the forest stand by removing trees (note that empty space does not map directly to physical space since it depends on seedling suppression, for example by light and water requirements). To control the disease the tree density is approximately halved, but this is still consistent with observed stand densities (Cobb et al., 2012). Figure 6.5(d) compares the objective function values for open-loop, MPC, no control, and without any disease. We can see that MPC lowers diversity performance slightly compared to open-loop, but in doing this retains significantly more tanoak.

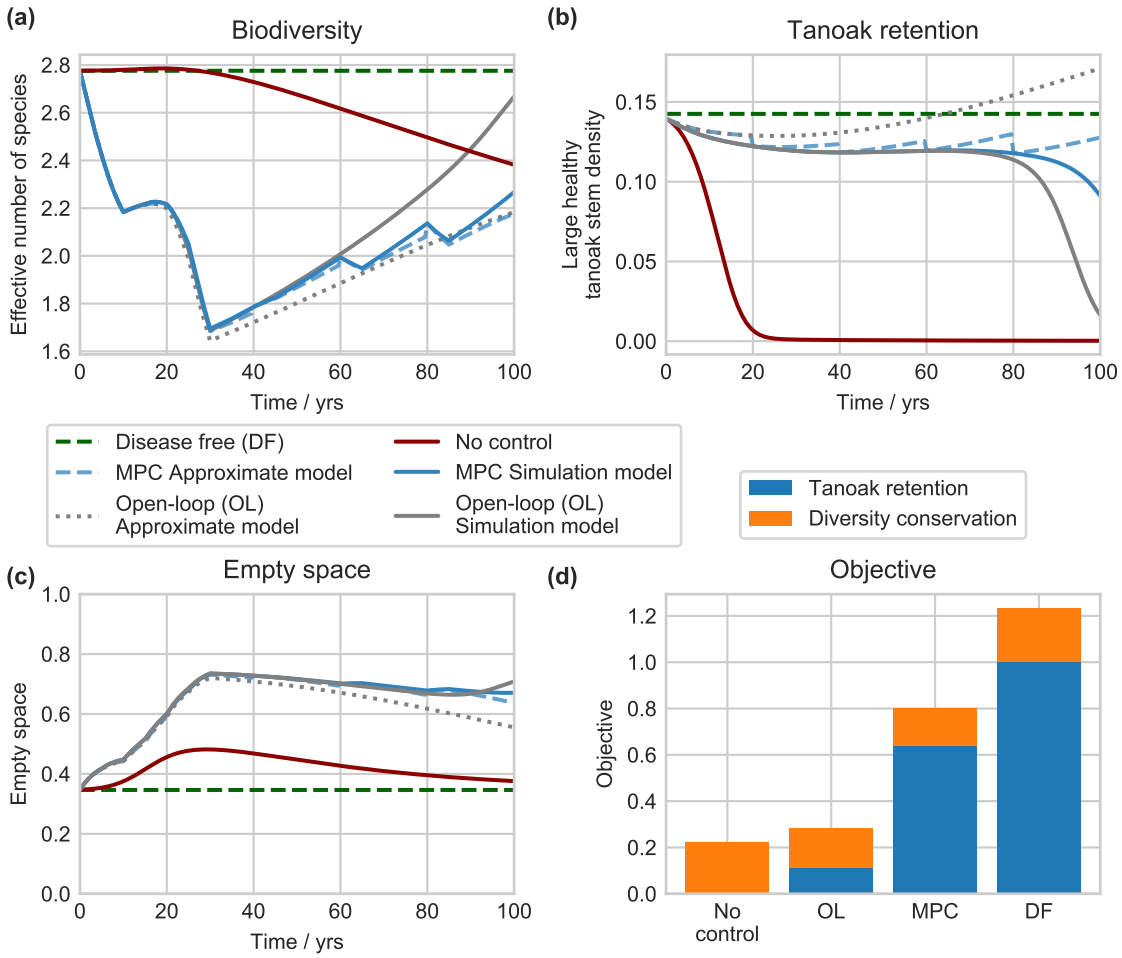

Figure 6.5: Comparing performance of the open-loop and MPC frameworks. **(a)** shows the effect of control on the biodiversity in the forest stand, presented as the effective number of equally-common species. Any control—be it optimised via open-loop or MPC—is damaging to diversity, but MPC in fact has a slightly larger impact. The simulation and approximate model dynamics are shown, with the approximate model resetting every 20 years under MPC. **(b)** shows the stem density of healthy large tanoak over time, with significant decline under no control. The open-loop strategy slows tanoak decline, but MPC is more effective. **(c)** shows the changes to empty space in the forest stand caused by the control. In **(d)** the overall performance of the strategies as measured by the objective function is shown. MPC is slightly more damaging to diversity than open-loop, but this is balanced by retaining significantly more healthy tanoak.

Since the updates in MPC improve control, an important question is how often to update the approximate model. Figure 6.6(a) shows the effect of changing the update period on the objective. We can see that as updates are made more frequent, control performance generally improves. This is because the approximate model can more closely match the simulation and hence appropriate control decisions can be made. There is, however, a dip in performance at update periods of around 50 years. This is due to the precise timing of the updates. The late disease re-emergence occurs after around 80 years. Update periods of around 50 years will not update close to this outbreak, and so the MPC framework cannot respond to that unexpected increase in infection. This results in ineffective control.

Figures 6.6(b) and (c) show the MPC control for update periods of 5 and 100 years respectively. The low frequency update corresponds to open-loop control. We can see that the main difference as update frequency increases, is additional continued thinning of bay. This results in less roguing being required later in the epidemic.

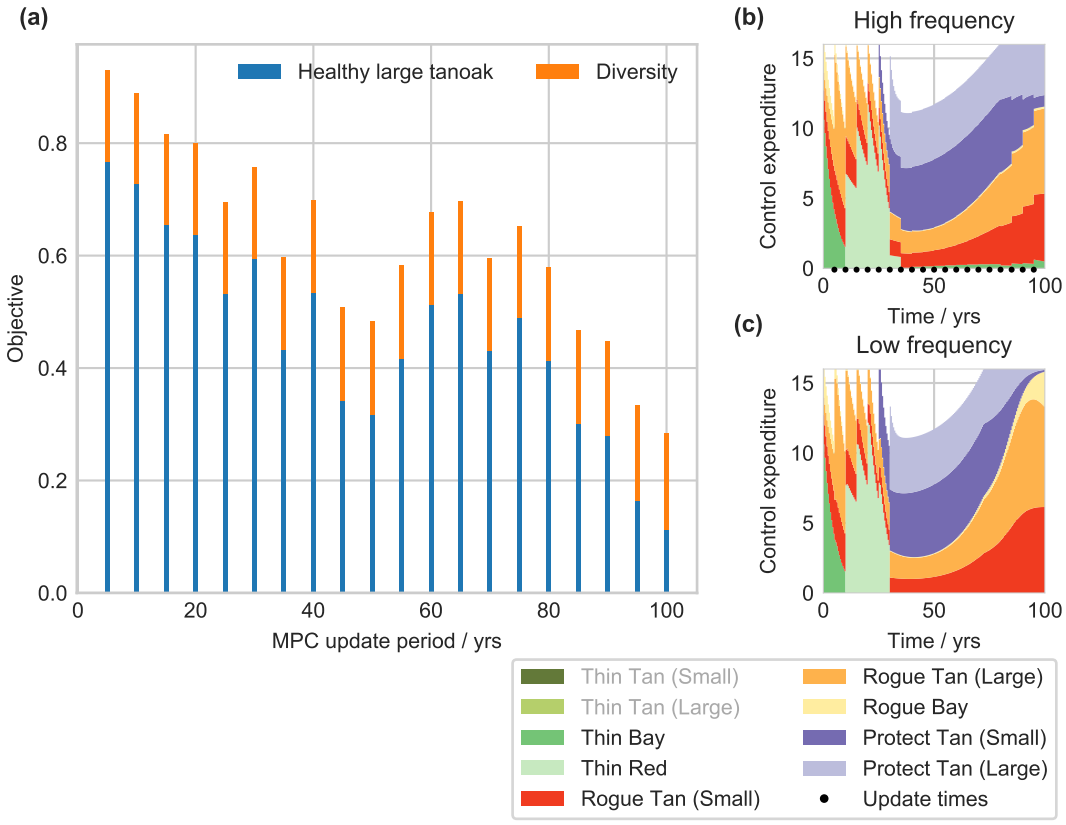

Figure 6.6: Effect of the MPC update period on control performance. **(a)** shows the objective value as a function of how often the MPC re-optimises control. Reducing the time until the next update generally improves control, although update periods of around 50 years perform worse than expected. This is because these periods do not update close to the late outbreak, and therefore miss the unexpected increase in infection. **(b)** and **(c)** show the control allocations for update periods of 5 and 100 years respectively. The low frequency control here corresponds to open-loop control. The high frequency control results in more continued thinning than in the low frequency control.

6.3.2 Robust control

In this section we test the robustness of the results so far. In particular, we analyse how parameter choices affect the control strategies selected by the open-loop and MPC frameworks, and their performance. We then investigate how the open-loop and MPC frameworks handle uncertainty in parameters and imperfect state estimation, i.e. not carrying out perfect sampling of the forest stand at the MPC update steps. These robustness analyses are important since the data available for fitting within-stand dynamics is limited.

Budget sensitivity

First we analyse the effect of the budget constraint. The maximum expenditure was chosen arbitrarily, so how much effect does it have on the optimal control strategy? Figure 6.7 shows that in general the control strategy does not depend strongly on the budget. Across all budgets apart from the very smallest, largely the same set of control methods is used. As the budget increases, open-loop can allocate more resources to thinning, and with improved control both frameworks increase resources allocated to protection. It can also be seen that performance generally increases with increasing budget as might be expected. Kinks in this trend are due to some levels of control leading to a closer fit between the simulation and approximation, and hence improved control. At very high budgets control performance starts to degrade with budget. This is because in these cases the control is very effective in the approximate model, and so less control is carried out resulting in worse control in the simulation. As explained on page 112, under-allocation of resources is not corrected for, although in this case the problem arises because a higher proportion of the available resources is allocated to protective controls. This results in less resource allocated to roguing, and hence less effective control.

Interestingly, at very low budgets—where neither control framework is very effective—we see thinning of small tanoak. In these cases there are not enough resources to control the disease effectively with roguing of tanoak alone. The next best option in this case is thinning of small tanoak, which reduces infection but also reduces numbers of healthy tanoak. This is therefore only chosen as a control method when roguing alone is not enough, and only in small amounts.

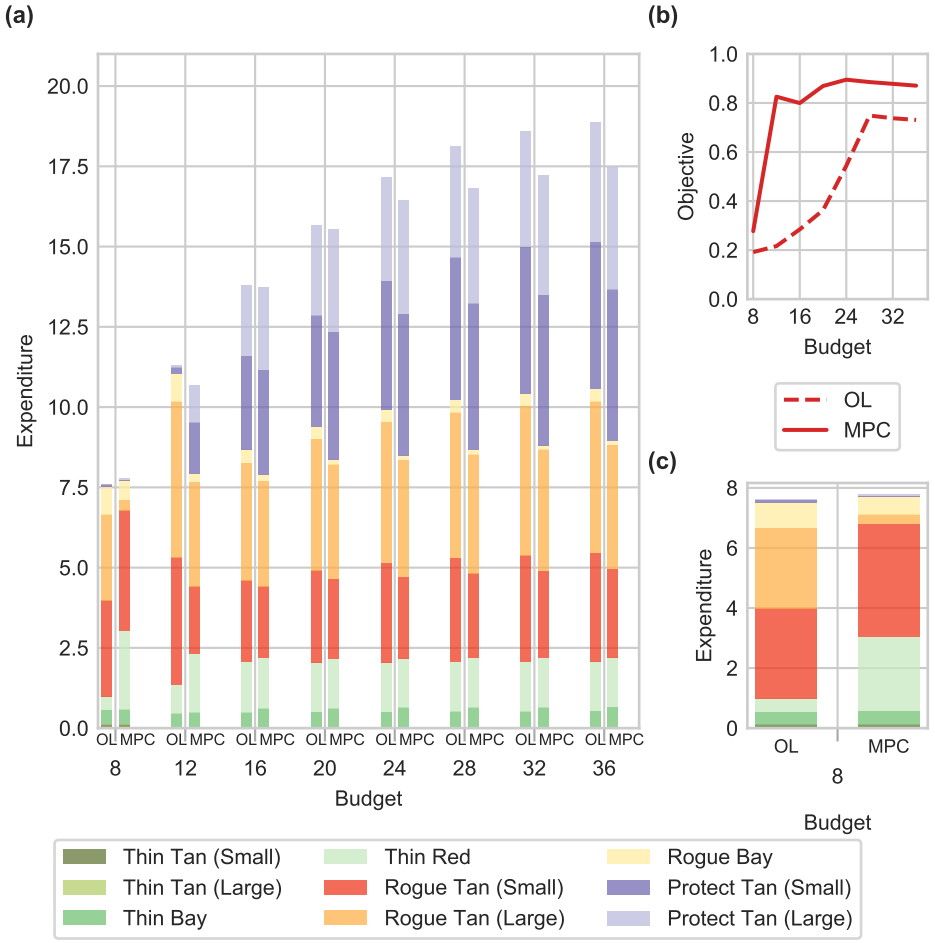

Figure 6.7: Effect of resource budget on control. **(a)** shows overall allocation of resources to each control method as a function of the maximum budget. Left hand bars show the results for open-loop, right hand bars for MPC. We can see that MPC generally allocates more to thinning, and requires less roguing. **(b)** shows the corresponding simulation objectives for the open-loop and MPC frameworks. Control generally improves as the budget increases. Interestingly, at the lowest budget resources are allocated to thinning of small tanoak, as shown in **(c)**.

Diversity benefit sensitivity

Next we test the sensitivity of the control strategies to the relative importance of biodiversity conservation and tanoak retention. By default, the diversity benefit is chosen such that in the disease free case, the diversity term will be 25 % of the tanoak retention term in the objective function (as described in Section 6.2.1). In Figure 6.8 the control allocations and performance are shown, scanning over relative diversity benefit from 0 to 100 %. It can be seen that, unsurprisingly, the best protection of tanoak, using both open-loop and MPC frameworks, is possible when there is no diversity benefit. As the diversity benefit increases, open-loop allocates fewer resources to thinning, and performance degrades. MPC on the

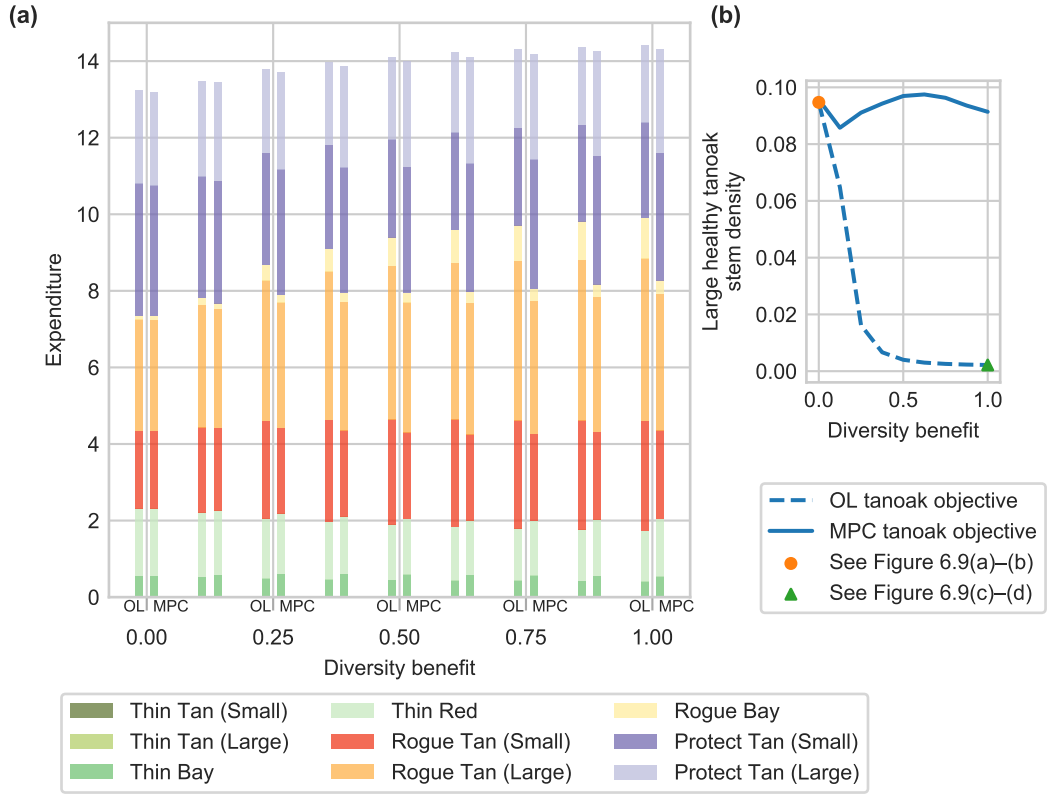

Figure 6.8: Changes in control strategy and performance as the relative benefit of diversity is changed. **(a)** shows overall allocation of resources to each control method as the diversity benefit is varied. **(b)** shows the stem density of large healthy tanoak retained in the stand. Low diversity benefits result in more thinning, but improved retention of tanoak in the open-loop case. MPC is able to respond to changes in diversity and retain tanoak for all values of the diversity benefit.

other hand, is able to adapt and maintains high levels of tanoak protection through to the highest diversity benefits.

In Figure 6.9 we compare the control strategies and host dynamics with no diversity benefit, and the highest level of benefit. We can see that when there is no benefit to diversity protection, high levels of thinning are carried out that remove all bay and redwood trees. This leads to very effective disease control but for many forest managers and conservationists this would be an unacceptable cost to slow decline of a single important species Noss (2000). This demonstrates why diversity should be accounted for in the management objective. At the highest diversity benefit all species are retained in the system, but under open-loop control the disease is not controlled well towards the end of the epidemic. The same patterns are seen with the MPC framework, but the updates allow tanoak retention whilst also preserving diversity.

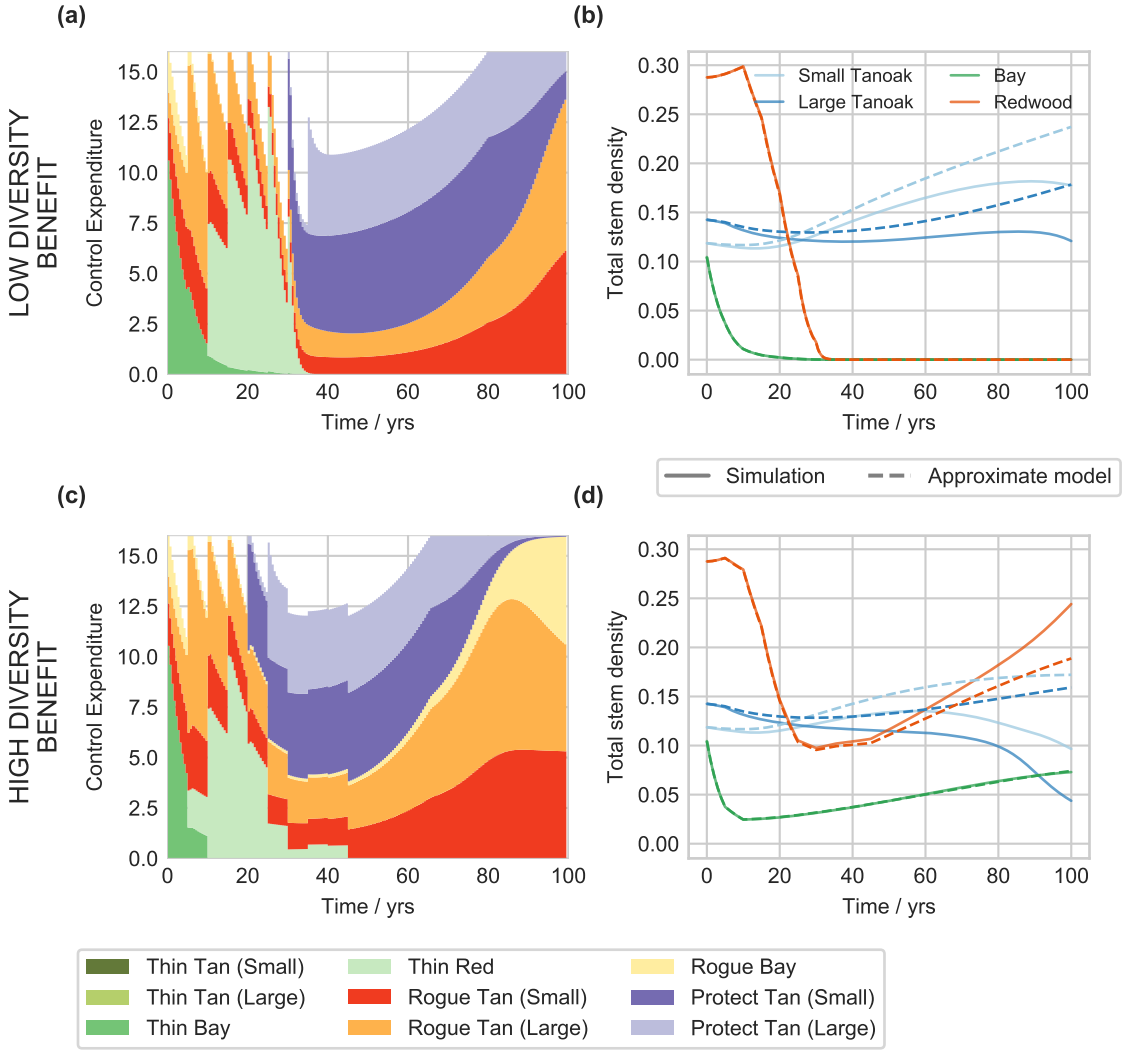

Figure 6.9: Open-loop control and host dynamics are shown for low diversity benefit (**(a)** and **(b)**) and high diversity benefit (**(c)** and **(d)**). Additional thinning is carried out when diversity benefits are low, resulting in complete removal of all bay and redwood trees. Similar strategies are seen using the MPC framework.

Parameter sensitivity

We next analyse the sensitivity of the control strategies to the underlying model parameterisation. As in Section 5.2.3, all parameters from Table 5.1 were randomly perturbed using a truncated normally distributed error with standard deviation of 25 %. For each perturbed parameter set the approximate model was re-fitted, and control optimised using both the open-loop and MPC frameworks. The control strategies are compared across parameter sets by visualising the proportion of the budget that is allocated to each control class (thinning, roguing and protecting) over time. By sorting the parameter sets according

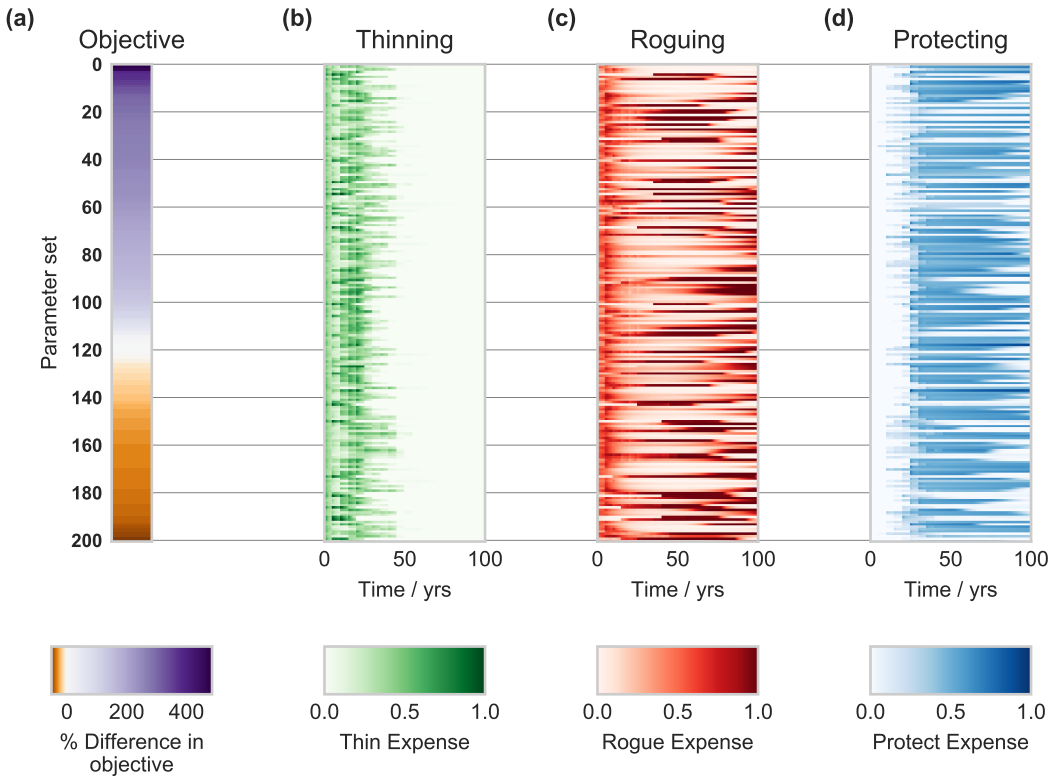

Figure 6.10: Sensitivity of the open-loop control strategy to parameterisation. Each row corresponds to a single random parameter perturbation, with the rows ordered by the objective value under open-loop control. **(a)** shows the change in objective from the baseline parameterisation, with negative values shown in orange signalling worse epidemics. **(b)**, **(c)** and **(d)** show the allocation of control resources over time to thinning, roguing and protecting control methods respectively. All strategies carry out thinning early in the epidemic, with roguing throughout and protection only after thinning is carried out.

to the objective function value, we test for any systematic differences in control strategy, for example variation in thinning intensity with epidemic size.

Figure 6.10 shows the ordered control strategies using the open-loop framework. There is a very clear shared structure to all control strategies, with thinning always carried out early in the epidemic. Roguing is used throughout, and protection is only used once resource-intensive thinning has lowered bay and redwood densities enough for disease suppression and tanoak promotion in the approximate model. There is no strong systematic pattern to the strategies once ordered by objective value.

Figure 6.11 shows similar results using the MPC framework. Here we can clearly see the additional thinning carried out over the course of the epidemic for many of the parameter sets. The strategies remain similar in structure, with most of the thinning carried out early in the epidemic, roguing throughout, and protection after the initial thinning regime. The difference in objective is here calculated relative to the baseline MPC strategy. Once again, there is little systematic structure when ordered by objective.

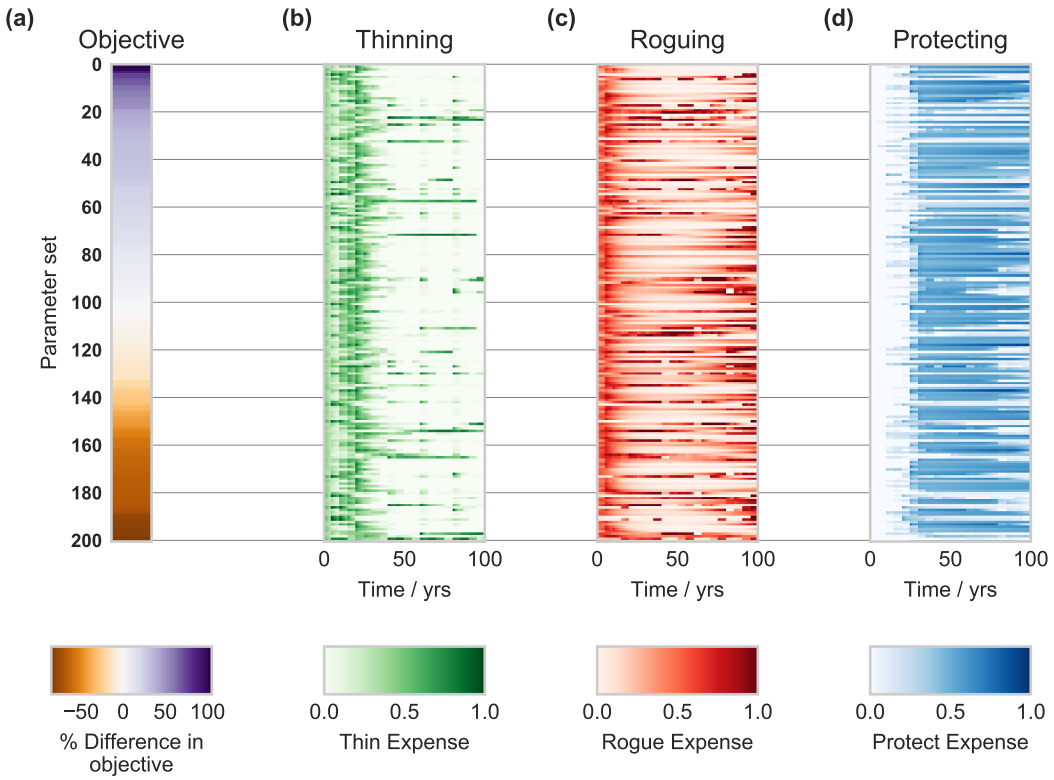

Figure 6.11: Sensitivity of the MPC strategy to parameterisation. As in Figure 6.10, the control allocations are shown ordered by objective value, but here using the MPC framework updating every 20 years. Additional thinning is seen compared to open-loop, but no strong systematic pattern is seen.

Parameter uncertainty

In reality, precise values of infection rates are never known. These parameters are often fitted to limited data with Bayesian techniques, giving a probability distribution of values (e.g. Kleczkowski and Gilligan, 2007; Parry et al., 2014; Thompson et al., 2018; Cunniffe et al., 2014). In this section we test how the open-loop and MPC frameworks handle this type of uncertainty in the system dynamics. Uncertainty is introduced by sampling values from a distribution of infection rates for each species in the simulation model. What benefit does feedback in the MPC framework have when the approximate model cannot accurately capture the dynamics of each individual realisation?

The parameter distribution is chosen to be normal (truncated so that infection parameters remain positive), and the standard deviation for each infection rate is set to 40 % of the parameter value. The approximate model is re-fitted to an ensemble of 200 simulations (Figure 6.12(a) and (b)), with each simulation using a set of infection rates drawn from the distribution. The approximate model uses the same roguing rate scaling factor as previously found (Section 6.2.3, page 111).

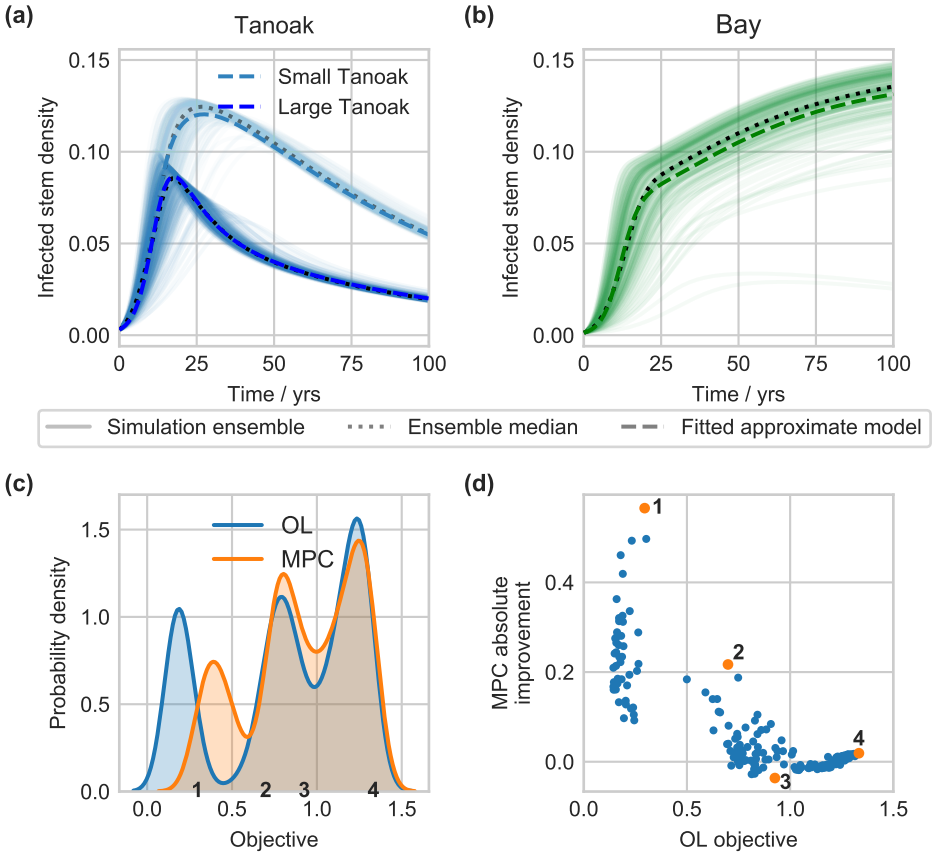

Figure 6.12: The effect of parameter uncertainty on control performance. The approximate model is fitted to an ensemble of simulation runs without control, with infection rate parameters drawn from a truncated normal distribution. **(a)** shows the ensemble and fitted infected tanoak dynamics, and **(b)** the ensemble and fitted infected bay dynamics. **(c)** shows the distribution of objective values using open-loop and MPC across 200 draws of simulation parameters. The absolute improvement of the MPC strategy over open-loop is shown in **(d)**, as a function of the open-loop objective. The MPC framework performs well in the worst-case scenarios, improving control to the largest extent when open-loop performs badly. Four individual cases have been highlighted in panels **(c)** and **(d)**. Further details for each of these, highlighting how rates of spread drive the differences in performance, are shown in Appendix B.1, p. 185.

To test control on these parameter distributions, the fitted approximate model is used to run the open-loop and MPC frameworks for a single draw of infection rates from the distribution. This is repeated for 200 draws from each distribution. The distribution of the resulting objective values under the open-loop and MPC frameworks shows that MPC improves the worst-case scenarios, i.e. the MPC updates are most beneficial when the disease is hardest to manage (Figure 6.12(c)). The open-loop framework gives a distribution of objectives with a worse minimum value than MPC. The continued surveillance of MPC generally improves control, but the greatest improvements are seen when the epidemic is difficult to control, making the open-loop framework ineffective (Figure 6.12(d)). When objective values are high, and so the epidemic is easy to control, there is little difference

between open-loop and MPC. MPC is therefore useful for limiting the worst-case scenarios under parameter uncertainty, highlighting the importance of continued surveillance when disease progression cannot be predicted accurately.

In the open-loop case there is a large peak at very low objective values. In these cases the control strategy fails to manage the epidemic and almost all tanoak is lost. The peak is less prominent in the MPC case since the updates allow control to respond to unexpected changes. This can result in more intermediate performance because outbreaks can be caught at an earlier stage.

Observational uncertainty

The re-optimisation of control in the MPC framework requires accurate information about the current state of the forest at each update step, i.e. every 20 years by default. However, full forest stand surveys are expensive: a cost that is not accounted for in the budget constraint. We test here whether the intensity of these update surveys can be reduced whilst maintaining effective control, and whether open-loop strategies can ever outperform MPC with low quality surveillance.

At each MPC update time a proportion of the forest stand is sampled. The measured state is then used as the initial condition in the approximate model for re-optimisation of the control. For the sampling, each cell in the simulation model is split into 500 1 m^2 discrete units. Surveillance at update times is then carried out by observing a fixed number of units in each cell across the landscape, without replacement. The infection status of tanoak and bay hosts is determined randomly, with probabilities matching the proportion of that host that is infected.

As the proportion of area sampled decreases, the uncertainty in the outcome of MPC increases (Figure 6.13(a)). The median performance of MPC also decreases. This is because as less of the forest is sampled, there is a higher chance that infected hosts will be missed during surveillance, and so the rate of disease spread will be underestimated. We introduce a fixed cost per unit area sampled for the update step surveys, and negate the objective function value as a measure of management benefits. The overall cost is then given by:

$$\text{cost} = k_1 N p - k_2 J \tag{6.9}$$

where k_1 and k_2 are arbitrary constants, N is the number of surveys carried out surveying a proportion p of the forest, and J is the management objective function. The choice of constants k_1 and k_2 is arbitrary, and must be carried out by a forest manager since it balances management and surveillance costs with diversity and tanoak retention benefits. We here choose, as an example, $k_1 = 10$ and $k_2 = 45$. As the proportion of the stand that

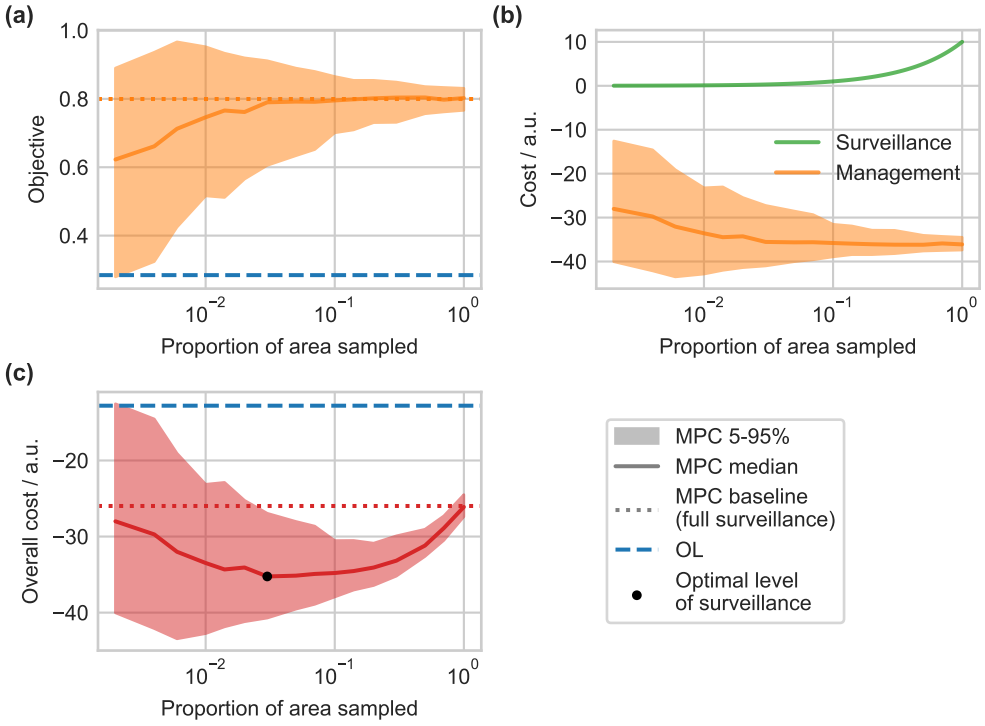

Figure 6.13: **(a)** shows the objective value when there is observational uncertainty at MPC updates. The uncertainty is modelled as random sampling of a proportion of the forest stand area. Shaded areas show the 5th to 95th percentiles of the objective values. In **(b)** the costs associated with control and update surveys are shown. Surveillance costs increase as surveillance becomes more intensive, but the benefits of improved control (negative costs) also increase. **(c)** shows the corresponding overall cost under observational uncertainty, including costs for MPC update surveys. Reduced costs of less intensive surveying must be balanced with less effective control when there is uncertainty about the pathogen extent.

is surveyed increases, so too do the surveillance costs (Figure 6.13(b)). The disease costs however, reduce with more tanoak retained as a result of more informed and effective control strategies.

Balancing these two costs results in an optimal level of surveillance effort (Figure 6.13(c)). The precise location of this optimum depends on the balance between tanoak retention, biodiversity conservation, and surveillance costs: a decision that must be made in the context of local forest management goals. It is clear however, that some level of continued surveillance and re-optimisation through the MPC framework is necessary for effective control. It can also be seen that to minimise the 95th percentile a higher intensity of surveillance is required. This would represent the control policy of a highly risk-averse manager.

6.3.3 Refined optimal strategy

When the approximate model was parameterised, the roguing rate was scaled so that the model matched the simulation as accurately as possible under control. So as not to bias results, no assumptions were made about the form of the control, and the scaling was carried out using a constant roguing rate. Now that the optimal strategy has been found, we are in a position to re-consider this scaling to parameterise the best possible fitting approximate model. Whilst in most real applications of open-loop and MPC this would not be possible due to uncertainties in the simulation model, lack of data or stochasticity for example, here we can test the performance of a refined, optimal approximate model.

The roguing rate in the approximate model was rescaled using the open-loop control strategy found in Section 6.3.1. This ensures the approximate model fits best under the type of control that will be optimal. As before in Figure 6.3, the roguing rate was scaled to best match the final number of tanoaks in the simulation model. As shown in Figure 6.14(a) this gives a much lower scaling factor than was previously obtained.

When this new approximate model is optimised, the control strategy found is very similar to the previous open-loop strategy (Figure 6.14(b)). As with the open-loop framework, control initially focusses on thinning of bay then redwood, with additional roguing. The refined approximate model fits the simulation much better than previously and so the control using the refined approximate model is more effective than either the open-loop or MPC strategies found previously (Figure 6.14(d)). As a result of the increased control efficacy, the strategies are different after the initial phase of control. Where previously further roguing was required, using the refined model additional resources can be used for protection. Since the approximate model fits better, there is little benefit to the repeated updates of the MPC framework, and so MPC using the newly parameterised model is only a marginal improvement over open-loop (Figure 6.14(c)).

This refined model demonstrates how effective optimal control can be if the system is modelled perfectly. However, in the real world the approximate model is unlikely to capture the disease dynamics as precisely, and so suboptimal control frameworks such as open-loop and MPC must be used. We see here that the refined model shows only a modest improvement in control over the suboptimal MPC framework, suggesting that MPC can be used to design effective control strategies.

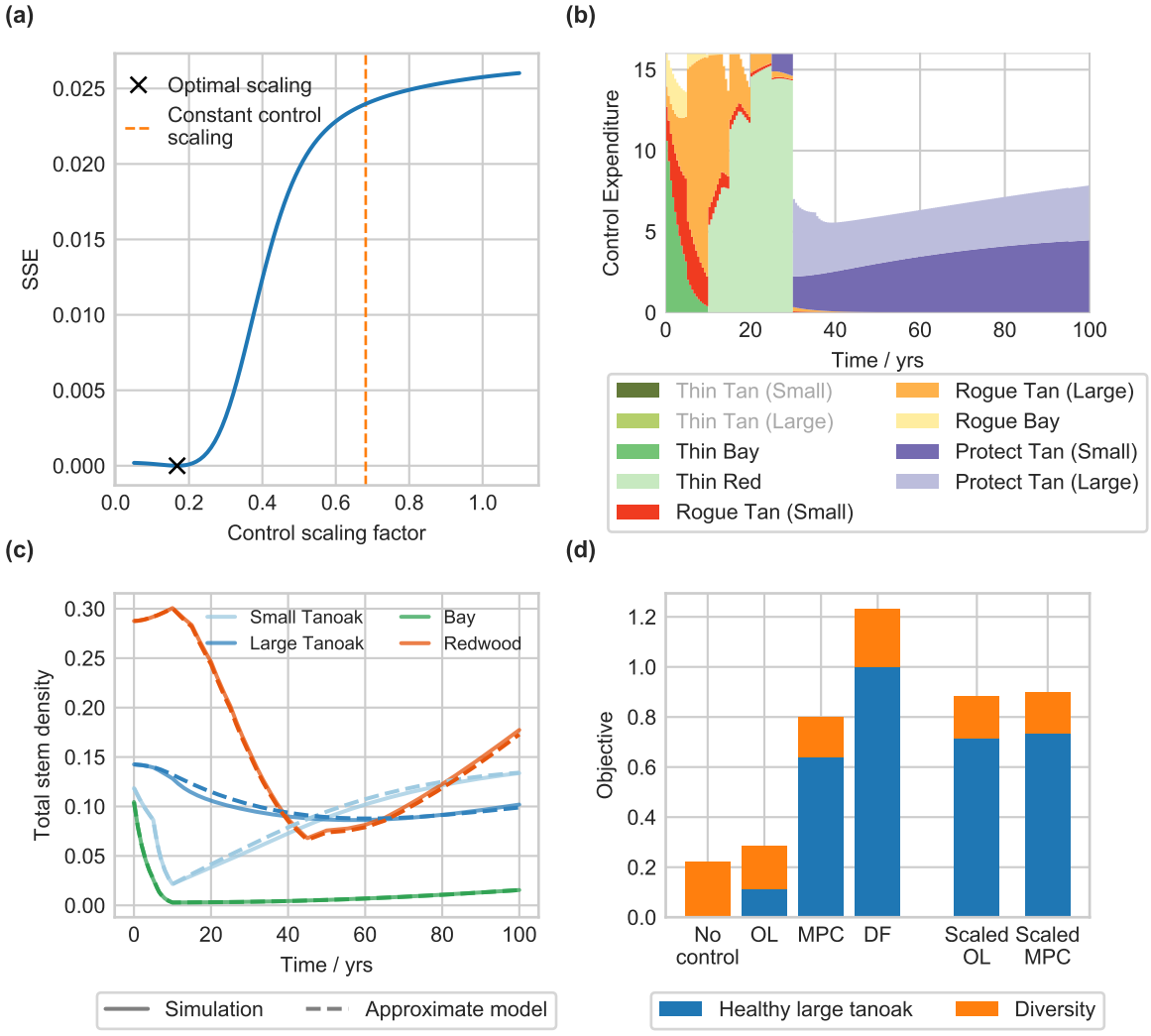

Figure 6.14: Rescaling the roguing rate in the approximate model to find a refined optimal strategy. **(a)** shows the optimisation of the new roguing rate scaling factor, chosen to minimise the deviation in final tanoak numbers between approximation and simulation. The optimal scaling factor under the open-loop control strategy was found to be 0.168 (3 s.f.), compared with the previous scaling of 0.682 (3 s.f.) using constant controls. **(b)** shows the open-loop control strategy found using this newly parameterised model, with the state dynamics shown in **(c)**. The approximate model matches the simulation dynamics very closely. This allows improved disease management, as shown in **(d)**. The newly scaled frameworks outperform the previous strategies.

6.4 Discussion

In this chapter we have seen how the frameworks developed in Chapter 4 can be applied to a real-world scenario: control of sudden oak death in forest stands. As we found in Chapter 4 in a more theoretical setting, the feedback in the MPC framework improves management, and leads to more robust control. In this section we will discuss how these results relate to practical management of SOD.

6.4.1 Practical application of management strategies

Current advice for management of SOD centres around removal of the spreader species, bay laurel, as well as infected tanoak and bay, and susceptible tanoak close to known infections (California Oak Mortality Task Force, 2014). Application of protective chemical treatments is also recommended for high value trees that are close to infections and known not to be currently infected (Lee et al., 2010). The strategies found in this chapter using OCT are broadly similar in nature to these recommendations. We have found that thinning of bay laurel is very important to the success of disease management, echoing results of previous modelling studies (Cobb et al., 2017; Ndeffo Mbah and Gilligan, 2010b). Management advice from the US Forest Service (Swiecki and Bernhardt, 2013) also suggests removal of bay trees, and even complete area-wide removal of bay in some cases. Area-wide removal is only recommended when bay is a minor forest component or consists of only small trees, and only when this removal is consistent with other management goals. We found that roguing of tanoak is more important than roguing of infected bay, as thinning would reduce the bay population density sufficiently for disease control. Our results show that with continued surveillance and careful optimisation of controls, disease can be successfully managed. This can be done whilst maintaining a bay population which may be ecologically important. However, when biodiversity benefits are less important, control is always easier and more effective with additional removal of bay laurel.

Application of chemical protectants is only recommended in practice for individual high value trees close to known infections (California Oak Mortality Task Force, 2014). However, the strategies we have found deploy significant protection resources. In fact, the protectant application only has a very minor effect on the performance of the strategies (see Appendix B.2, p. 190), and is unlikely to be cost-effective. Our formulation of the budget constraint implements a maximum expenditure where a fixed amount of money is put aside for SOD control, rather than minimising total costs. This captures governmental allocation of money for SOD control. In our model, when the optimal levels of roguing and thinning do not use the entire budget, the surplus can be allocated to protectant application. In practice though, control methods will also be individually assessed for cost-effectiveness, and given the limited effect of the protectant strategies it seems unlikely that they would be used.

A possible limitation of the strategies we have found is their complexity. Control inputs were held constant over 5 year stages so that resources do not have to be continually moved, but the strategies are nevertheless still complex in their time dependence and relative allocations to multiple control methods. However, the findings of our results could still be useful. Building up an intuition about what drives the optimal strategy could lead to more practical advice. For example, in the open-loop and MPC strategies thinning of bay is carried

out early before switching to thinning of redwood. These species are reduced to a threshold density that OCT has identified as sufficiently low to suppress pathogen spread, and promote tanoak restoration. This type of insight about optimal densities of different species in mixed stands over the course of an epidemic could provide actionable advice for foresters.

6.4.2 Choosing management goals

The complexity of the optimal strategies found comes in part from balancing multiple costs and benefits: tanoak retention, biodiversity conservation, control expenditure and surveillance costs. The valuation of the cultural and ecological benefits against more direct economic costs is a difficult decision that must be taken by forest managers. These decisions must be made in the context of the local area, as well as other forest management goals such as fire reduction and timber production (Cobb et al., 2017). Decisions about management goals however, either locally or through larger scale regulation, can lead to conflicts that may impact the effectiveness of control (Alexander and Lee, 2010). The value of tanoak retention must be balanced carefully with the wider impacts of the control on the forest.

As well as valuation, the formulation of the different cost functions is important. Here, we used a metric for biodiversity conservation which was integrated over time and so ensures biodiversity is conserved at all times. This captures the importance of continued biodiversity for wildlife habitats, but also avoids introducing edge effects, for example thinning very late in the epidemic to meet a biodiversity target. Both the biodiversity and tanoak objectives introduce a dependence on the chosen time horizon of 100 years, but the tanoak objective only depends on the amount of tanoak at the final time. This final time dependence is appropriate for a restoration-type management goal, such as ensuring a resource is available in the future. There is still a flexibility in the form of the objective function chosen, and the precise choice of objective does impact disease control (Probert et al., 2016).

The chosen time horizon of 100 years is appropriate for analysing structural changes to the forest, including tanoak restoration and biodiversity effects. However, our results are also robust to changes in this time horizon. Extending the final time out beyond the duration of the epidemic to 200 years results in MPC retaining a tanoak population twice as large as under open-loop control (see Appendix B.3, p. 191). Furthermore, control in the short term is similar under a much shorter time horizon of 20 years. Decision makers may optimise control for short-term benefits rather than future tanoak restoration and biodiversity conservation. Despite this, control strategies are similar with focus still on thinning of bay laurel early in the epidemic (see Figure B.8, Appendix B.3, p. 193).

It is also important to consider the effect of the control strategies on the forest. Whilst the effects of the disease can be devastating—particularly to tanoak populations—the optimal

control strategies remove large numbers of trees, including trees that would not have died from the disease. This could be more damaging to the forest ecosystem than the disease effects would have been, particularly given the simplistic modelling approach for ecosystem services and the difficulty of valuing tanoak restoration against biodiversity. Alexander and Lee (2010) write that one land manager who supported removal of tanoak and bay laurel feared an outcome of 'destroy the village to save it'. A difficult problem for forest managers is when control should be abandoned, as has been done in the 'Generally Infested Area' in Oregon (Hansen et al., 2019).

6.4.3 Continued surveillance and re-optimisation

We have shown that by repeated surveillance and re-optimisation of control, the damaging effects of the worst-case scenarios of pathogen spread can be limited. As found by Cobb et al. (2017), disease control is only effective when there is long-term commitment to management projects. However, effort put into this long-term surveillance has to be cost-effective. With imperfect surveillance introducing observational uncertainty, an optimal balance between survey costs and epidemic control was found. Other modelling studies of SOD have also found that resource constraints lead to a trade off between detection and control (Ndeffo Mbah and Gilligan, 2010a; Cunniffe et al., 2016). However, our analysis does not incorporate the risk of disease re-emergence. Our model is deterministic, so also does not capture stochastic re-introductions. Furthermore, the state of the epidemic in the wider region will impact the risk of re-emergence through potentially increasing inoculum pressure, for example from the advancing wavefront of other epidemics in the local vicinity. Forest managers must take into account these other factors in determining optimal levels of surveillance as well as in designing controls, but regardless we have shown that vigilance to disease progression is important. Alongside this vigilance must be a willingness to adapt control measures, re-optimising control to suit the current state of the epidemic and changing local management goals.

6.4.4 Robust control

We showed in Section 6.3.2 that the general form of the optimal control strategy is robust to changes in parameterisation. In all cases a period of thinning is carried out before protection starts, and roguing is carried out throughout the epidemic. However, the relative performance, both in parameter sensitivity and parameter uncertainty studies, can vary significantly. In Figure 6.12 we saw that, under high levels of parameter uncertainty, some epidemics become much easier to control, and others much harder. In the open-loop case this leads to a split in the distribution of objective values into those where initial control

manages the disease, and those where re-emergence leads to significant loss of tanoak. This shows how a poorly designed strategy can lead to significant disease impacts.

In testing robustness we showed that the MPC framework is able to mitigate the effects of the most damaging epidemics, improving management performance in the worst-case scenarios. MacCleery (2015) states that a major barrier to US Forest Service management is the opposition to adaptive management, in which ongoing monitoring is used to update management advice. This is seen as too 'experimental' and increases short-term risk, but imposing fixed interventions means strategies cannot be adapted based on what is seen 'on the ground'. Here we have shown a clear benefit to the ongoing surveillance and re-optimisation of control, with MPC as a possible formal framework for adapting strategies. However, even the simpler open-loop framework still significantly slowed tanoak decline, and slowing pathogen spread is still a useful goal allowing time to prepare for ecosystem impacts (Cobb et al., 2013).

6.5 Conclusions

In this chapter we have optimised strategies for slowing, or even halting, pathogen-induced decline of tanoak in mixed species forest stands. OCT was used to find optimal time-dependent deployment of thinning, roguing, and protectant control resources. The strategies found are broadly consistent with current expert advice: focussing on thinning of bay laurel and roguing of infected trees. However, the strategies we found show significant time dependence. Continued surveillance and re-optimisation of the control strategy by using the MPC framework improves control performance. MPC leads to robust strategies that can effectively respond to unanticipated disease dynamics and system uncertainties, and so manage SOD to protect valuable tanoak trees whilst also conserving biodiversity.

7 Optimising spatial strategies to protect Redwood National Park

7.1 Introduction

In this chapter we will extend the optimal control methods we have developed to manage SOD in a more complex, spatial setting. So far our application of OCT has identified strategies that are non-spatial, in the case of protecting tanoak in forest stands in Chapter 6, or of limited spatial detail, e.g. the metapopulation models of Chapters 3 and 4. Here we extend the setting of pathogen spread to a continuous landscape and show how OCT can be used to find spatially complex strategies that can more effectively control the spread of SOD. These optimal spatial strategies can vary in time and across 120 metapopulation cells in the landscape, allowing significantly increased spatial resolution in the control strategies. Plant disease management is most effective when the scale of control matches the scale of the epidemic (Gilligan et al., 2007; Gilligan and Bosch, 2008; Cunniffe et al., 2015), meaning that necessarily control must depend on the pattern of invasion. Spatially optimising control can target key locations that could link multiple regions (Minor and Gardner, 2011), or that if infected would lead to large epidemics (Hyatt-Twynam et al., 2017). Prioritising management based on host risk can improve management of plant diseases (Cunniffe et al., 2016), but also more broadly of animal (Tildesley et al., 2006) and human diseases (Fraser et al., 2004). How can OCT be used to design these strategies?

As a case study, we use the 2010 SOD invasion along Redwood Creek, near to Redwood National Park in California. We will investigate how OCT can be used to identify spatial strategies that are designed to protect the national park from the impacts of SOD. The methodology we employ could be applied equally well to, for example: other plant diseases threatening important natural or commercial hosts, diseases threatening commercial animals or human diseases invading novel environments. With widespread control of SOD impossible (Cunniffe et al., 2016), designing the most effective strategies to protect valuable resources is essential. The strategies we identify will be compared with the control that was actually carried out in practice, a simple 100 m cull radius. We demonstrate the benefit of using OCT for strategy design[1].

[1]All code for this chapter is available at `https://github.com/ehbussell/RedwoodCreekAnalysis`

7.1.1 Redwood Creek sudden oak death outbreak

Redwood National Park is located on the coast of northwestern California, just south of the Oregon border (Figure 7.1(a)). The park was established in 1968 to protect redwoods from extensive logging, and the combined Redwood National and State Parks (RNSP) now contain 45 % of old-growth (never been harvested) redwood forest in California (National Park Service, 2019). As well as redwoods, the parks preserve the ecosystem and natural biodiversity, conserving flora, fauna and natural features such as rivers and streams (National Park Service, 2019). The area is also of importance to northwestern Californian Native Americans, with Hoopa and Yurok tribal lands nearby. Redwood Creek is a river passing through the national park, from the southern boundary up to Orick (Figure 7.1(b)).

In May 2010, stream baiting near Orick identified the presence of *P. ramorum* for the first time, over 80 km north of the nearest Californian SOD infestation (Valachovic et al., 2013a). Stream baiting places mesh bags containing rhododendron leaves that are susceptible to infection by the pathogen in rivers and streams. The leaves are periodically tested in the laboratory for *P. ramorum* presence and replaced with new leaves. The 2010 detection near Orick confirmed that the pathogen causing SOD was present, but the location of the inoculum source could have been anywhere within the 80,937 ha watershed of Redwood Creek (Valachovic et al., 2013a). Because of the importance of the region—due to the proximity of RNSP, tribal lands and USDA Forest Service lands—a large surveillance effort was carried out to identify the source, including additional stream baiting. In July 2010 one potential source was coincidentally found and confirmed as a small infestation in a residential area across several private properties in Redwood Valley (Figure 7.1(b)) (Valachovic et al., 2013a).

Subsequent control of the infestation involved collaborations between public, private and tribal land managers, and resulted in changes to legislation to offset some of the management costs for commercial landowners (Valachovic et al., 2013a). The management carried out removed all detected infected trees, and all tanoak and bay laurel within 100 m of the infected trees (Valachovic et al., 2013a). This level of management was based on experimental treatments in southern Humboldt county (Valachovic et al., 2010; Valachovic et al., 2013b) and on experiences from management in Oregon (Goheen et al., 2010). Over 150 ha were treated with funding from a number of collaborators but, despite the scale of the treatment efforts, weather conditions conducive to the pathogen and the short cull radius ultimately hampered the effectiveness of control (Valachovic et al., 2013a). The disease was discovered inside the national park in two locations in 2014 near Bridge Creek and Bond Creek (Stark et al., 2014). Management continues with the aim to protect the national park from further infestations.

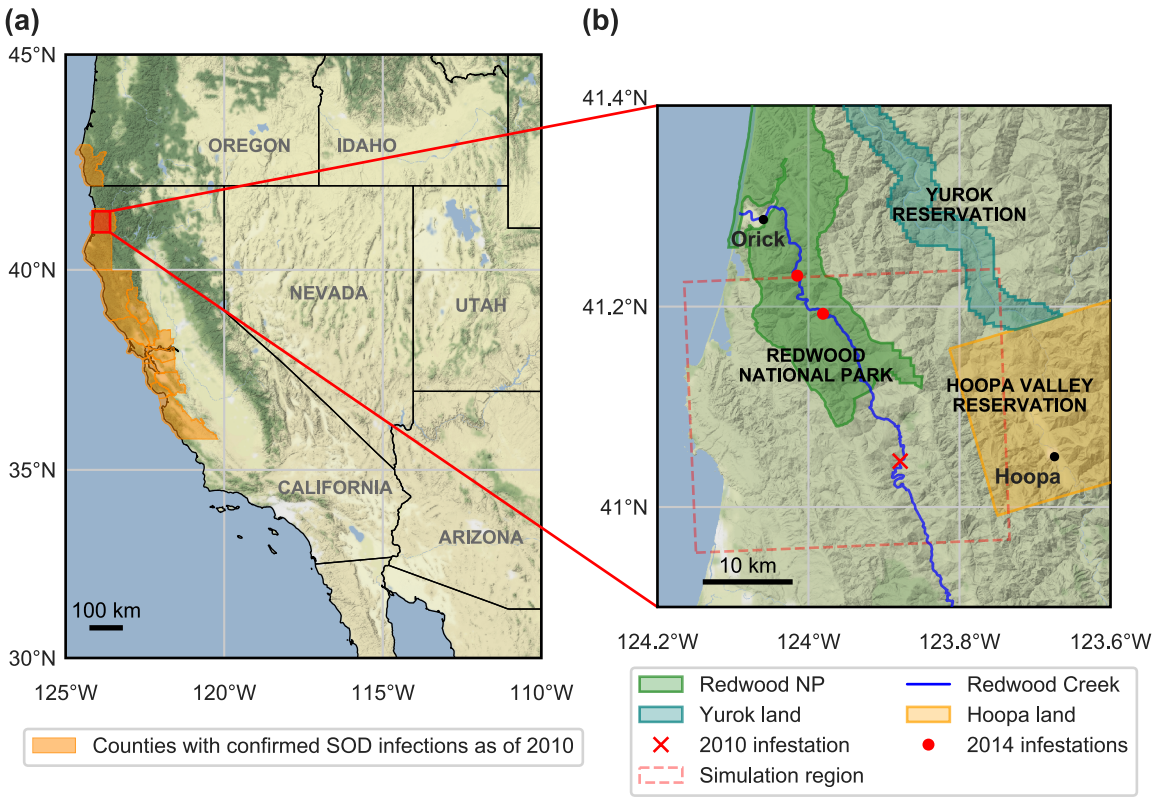

Figure 7.1: Redwood Creek study area and SOD outbreak, with map background showing hill shading and natural vegetation colours. **(a)** shows the counties with confirmed SOD infestations in 2010. The red box in **(a)** shows the study region around the southern tip of Redwood National Park in California, shown in more detail in **(b)**. The national park is close to Yurok and Hoopa tribal lands. In 2010, SOD was detected through stream baiting of Redwood Creek near Orick. The infestation was later located at Cookson Ranch (red cross south of the national park). In 2014 the disease was found inside the national park at Bridge and Bond Creeks, shown by the red dots. The region over which simulations are carried out in this chapter is shown by the dashed red box.

7.1.2 Aims and key questions

In this chapter we will investigate how to use OCT to design spatial strategies to protect Redwood National Park. Using the open-loop framework previously developed, we will test the strategies on a complex, spatially-explicit model of SOD spread at the landscape scale, and compare the strategies with the 100 m buffer scheme that was used in practice. We seek to answer two main questions:

1. How can OCT be used to design optimal spatial strategies to protect a high value region?

2. For the Redwood Creek case study, how do the strategies identified using OCT compare with the management that was carried out in practice?

We start by describing the simulation model used in this chapter, and which approximate models are appropriate for spatial optimisation. Using reduced resolution ODE models of the SOD invasion, we optimise objectives to protect the national park whilst also varying the relative benefit of protecting hosts outside the park. We compare the strategies and their wider effects on the surrounding regions. This chapter can be considered an extension of the simple two-patch model in Chapter 3 to continuous landscapes, applied to the Redwood Creek infestation.

7.2 Simulation model

In Chapter 1 (Section 1.3.2, p. 23) we reviewed models of SOD spread. Two models captured SOD dynamics at the landscape scale, the scale appropriate for modelling the invasion into Redwood National Park. The first model by Meentemeyer et al. (2011) was fitted to data and shown to capture spread across California, and has subsequently been used to assess potential large scale management efforts (Cunniffe et al., 2016). The second model by Tonini et al. (2018) is similar in structure, but as it is integrated with the forest simulation model LANDIS-II (Scheller et al., 2007), the scope for complex control strategies of the form we consider is limited. We therefore use a reimplementation of the model from Meentemeyer et al. (2011) in this chapter. The controls implemented by Cunniffe et al. (2016) include area-wide removal of susceptible and infected hosts, and test a number of prioritisation strategies, but the controls are only implemented across the entire state. The controls at this scale are ineffective, or at least require infeasible levels of host removal, but the model was not used to investigate whether local controls could be used to protect local regions. Here we will use this model on a smaller scale, modelling the infestation near Redwood National Park.

7.2.1 Model structure

The model is a stochastic, spatially-explicit, raster-based simulation, with each simulation cell containing a number of host units that each represent hosts susceptible to SOD. When conditions are conducive to pathogen sporulation, infected hosts produce inoculum—*P. ramorum* spores—that are distributed according to a dispersal kernel and can infect susceptible host units across a heterogeneous landscape. Conduciveness depends on the host type, temperature and moisture conditions, and time of year. In the form described by Meentemeyer et al. (2011), the model does not include host demography, and individual species of host are not tracked separately but are instead combined into one amalgamated

class of host. Host units can be susceptible or infected (Figure 7.2(a)). As described in Meentemeyer et al. (2011) an infected cell i infects cell j at time t at rate:

$$\psi_{ijt} = \beta \left(\chi_t(f_i) m_{it} c_{it} I_{it}\right) \left(\chi_t(f_j) m_{jt} c_{jt} S_{jt} / N_{\text{max}}\right) K_{ji} \tag{7.1}$$

where:

- β is the overall maximum spore production rate from each infected host unit;
- $\chi_t(f_i)$ is a seasonal indicator variable for hosts of type f_i, either 0 or 1 dependent on whether hosts in cell i are able to infect and be infected at time t;
- m_{it} and c_{it} measure how conducive the weather conditions are for pathogen sporulation (moisture and temperature respectively) for cell i at time t, each between 0 and 1;
- I_{it} and S_{it} are the numbers of infected and susceptible host units in cell i at time t;
- N_{max} is the maximum number of host units in any cell; and
- K_{ji} is the dispersal kernel giving the probability of a spore travelling from cell i to cell j.

The dynamics are simulated using the Gillespie direct method (Gillespie, 1977), giving stochastic simulations. Note that in the original implementation a discrete time approximation was made for computational ease, an approximation we do not make here because of the smaller study region. All parameter and variable meanings and default values are given in Appendix C.1.

We use a spatially restricted subset of the same host landscape used by Meentemeyer et al. (2011), in which each 250 m $\times$ 250 m cell is given a localised host index value that combines abundance and susceptibility across all hosts in the cell. This index is then discretised to give a number of host units per cell, with a maximum of 100 in any cell (Figure 7.2(c)). Since the hosts are conducive to pathogen sporulation at different times of the year, each cell is classified as predominantly redwood and tanoak, or mixed evergreen forest. This forest type mask sets the seasonal indicator variable for each cell ($\chi_t(f_i)$). Redwood/tanoak forest can infect and is able to be infected for the first 28 weeks of the year. Mixed evergreen forests are suitable for the pathogen from week 7 through to week 28.

Model fits by Meentemeyer et al. (2011) indicate infected host units produce spores at a rate of 4.55 wk^{-1} (Cunniffe et al., 2016), which are distributed in space according to

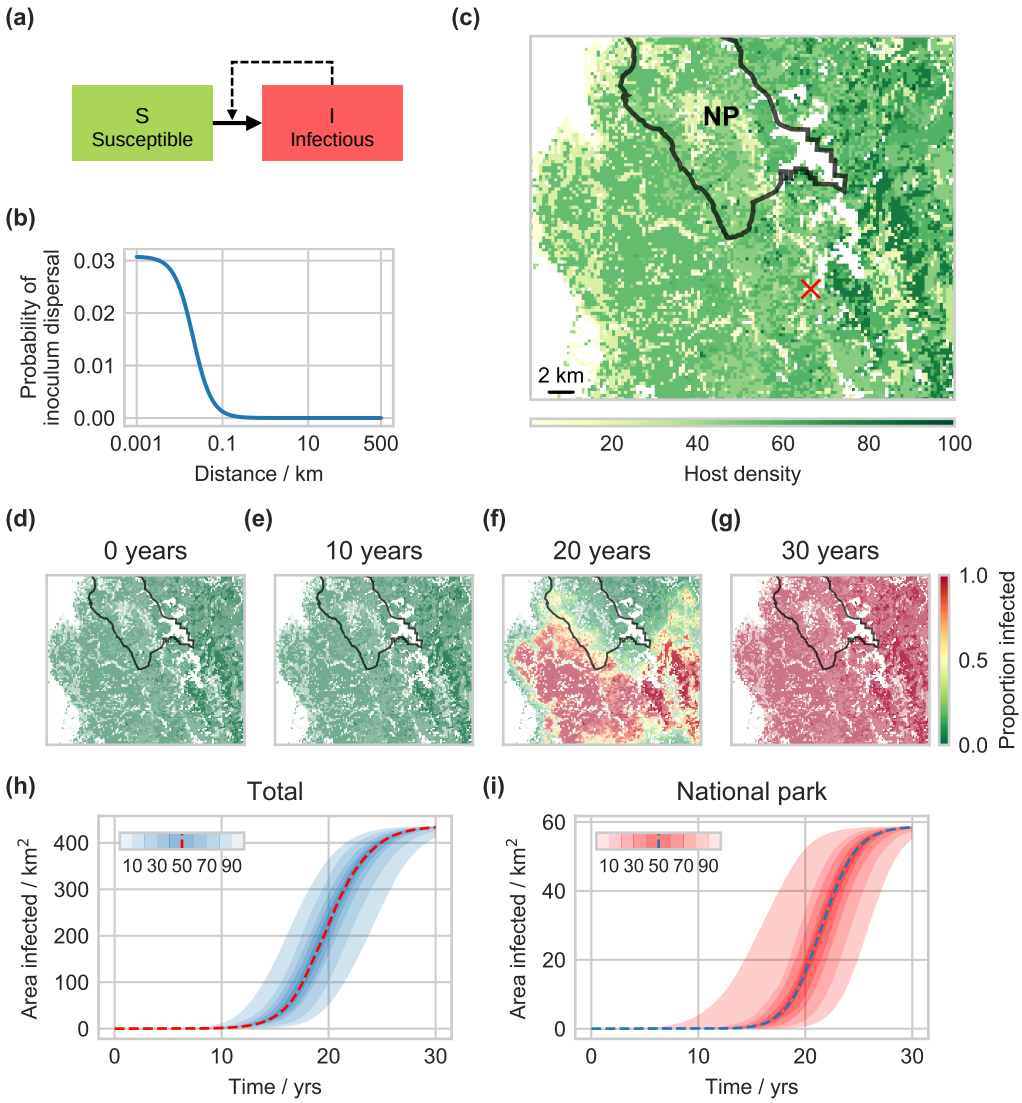

Figure 7.2: The simulation model capturing the spread of SOD into Redwood National Park. The hosts transition between susceptible and infected, as shown in **(a)**. Infected hosts produce inoculum which is then distributed according to the kernel shown in **(b)**, a combination of a long- and short-ranged Cauchy kernel. In **(c)** the host landscape around the south of the national park is shown, with the red cross indicating the initial infection site and the black line marking the edge of the national park (NP). **(d)**–**(g)** plot the median level of infection across 100 realisations of the simulation model over 30 years. The colour indicates the proportion of hosts infected, and the transparency indicates the host density. **(h)** and **(i)** show the disease progress curves across the landscape and just in the national park respectively. The deciles of the distributions and the median are shown in each case.

the dispersal kernel found by Meentemeyer et al. (2011). The kernel is a combination of 2 Cauchy type kernels, giving the probability of a spore dispersing a distance d:

$$K(d) = \gamma \left(\frac{N_1}{1 + (d/\alpha_1)^2} \right) + (1 - \gamma) \left(\frac{N_2}{1 + (d/\alpha_2)^2} \right) \tag{7.2}$$

where $\gamma = 0.9947$ is the proportion of dispersal events distributed according to the short range kernel, $\alpha_1 = 20.57$ m is the short range kernel scale, and $\alpha_2 = 9.504$ km is the long range kernel scale (Figure 7.2(b)). The constants N_1 and N_2 are normalising constants for the two kernels in the expression reported here. Infection is seeded in the cell corresponding to Cookson Ranch, where the initial infestation was found, with that cell starting fully infected. Overall, simulating the model around the southern tip of Redwood National Park leads to widespread infection over 30 years (Figure 7.2(d)–(g)). The pathogen tends to spread more to the west where conditions are more conducive to pathogen spread (Figure 7.3(a)–(b)).

7.2.2 Sporulation conditions

The simulation model used by Meentemeyer et al. (2011) uses the forest type mask that varies pathogen suitability in space and time, and weather data (m_{it} and c_{it} in Equation 7.1) that also varies by cell and each week. This gives rise to the increased spread to the west seen in Figure 7.2(d)–(g). Since we will be approximating the simulation model using an ODE system, removing this time dependence would make the ODEs simpler and make convergence using OCT easier. It would also improve the computational efficiency of running the simulation model. We tested what effect averaging these time-dependent variables has on the simulations, generating a single time-independent scaling factor for each cell that captures the effects of temperature, moisture and forest type (Figure 7.3).

To calculate the average, we take for each cell i the root mean square value of the weather and forest type mask over time:

$$M_i = \sqrt{\frac{\sum_t \left(\chi_t(f_i) m_{it} c_{it}\right)^2}{N_t}} \tag{7.3}$$

where N_t is the number of time points. The root mean square is used because in the infection rate (Equation 7.1) these terms appear in the susceptible and infected terms, and so are effectively squared. Taking a simple mean leads to significantly lower levels of infection. Since the kernel is very short ranged, taking this local average is a good approximation, since in reality the value is not squared but multiplied by the value in another cell. To understand why this is necessary, consider a single cell with conduciveness equal to 0.5 for half of each year, and zero otherwise. Whilst using the mean conduciveness of 0.25 leads to a relative infection rate of 0.0625 (0.25^2), the actual mean relative infection rate is 0.125 (0.5^2 for

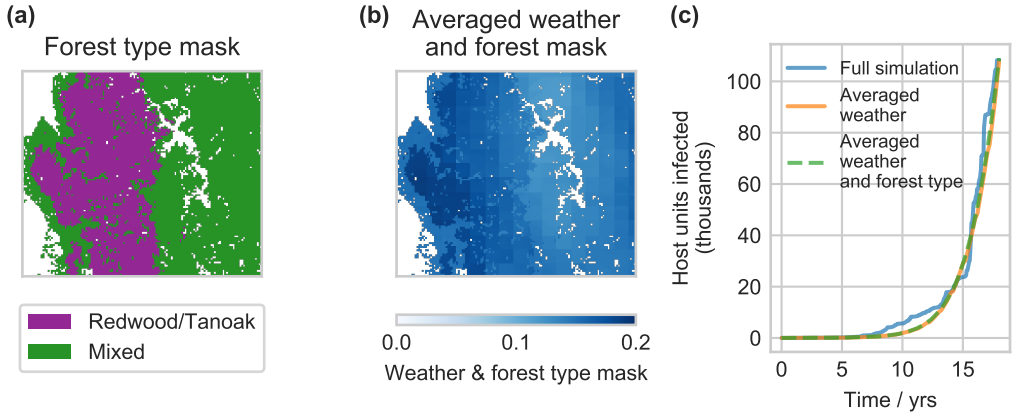

Figure 7.3: Accounting for varying weather and forest type in the simulation model. In **(a)** the forest type around the national park is shown, with regions of mixed evergreen forest, and redwood and tanoak areas. The redwood/tanoak forest is susceptible and can sporulate for the first 28 weeks of the year. Mixed evergreen forests are unsuitable for the pathogen until the 7^{th} week of the year, and become unsuitable again after the 28^{th} week. **(b)** shows the average mask across the region, with weather and forest type averaged over time. The weather and forest type are more suitable for the pathogen closer to the coast. In **(c)** the effect of averaging the weather and forest type mask is shown. The lines show the medians of 250 simulation realisations. The full simulation has weekly varying weather, and differences in pathogen suitability due to forest type. Averaging just the weather over time, or averaging both the weather and forest type suitability over time, does not significantly affect the simulation dynamics. We therefore use the averaged weather and forest type mask for simulations going forward, because of the increased speed of simulation.

half the year, and zero otherwise). The root mean square gives the correct relative infection rate.

It is clear that the redwood/tanoak forest nearer the coast is generally more suitable for the pathogen, because of both the weather and forest type mask (Figure 7.3(a)–(b)). We ran 250 simulations over 18 years using the full weekly weather and forest type, averaging just the weather, and averaging both the weather and forest type over time. Whilst averaging these effects over time does reduce within year variation, it does not have a large effect on the median simulation dynamics (Figure 7.3(c)). We therefore use the averaged weather and forest type mask for simulations in this chapter.

7.3 Approximate model

7.3.1 Reduced resolution model

We require an approximate model of the simulation dynamics which must be simple enough to allow optimisation using OCT, but with enough spatial detail to allow spatially resolved strategies to be identified. As mentioned in Chapter 4, the main factor affecting convergence in OCT is the number of variables that must be optimised. As spatial detail is

added this increases the number of state and control variables, but also increases the spatial resolution of the control strategy. We therefore seek an approximate model that captures as much spatial detail as possible whilst remaining optimisable. We propose using reduced resolution ODE models of the simulation, which are raster-based ODE approximations of the dynamics on a grid of a larger spatial scale (Figure 7.4(a)–(b)). The host landscape used is an aggregated version of the simulation host landscape, with each cell in the approximate model using the average host density of the 250 m simulation cells contained within it. A buffer region is included in the simulation model around the approximate landscape (Figure 7.4(a)). This ensures that any edge effects from infection leaking around the managed region are accounted for. This more realistically models the effect of inoculum pressure from outside the area under management.

The ODE approximation follows the same susceptible-infected dynamics as the simulation model, with dispersal across the landscape according to a different kernel. Since the approximate model uses grid cells of a different scale to the simulation and is deterministic, the same kernel cannot be lifted from the simulation. We use the averaged weather and forest type mask to modulate the susceptibility and infectiousness of each cell, with the average now calculated for each cell in the aggregated landscape. The ODE system for susceptible ($\tilde{S}$) and infected ($\tilde{I}$) hosts in aggregated cell i is therefore given by:

$$\dot{\tilde{S}}_i = -\tilde{\beta} M_i \tilde{S}_i \sum_j \left(k_{ij} M_j \tilde{I}_j\right) \tag{7.4a}$$

$$\dot{\tilde{I}}_i = \tilde{\beta} M_i \tilde{S}_i \sum_j \left(k_{ij} M_j \tilde{I}_j\right) \tag{7.4b}$$

where $\tilde{\beta}$ is the infection rate to be fitted, k_{ij} is the kernel also to be fitted, M_i is the averaged weather and forest type mask, and the sum is over all aggregated cells in the landscape.

7.3.2 Metrics for comparison

To choose the most appropriate resolution for the approximate model, by testing the quality of fit and plausibility of control optimisation, we require metrics that can compare the approximate model with simulations across different resolutions. This will ensure that we can assess different resolution approximate models using a single consistent metric. There are three obvious choices of scale for comparing a single approximate model with simulations:

Landscape scale: total infection across the landscape;

Aggregated scale: simulation data aggregated to the approximate model resolution;

Divided scale: approximate model data resampled down to the 250 m simulation scale.

For clarity, in a simulation the number of infected host units in cell i is given by $I_i(t)$. This is the divided scale. At the landscape scale the numbers are simply summed over all cells ($\sum_i I_i(t)$). At the aggregated scale the numbers are summed over the set of cells contained within the aggregated cell j (A_j):

$$\tilde{I}_j = \sum_{i \in A_j} I_i \tag{7.5}$$

To calculate the number of hosts at the divided scale from the approximate model, the hosts are homogeneously distributed across simulation cells contained within the aggregated cell:

$$I_i = \frac{\tilde{I}_j}{N_j} \quad \forall i \in A_j \tag{7.6}$$

where N_j is the number of simulation cells contained within aggregated cell j.

For comparing across different resolutions, the aggregated scale cannot be used since this will be a different absolute scale for each approximate model. The landscape and divided scales can be used to compare across resolutions, though. The metric we choose for comparison is the root mean square error (RMSE) between the simulation and approximate model disease progress curves (DPCs). This metric is a scaled version of the summed squared errors (SSE) used in previous chapters. We use the RMSE in preference to the SSE because the value is more meaningful for comparison, since it is measured in the same units as disease progression. At the landscape scale, the RMSE for the landscape DPCs is used. At the divided and aggregated scales the RMSE is calculated over DPCs for each cell over 30 years with a time step of 2 weeks.

7.3.3 Fitting

We fit two forms of kernel (k_{ij} in Equation 7.4), an exponential type:

$$k^1_{ij} \propto \exp\left(-d_{ij}/\sigma_1\right), \tag{7.7}$$

and a Cauchy type kernel:

$$k^2_{ij} \propto \frac{1}{1 + \left(d_{ij}/\sigma_2\right)^2}, \tag{7.8}$$

where d_{ij} is the distance between cell centres for cells i and j, and σ_1 and σ_2 are scale parameters for the two forms of kernel. Whilst other forms of kernel could be considered—for example an exponential power distribution as considered by Skelsey and Garrett (2013) that can capture both short- and fat-tailed shapes—there was little difference in fit between the exponential and Cauchy kernel models. We therefore only consider these possibilities. The scale parameters and infection rate ($\tilde{\beta}$) are fitted by minimising the RMSE at the aggregated scale using the Nelder-Mead simplex algorithm (Gao and Han, 2012) from the SciPy library (Jones et al., 2001–). This scale is used to maintain spatial information whilst avoiding

repeated resampling operations that would be required to calculate the metric at the divided scale.

The fitted dynamics match the simulation very closely at the landscape scale (Figure 7.4(c)), with the largest errors spatially close to the infection introduction site (Figure 7.4(d)–(f)). In spatial pattern, all reduced resolution models capture the increased spread to the west where weather and forest type conditions are more suitable for the pathogen (Figure 7.5). Fitted parameters for each approximate model are given in Appendix C.2.

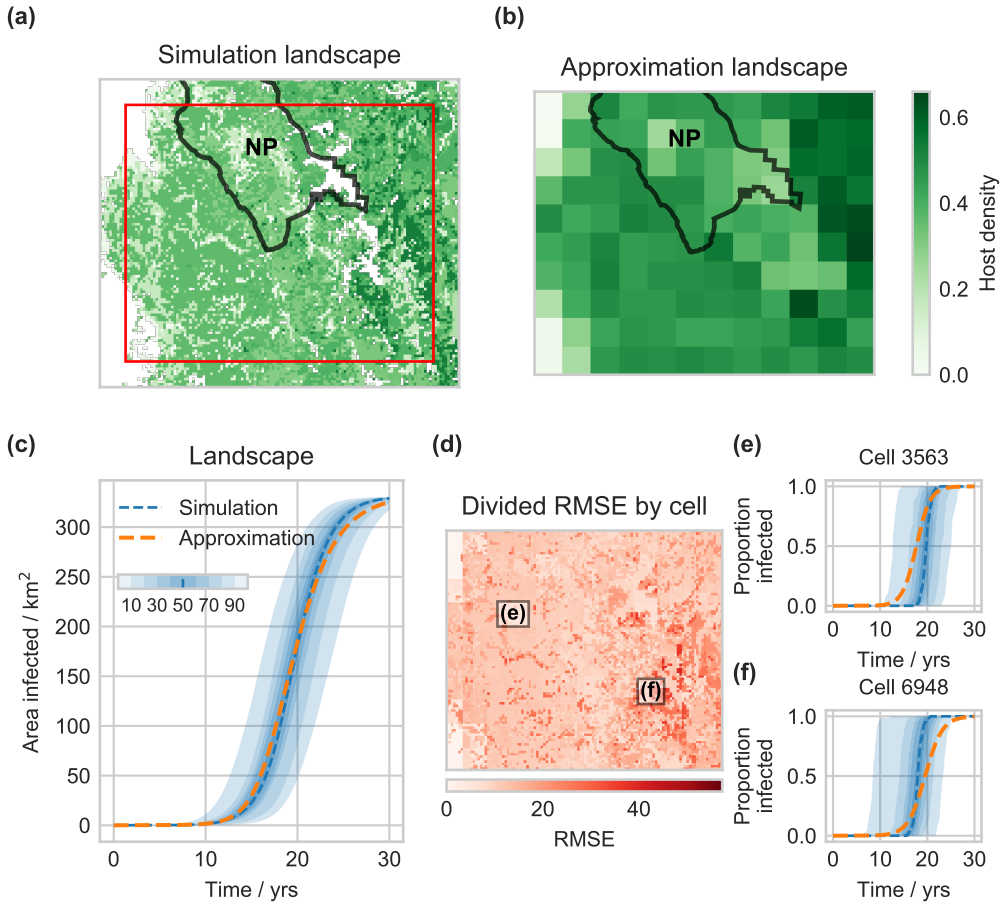

Figure 7.4: Approximating the simulation model with a reduced resolution model. The simulation host landscape in **(a)** includes a buffer to eliminate edge effects. The red box shows the region approximated by the reduced resolution host landscape, shown using a 2,500 m resolution in **(b)**. The ODE approximate model is fitted to the simulation model by matching DPCs for each cell in the approximate model, here shown for the 2,500 m resolution model with a Cauchy kernel. In **(c)** the simulation and approximate disease progress curves across the whole landscape are shown, with the approximate model capturing the dynamics well. In **(d)** the root mean square error (RMSE) for each cell at the scale of the simulation is shown (divided metric). The largest errors are seen close to the initial seed infection. **(e)** shows the disease progress curves for a single 250 m cell far from the initial seed infection, capturing the dynamics well. **(f)** shows the same for a cell close to the initial infection. The low resolution approximate model cannot precisely capture the number of hosts in each 250 m cell, and so the dynamics are captured less well.

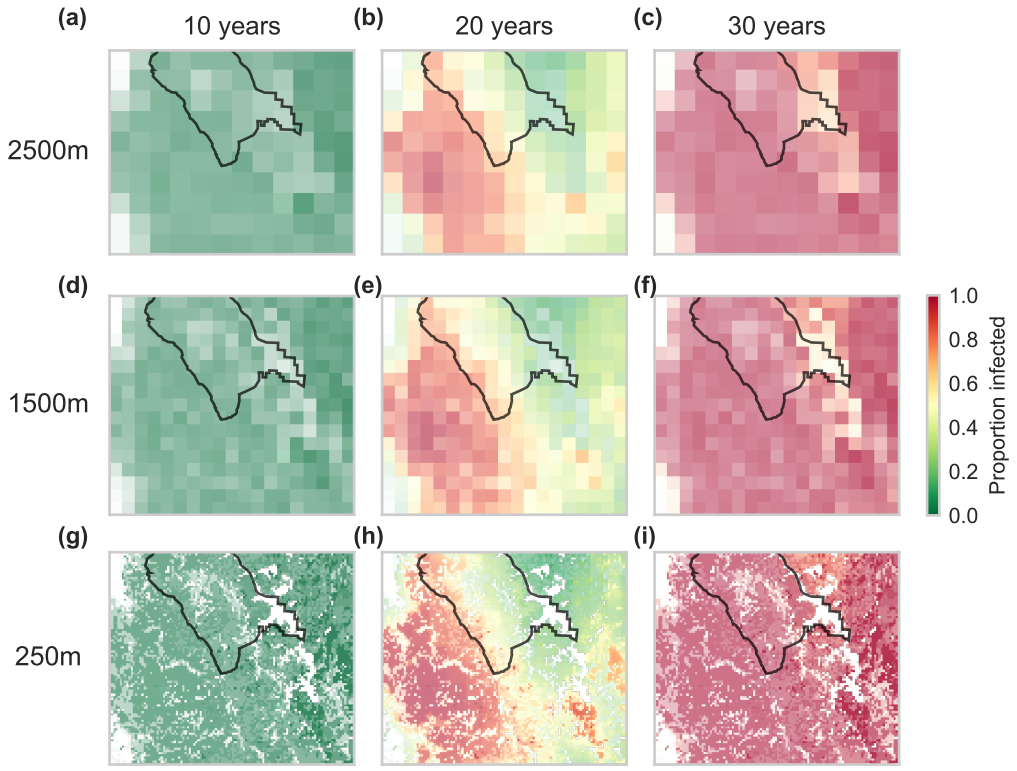

Figure 7.5: Comparing disease spread across approximate model resolutions. The top row shows infection spread in the 2,500 m resolution approximate model after 10, 20, and 30 years. The second and third rows show the same for the 1,500 m and 250 m resolution models respectively. All model resolutions capture the same pattern of spread, with infection first spreading towards the coast.

7.4 Control and optimisation

7.4.1 Control methods

In the simulation and reduced resolution approximate models, we implement control methods that remove infectious or susceptible hosts at a continuous rate. Optimising control finds a partition between removing susceptible hosts and infected hosts across the landscape, in order to achieve an objective. The roguing control method removes infectious hosts, and the thinning method removes healthy, susceptible hosts. The time-dependent variables $u_i(t)$ and $v_i(t)$, each between 0 and 1, give the level of thinning and roguing respectively in cell i of the approximate model. With η as the rate of removal, the approximate model dynamics become:

$$\dot{\tilde{S}}_i = -\tilde{\beta} M_i \tilde{S}_i \sum_j \left(k_{ij} M_j \tilde{I}_j\right) - u_i(t)\eta \tilde{S}_i \tag{7.9a}$$

$$\dot{\tilde{I}}_i = \tilde{\beta} M_i \tilde{S}_i \sum_j \left(k_{ij} M_j \tilde{I}_j\right) - v_i(t)\eta \tilde{I}_i \,. \tag{7.9b}$$

We impose a simple constraint on the thinning and roguing levels:

$$\sum_i \left(u_i(t) + v_i(t)\right) \leq 1 \quad \forall t \,, \tag{7.10}$$

where the sum is over all cells in the approximate model. This means the total control rate η is partitioned into thinning and roguing, and across all cells in the approximate model. This captures the logistical constraint of limited resources to allocate to particular locations in the landscape. Note that the constraint here is on intensity of control, not the number of hosts removed. This captures the difficulty of resource location; control resources allocated to cells with fewer hosts to remove still have to be moved to that location. This constraint is simpler in form than those used in previous chapters. By separating the state from the control constraint, convergence when optimising the control is significantly improved. Controls are lifted from the approximate model to the simulation model as expected, with control homogeneously applied across simulation cells contained within each approximate model cell. As before, disease is seeded at Cookson ranch, and control starts after 3 years of uncontrolled spread to allow the infestation to establish.

7.4.2 Objective function

We use an objective function that maximises the number of susceptible hosts at the final time $T = 30$ years, with a weight for each cell in the approximate model landscape. The weight for hosts inside the national park is equal to 1. For cells that are partially outside the park, this weight is multiplied by the proportion inside the national park. We test different

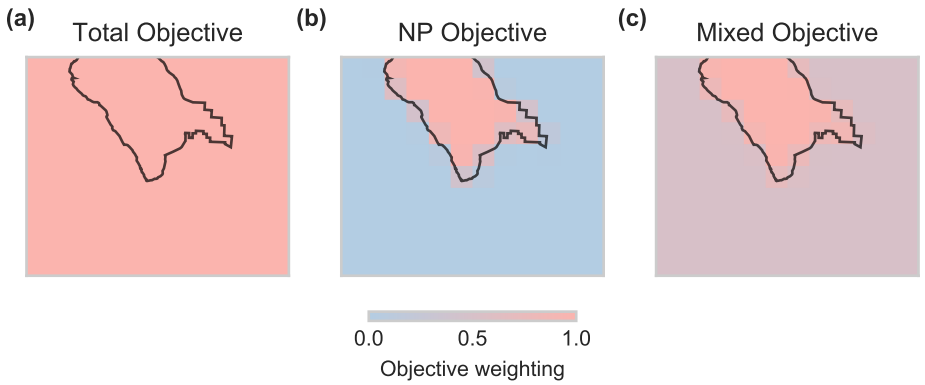

Figure 7.6: The different objective functions under test. **(a)**, **(b)** and **(c)** show the Total, NP and Mixed objective functions respectively.

objectives, varying the weight of hosts outside the national park. The overall objective is given by:

$$J = f_1 \sum_i p_i S_i(T) + f_2 \sum_i (1 - p_i) S_i(T) \tag{7.11}$$

where f_1 and f_2 are the weights for hosts inside and outside the national park respectively, and p_i is the proportion of cell i that is inside the national park.

We test three different objectives (Figure 7.6). The Total objective uses $f_2 = 1$ so that all hosts across the landscape are weighted equally. The NP objective only considers the national park, with $f_2 = 0$. Finally, the Mixed objective uses $f_2 = 0.5$, as a balance between the first two objectives.

7.4.3 Large scale optimisation

Despite reducing the resolution of the approximate model, the optimisation problem is still very large. For a 2.5 km resolution approximate model, there are 120 cells each with two state variables and two control variables. Discretising the 30 year time period into 120 steps, each step representing 7 weeks as infection is only possible for 28 weeks of the year, gives approximately 58,000 variables to optimise using the direct method. The software used in previous chapters to carry out the direct optimisation, BOCOP (Team Commands, Inria Saclay, 2017), is unable to handle problems of this size. For the optimisations in this chapter we use the optimiser underlying BOCOP directly, Ipopt (Wächter and Biegler, 2006), and generate the non-linear programming (NLP) problem ourselves.

The calculation of the NLP jacobian, i.e. the derivative of the NLP constraints with respect to the NLP variables, is the limiting factor in using BOCOP for large scale optimisation, since BOCOP uses automatic numerical differentiation to calculate this derivative. Since we are able to compute the equations and hence the derivative exactly, we implement exact

derivatives in our interface to Ipopt. This allows for faster convergence and lower memory usage. We describe the formulation of the NLP problem in Appendix C.3 and the optimisation details in Appendix C.4. Despite these improvements in optimisation, the highest resolution for which optimisation is successful is the 2.5 km resolution approximate model.

7.4.4 Buffer strategy

We compare the optimal spatial strategies with the control that was carried out in practice. Once the source of infection was located along Redwood Creek, a 100 m radius buffer control was implemented. Any infected hosts that are detected are removed along with all susceptible hosts within 100 m. To implement this strategy using our model, within each cell all host units are assigned a random position. At the end of each year, surveillance is carried out and infected hosts are detected according to a Bernoulli trial with probability of success $p = 0.7$, following the same method used by Cunniffe et al. (2016). These detected hosts and any hosts within 100 m are then removed, following the strategy used in practice.

7.5 Results

7.5.1 Resolution testing

We first describe the results of resolution testing: assessing different resolution approximate models. We use the RMSE metric at both the landscape and divided scales to compare across resolutions. Under no control, resolutions from 250 m to 2.5 km all perform similarly at the landscape scale, with the Cauchy kernel giving slightly lower metric values (Figure 7.7(a)). At the divided scale higher resolution models perform better, as predictions are more accurate since host density information is available at a finer spatial scale. RMSE values are much lower at the landscape scale because errors in host numbers in any individual cell are averaged out across the landscape.

To ensure the approximate models fit when control strategies are implemented, we also test the approximate models under an optimised non-spatial control strategy. A non-spatial strategy is used so that exactly the same control can be used across all resolution models. The strategy is optimised as described in Section 7.4 using the 2.5 km resolution model and the Total objective function, with the added constraint that control is allocated evenly across the landscape. The resulting control carries out thinning first, before switching to roguing (Figure 7.7(c)). Under this control strategy all resolutions fit closely to simulations realised under the same strategy (Figure 7.7(b)). High and low resolution models match the simulation disease progress curves well and once again the largest errors are seen close to the introduction site (Figure 7.7(d)–(e)).

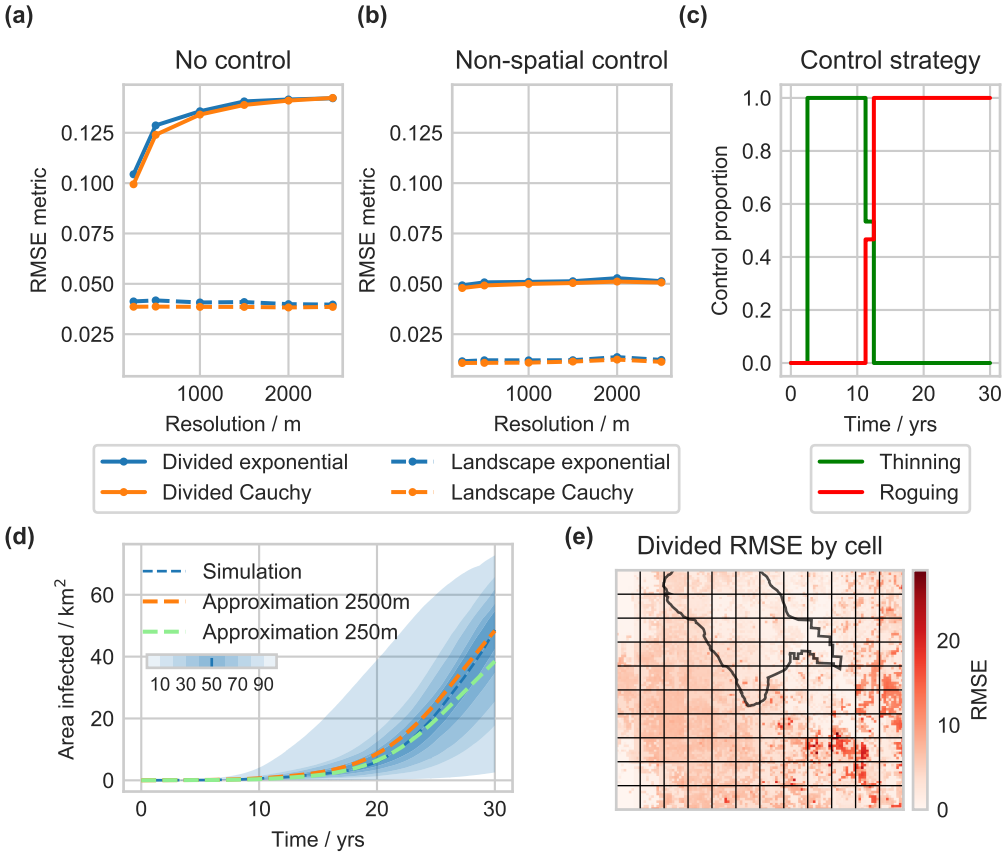

Figure 7.7: Using the RMSE metric the different resolution approximate models can be tested. In **(a)** the landscape and divided metrics are shown for the exponential and Cauchy kernel under no control. The Cauchy kernels show a slightly better fit across all resolutions, with higher resolution approximate models fitting better. Under a non-spatial control strategy, all resolutions fit approximately the same and, since the level of infection is lower, fit better than under no control, as shown in **(b)**. The control strategy is shown in **(c)**. The landscape disease progress curves are shown in **(d)**, and the divided RMSE metric across the landscape for the 2,500 m Cauchy model under control is shown in **(e)**.

We choose to use the 2.5 km resolution approximate model for the optimisations in the rest of this chapter. This resolution provides enough spatial detail to capture the increased spread to the west and allow spatially resolved control strategies, and also fits well to simulation data. Whilst higher resolution models provide improved fits, for resolutions higher than 2.5 km the optimal control problem is not tractable using our methods. Later in this chapter we will compare the resulting optimal control using this approximate model with that from a lower, 5 km resolution model, and show that control is improved by the additional spatial detail in the 2.5 km model.

7.5.2 Optimal spatial control

The control strategies optimised using the 2.5 km resolution approximate model are applied to the simulation model using the open-loop framework described in previous chapters. Using the NP objective, the optimal strategy initially rogues at the disease introduction site alongside heavy thinning in the highly pathogen conducive region to the west (Figure 7.8(d)–(e)). Thinning is then carried out more generally outside the national park, with intensive roguing inside the park. The disease progress curves across the landscape and inside the national parks (Figure 7.8(b) and (c)) show low levels of infection, and the approximate model matches the average landscape dynamics. Median simulation dynamics show significant loss of host outside the national park, but the park itself is well protected (Figure 7.8(f)). Note though, that significant amounts of infection spread around the edge in the buffer region where management is not carried out.

The NP objective protects Redwood National Park, but at a significant cost to the surrounding region. An alternative strategy, the Mixed objective, weights protection of hosts outside the national park by half as much as those inside (Figure 7.9(a)). The optimal strategy under this objective still thins to the west, but uses roguing rather than thinning in the east (Figure 7.8(d)–(e)). This results in higher levels of infection (Figure 7.8(b)–(c)), but retains more healthy host in the north east of the region. As with the NP objective, infection does spread around the buffer region.

We next analyse the performance of the control strategy that was carried out in practice, the 100 m buffer. The buffer strategy keeps levels of infection much lower than the previous strategies, because these hosts are directly detected and removed every year. In some cases the buffer strategy successfully manages the epidemic, keeping levels of infection very low and with few hosts removed (Figure 7.10(b)–(c)). In many cases though, the buffer of 100 m is not sufficient to control the spread and, whilst overall infection levels remain low, a large number of hosts are eventually removed in the national park. This means that the median simulation dynamics show large numbers of hosts removed across the landscape, but the strategy does still protect many hosts in the north east of the region.

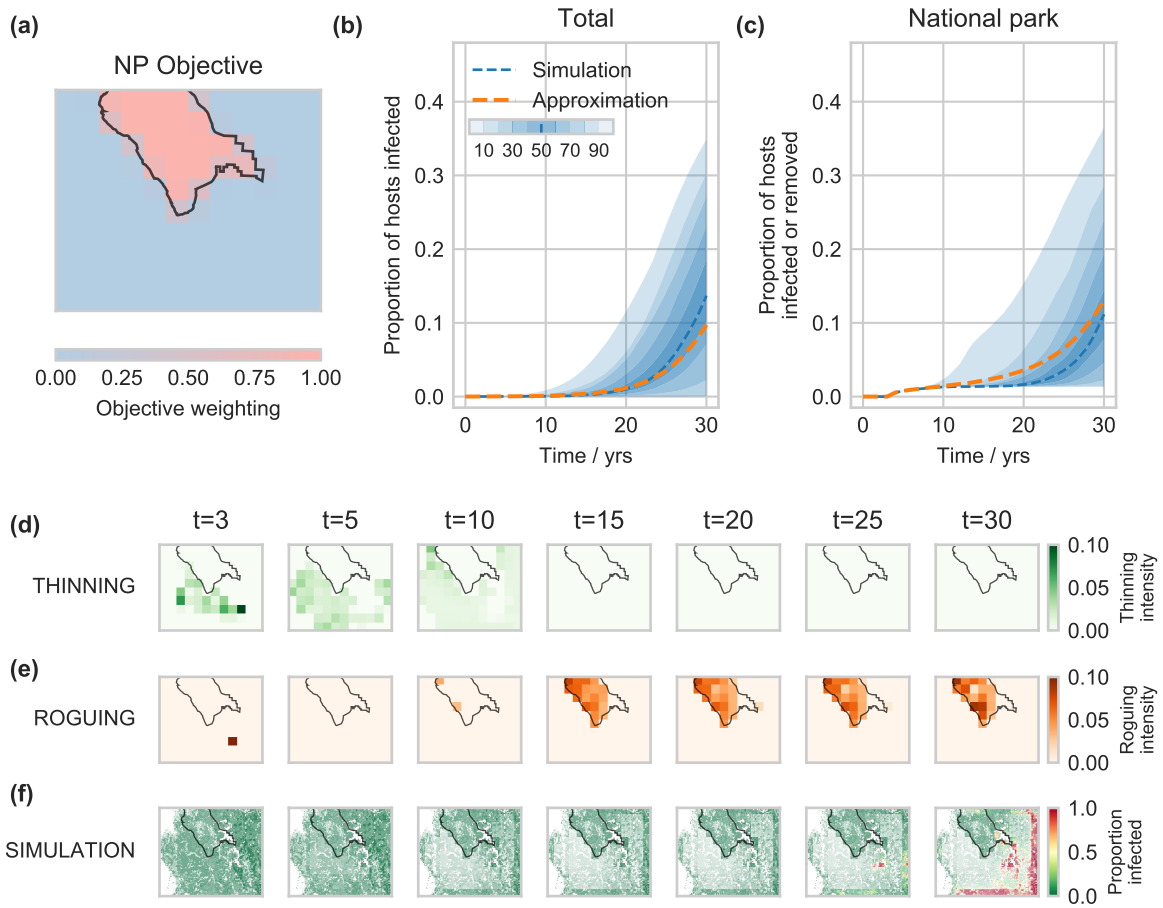

Figure 7.8: Optimal control strategy using the NP objective. **(a)** shows the objective raster, with control optimised to maximise the number of healthy hosts in the national park after 30 years. The simulation and approximate model disease progress curves across the landscape are shown in **(b)**, and the proportion of hosts infected or removed within the national park only in **(c)**. These DPCs do not include the buffer region in the simulation model. The control strategy is shown in **(d)** and **(e)**, with thinning initially focussed near the coast on the west, and roguing focussed in the national park. **(f)** The median proportion of hosts infected in the simulations show a reduction in hosts across the landscape apart from in the national park. The colour indicates the proportion of hosts infected, and transparency indicates the host density. There is little infection spread inside the managed area, but infection does spread around the edge in the buffer region.

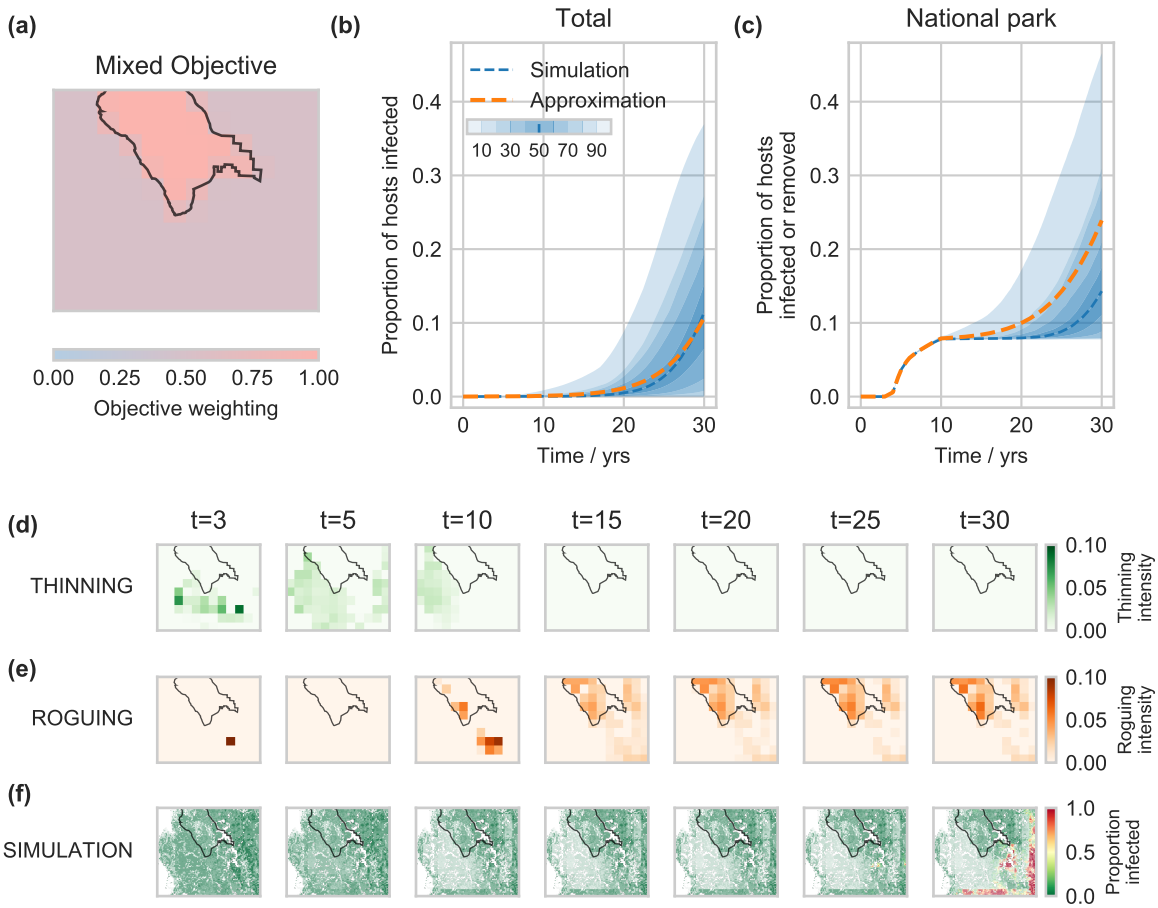

Figure 7.9: **(a)** The Mixed objective protects the national park, but with some value given to hosts outside. The disease progress curves across the landscape are shown in **(b)**, and the proportion of hosts infected or removed within the national park only in **(c)**. There is more disease than using the NP objective. The control strategy shown in **(d)** and **(e)** only thins in the environment conducive to sporulation. More roguing than thinning is carried out to the east, leading to more infection but also more host retained (**(f)**).

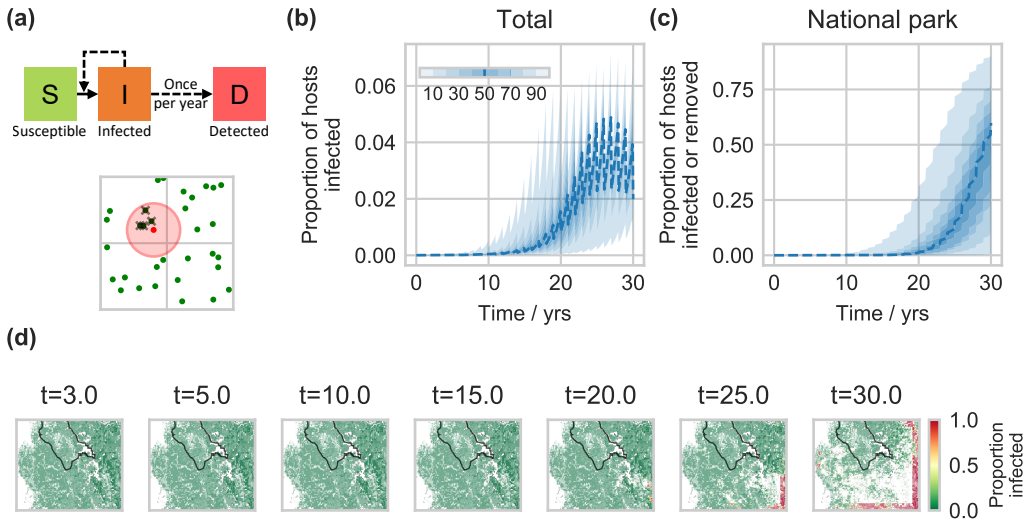

Figure 7.10: To implement the 100 m buffer strategy, a detected class is included, as shown in **(a)**. Once per year infected hosts are detected by a Bernoulli trial with probability of success $p = 0.7$. The 100 m buffer removes detected hosts, and all hosts within 100 m of the detection once every year, as illustrated by the red circular buffer. The strategy keeps levels of infection low (**(b)**) as it directly removes these hosts, but can lead to high numbers of hosts removed in the national park (**(c)**). As shown in **(d)**, large numbers of hosts are removed across the landscape.

7.5.3 Comparing strategies

The different control strategies vary in the pattern of hosts saved across the landscape (Figure 7.11(a)). The NP and Mixed objectives very effectively protect the national park, whereas without control spatially optimised to protect it, the non-spatial and 100 m buffer strategies cannot protect the national park effectively. In the cases where it fails to control the spread, the 100 m buffer strategy also results in more hosts culled than any of the other strategies (Figure 7.11(b)). Whilst this strategy does not remove many hosts early in the epidemic, if disease persists then very large areas of host may need to be controlled. Under the OCT strategies however, more hosts are removed pre-emptively, but this does result in improved disease management.

The OCT strategies (Total, NP, Mixed and Non-Spatial) all show a similar time dependence in allocation to thinning or roguing controls (Figure 7.11(c)), but with dependence on spatial location. Initially control is split between both methods. Thinning dominates the start of the epidemic, before switching to roguing after approximately 10 years. On average, the highest number of hosts outside the national park are protected by the Mixed strategy (Figure 7.11(d)). Hosts within the national park are best protected by the NP strategy, with the Mixed strategy finding a good balance of protection both in and out of the park.

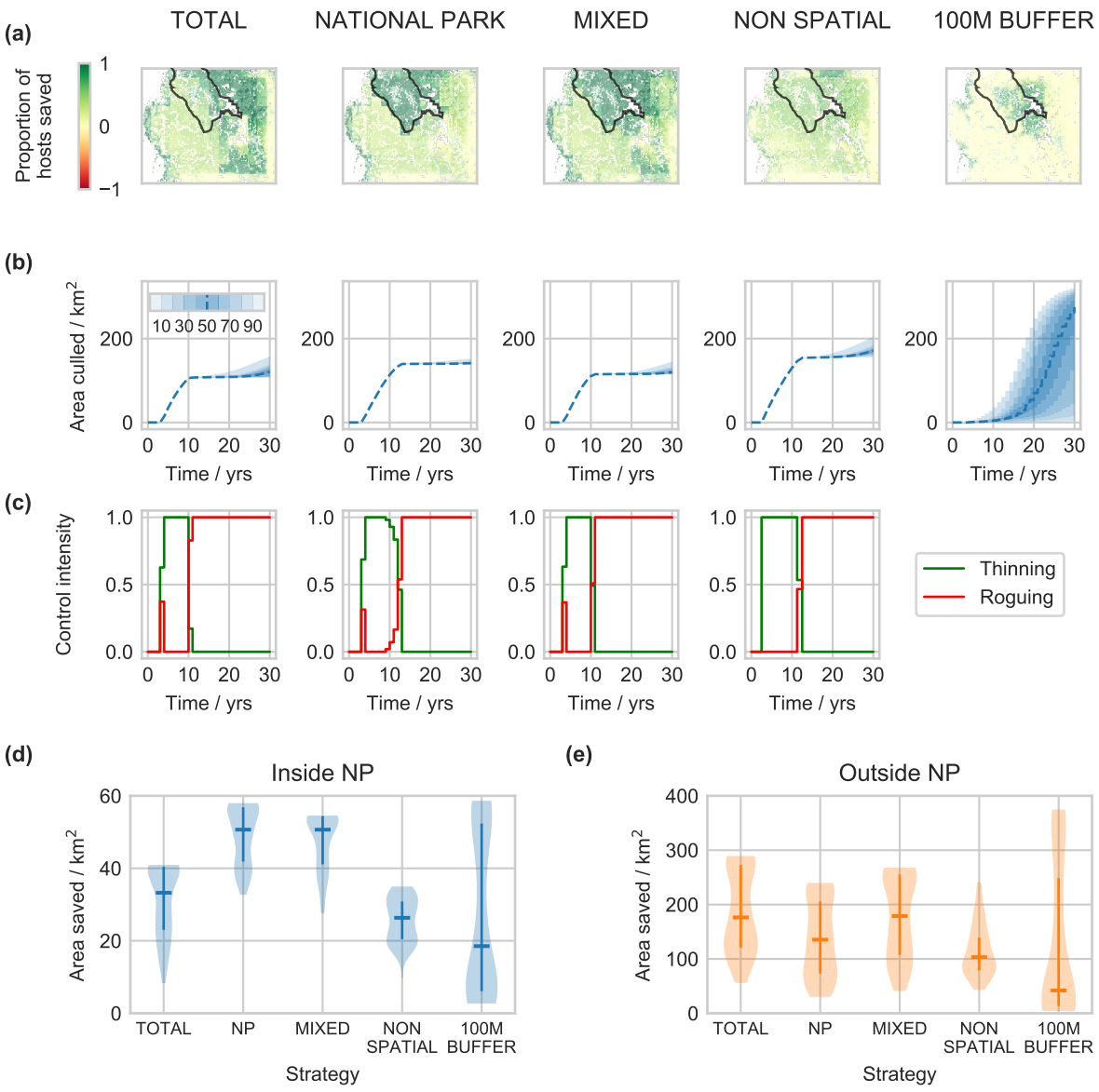

Figure 7.11: Comparison of all control strategies. In **(a)** the proportion of hosts saved relative to simulations under no control are shown for each strategy, with the national park outlined in black. **(b)** shows the area of host culled for each strategy, and **(c)** shows the levels of thinning and roguing for each of the OCT strategies. **(d)** and **(e)** give the area of host saved inside and outside the national park respectively. Horizontal lines indicate the median, and vertical lines span the 25th to 75th percentiles. In some cases the 100 m buffer strategy contains the epidemic and saves the most hosts, but can also result in culling large numbers of hosts. The NP and Mixed strategies protect the national park most effectively. Note that the OCT strategies are mostly capped with a maximum area saved. This is because the removal of trees by the control intervention limits the area of forest it is possible to save. The lower limits indicate the maximum rate of invasion under the control strategy.

7.5.4 Resolution choice

We here test how appropriate our choice of the 2.5 km resolution approximate model was. The motivation for this choice was that this model is the highest resolution for which the NLP optimisation is tractable. Higher spatial resolution allows more spatially detailed strategies, and hence improved protection of the national park. Figure 7.12 compares control performance of the 2.5 km model with a 5 km resolution model. Control in the lower resolution model is less spatially resolved around the national park. This leads to fewer hosts saved, both inside and outside the national park. It is clear that for effective control a higher resolution model is preferable, and so within our methodological limitations the 2.5 km resolution approximate model is best.

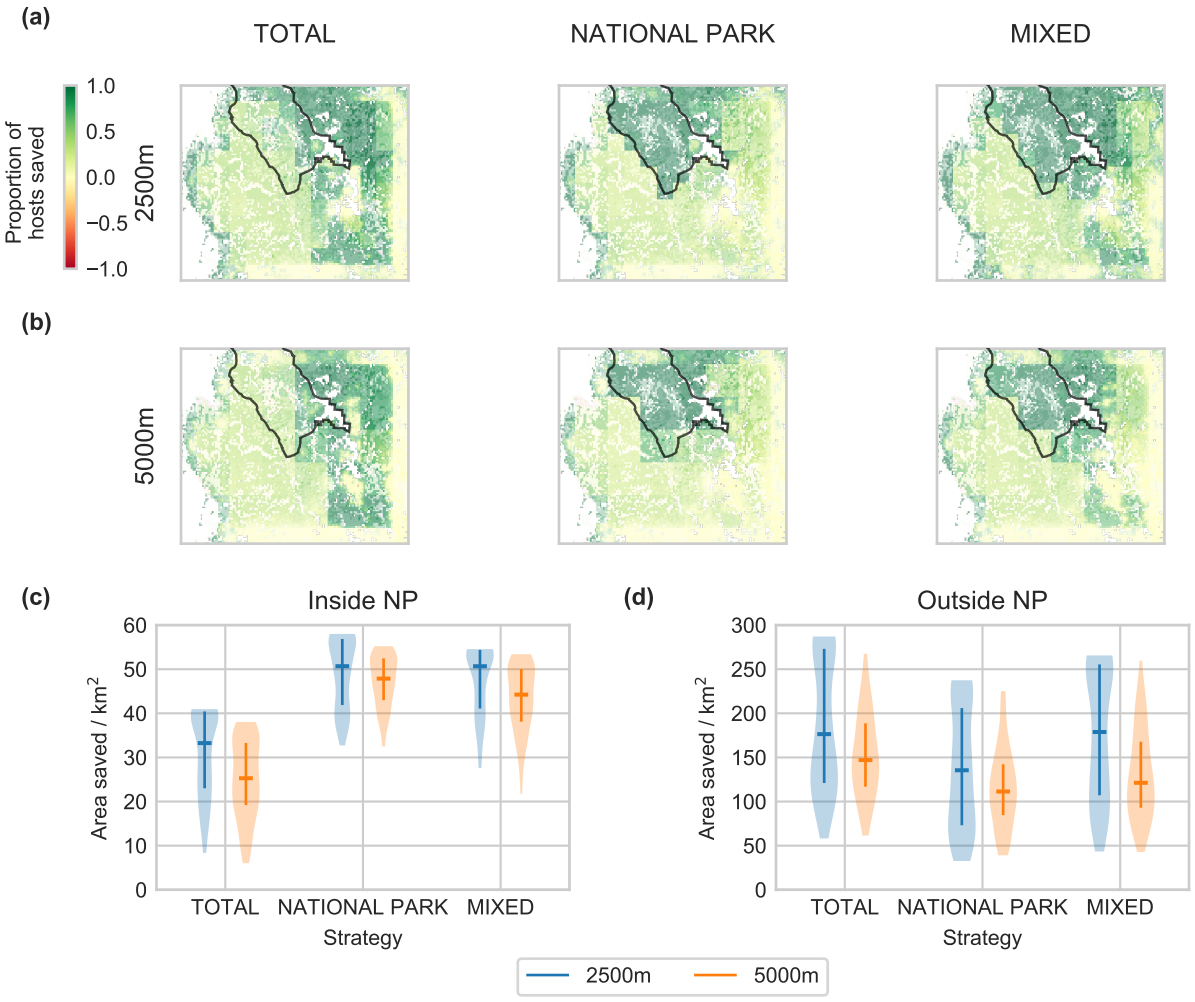

Figure 7.12: Comparison of optimal control using the 2.5 km and 5 km approximate models. **(a)** shows the proportion of hosts saved by each OCT control strategy using the 2.5 km resolution approximate model (i.e. repeats part of Figure 7.11(a)), with the same for the 5 km model shown in **(b)**. **(c)** and **(d)** show the area of host saved inside and outside the national park respectively, for each model and for each spatial OCT strategy. The higher resolution 2.5 km model saves more hosts on average because the control can be more spatially detailed, and hence it protects the national park better.

7.6 Discussion

In this chapter we have shown how spatially optimised strategies can protect valuable regions from disease. Using OCT to inform spatial control improves disease management, and better protects Redwood National Park from SOD than the control that was actually applied. The optimal strategies make use of epidemiological features which were ignored by the management carried out in practice. In particular, the optimal spatial strategies focus on thinning controls in areas where weather conditions are most conducive to pathogen spread.

7.6.1 Spatial optimisation

To allow for spatial optimisation of control, we used a reduced resolution deterministic model approximating the dynamics of the fully spatially-explicit simulation model. Capturing the full spatial detail of the simulation model results in a model for which optimisation is not tractable. In a similar setting, work by Epanchin-Niell and Wilen (2012) optimised spatial strategies for control of invasive species, but with much greater simplifications to the model dynamics. To allow spatial optimisation, Epanchin-Niell and Wilen (2012) reduce the underlying spread model to an integer problem where cells are either invaded or not, and the species potentially spreads to nearest neighbour cells in each time step. In Chapters 3 and 4 we simplified spatial dynamics by modelling the system with two or three metapopulations. By applying large-scale direct optimisation to the spatial optimisation problem, we have significantly reduced the simplifications that are necessary, and extended the optimal control to a discretised continuous landscape. This means that the resulting strategies are driven more closely by real-world dynamics, and so will be more effective if applied in practice. In testing the strategies on a fully spatially-explicit model—by applying the control lifting framework used in Chapters 4 and 6—we have demonstrated how these strategies would perform in reality, and shown that they outperform the management that was actually applied.

The constraint we used in this chapter partitioned control resources between thinning and roguing across the landscape. We did not use a constraint limiting the number of hosts controlled, as we used in Chapters 3 and 6 for example, because for an NLP problem of this scale these mixed constraints slow convergence. Including a mixed constraint could change the optimal strategy, but the general spatial structure of control, and the partitioning between thinning and roguing is unlikely to be significantly different. The optimal spatial strategies we found here show similar features to the results of previous chapters. We see a switching strategy that switches from thinning in a region that would threaten the national park, to prioritising roguing within the park. This is very similar to the switching strategies

found in Chapter 3, where a buffer region was first prioritised, before switching to control in the high value region.

Other methods exist for finding an optimal spatial control, including using OCT on PDE models of disease spread rather than the ODE approximate models we have used here. PDE models have been used previously to model plant disease epidemics (e.g. White and Gilligan, 2006). Reaction-diffusion type models can be used to capture the spatio-temporal dynamics of disease spread, and OCT can be applied to the resulting model to optimise spatial controls. For example, work by Neilan and Lenhart (2011) optimised spatial deployment of rabies vaccines to wild raccoons. They also found that optimal strategies make use of landscape features, in their case taking advantage of natural barriers to spread. Miyaoka et al. (2019) also optimise a PDE model of disease spread, there a model of Zika virus spread and optimal vaccine deployment. Whilst these studies use OCT to find spatial strategies, similar to the optimisations we have carried out, neither study tests the results on a more detailed simulation model. The advantage of our approach is that the effectiveness of the control can be shown under more realistic conditions, and control is designed to protect a high value region. Here we showed that the OCT strategies are still effective at protecting Redwood National Park under the stochastic spread in the simulation model.

7.6.2 Spatial resolution

We optimised spatial controls in this chapter using a 2.5 km resolution approximate model. This key choice of resolution was a balance between increasing spatial detail and a tractable optimisation problem. With the optimiser and NLP problem we use here, a higher resolution model is not solvable. We showed however, that lower resolution models result in less effective optimal spatial controls because of the lack of spatial detail. Whilst higher resolution models would be preferable, and could be used with sufficient computational resources, we here use the highest resolution model possible with our setup.

7.6.3 Practical implementation

Using the simulation model to test the OCT results allows potential practical outcomes of strategies to be compared. The OCT strategies tested here perform best on average, and protect the whole region against the very worst epidemics, where the 100 m buffer strategy results in removal of very large numbers of hosts. However, the strategies do result in large scale removal of hosts across the landscape. In the reduced resolution approximate model this extensive control is required to manage disease, but in the stochastic simulation—and by extension the real world—this may not be absolutely necessary. Implementing large scale removal can be difficult, particularly if centralised management leads to conflicts with local landowners who may rely on the forest for income (Alexander and Lee, 2010). Opposition

from affected parties can impact on deployment of control, as was the case with citrus canker (Gottwald, 2007). As shown by the testing of the 100 m buffer strategy, in some cases the disease does not spread rapidly and considerably lower levels of control are sufficient.

Deciding on the scale of control to carry out, or when to switch to larger scale control, is a difficult problem, particularly when estimates of future epidemic size can be very uncertain (Neri et al., 2014). With citrus canker, choosing this scale incorrectly eventually led to failure of control (Gottwald, 2007). Our models have shown that the more extensive OCT strategies perform better than the 100 m buffer on average, but there is a trade-off between implementing simpler controls more rapidly, and carrying out the optimal strategy with more knowledge of the system (Thompson et al., 2018). Whilst these decisions must be made in collaboration with policy makers and landowners, we have demonstrated that larger scale control is necessary to give the most effective control on average. For comparison, management of the isolated outbreak in Oregon treated approximately 2,510 ha of land between 2001 and 2015, costing $5000–15000 per hectare (Goheen et al., 2017). The total area of the approximate model we consider here is 75,000 ha, so the level of control necessary is extensive.

We have not accounted in our model for the importance of Redwood Creek. *P. ramorum* can spread through the watercourse, and this could lead to increased rates of spread into the national park. It could also allow for more effective surveillance and control strategies that prioritise areas along the river. Our work in this chapter has shown how the pathogen suitability of the landscape can influence optimal strategies, here through prioritising thinning where the weather is most conducive to pathogen spread. Similar methods could improve control when additional spread through the river is taken into account, but additional data would be necessary to fully inform this type of strategy. There is currently little information about the rate of *P. ramorum* spread through watercourses.

The broader message of the OCT strategies here though, is that management at larger scales is necessary to be sure of effective disease control. The 100 m buffer was clearly too short to protect Redwood National Park. This is consistent with work by Cunniffe et al. (2016) that found the optimal radius of treatment to be 187.5 m, but their result does depend on how regions are prioritised for treatment and the risk aversion of the decision maker. We also see in our results the benefit of front-loading identified by Cunniffe et al. (2016), i.e. allocating additional control resources earlier in the epidemic. Thinning resources in all OCT strategies are carried out early during management, in order to slow the spread of infection in the most susceptible regions. The control is carried out ahead of the epidemic wavefront in anticipation of future spread. The 100 m buffer strategy, on the other hand, is purely reactive to the ongoing epidemic, and as a result must try to catch up with the

epidemic wavefront. The benefit of this concept of front-loading has also been seen in other optimal control theory studies (Behncke, 2000).

7.6.4 MPC

A further extension to this work would be to include the feedback framework of MPC developed in Chapter 4 and used in Chapter 6. This could help to tailor the OCT strategies to disease progression, reducing levels of control when the epidemic is slow to spread. As we discussed previously, the 100 m buffer strategy shows that lower levels of control can be sufficient. The MPC framework could help adapt to an individual epidemic, to increase or decrease control intensity as required. Similar to the results in Chapter 4, this could improve control under the stochasticity of the simulation model. Allowing parameter inference between MPC update times could further improve control by updating the approximate model as the epidemic proceeds (Thompson et al., 2018).

7.7 Conclusions

In conclusion, this chapter has demonstrated the extension of the optimal control methods we have developed to optimising spatial strategies. These optimal spatial strategies can protect high value regions threatened by disease more effectively. The OCT strategies perform better on average than the control that was actually carried out, and protect against the worst-case scenario epidemics. The approaches we use here improve the resolution at which plant disease management strategies can be optimised compared with previous studies (c.f. Forster and Gilligan, 2007), and allow for more detail in the underlying spread model (c.f. Epanchin-Niell and Wilen, 2012).

8 Discussion

8.1 Thesis Summary

The threat from plant disease epidemics to the natural environment is severe (Freer-Smith and Webber, 2017), but mathematical models can be used to help design and optimise management strategies to better control spread (Lofgren et al., 2014; Cunniffe et al., 2015). The complexity of the simulation models used to inform policy however (e.g. DEFRA, 2014a), limits the extent to which dynamic controls can be optimised. For such models, multiple controls that vary in space and time can only be optimised through strategy testing: where a limited number of simple, plausible interventions are tested using a simulation model. In this thesis, our aim was to develop frameworks to aid optimisation of these simulation models, applied to the case study of determining effective and practical local management of SOD.

In Chapter 3 we showed how OCT can be applied to a simple model of a disease invading a highly valuable region. We found that in general it is best to prioritise control in the valuable region, but that switching strategies can perform better under certain parameter values. We tested direct and indirect numerical methods for solving the optimal control problem, showing that the direct method is accurate and more reliable than the indirect approach. The optimal strategy was not necessarily robust though, with significant errors when spread parameters were not known precisely. This motivates—at least in part—the problem solved in Chapter 4, where we developed a framework for applying optimal control results to complex simulation models. By using a feedback framework, model predictive control (MPC), we showed how the predictive power of simulation models can be coupled with insights from OCT obtained from approximate models to find effective disease management strategies.

In Chapters 5 and 6 we applied this framework to the practical question of how to control SOD in a newly invaded forest stand in order to protect valuable tanoak. This objective for local control is important, since widespread eradication of SOD is now impossible (Cunniffe et al., 2016), and so smaller-scale local goals that remain achievable and protect valuable resources are of interest. We showed how MPC finds strategies that align with management advice from the US Forest Service, but with additional important time dependence. The MPC framework finds strategies that are robust to parameter and observational uncertainty,

showing the greatest improvement in control in the worst-case scenarios. Finally, in Chapter 7 we applied OCT to spatial optimisation of control strategies to protect Redwood National Park from SOD invasion. We showed how the strategies that perform best make use of epidemiological features, such as prioritising control where weather conditions are most conducive to pathogen spread. The OCT informed management performs better than the 100 m buffer strategy that was actually applied. Again, we showed that the OCT strategies limit the impacts of disease in the worst-case scenarios.

8.2 Contributions and limitations

8.2.1 Optimal control theory

The overarching theme of this thesis is how to apply OCT to the problem of disease management, determining which control methods are most appropriate and how, when and where they should be deployed. The time-dependent strategies identified using OCT allow deployment of different control resources as an epidemic progresses. This switching, either between control methods or changing where control is applied, ensures that the optimal strategy is used at all times, allowing management to respond to changing risk (Hyatt-Twynam et al., 2017) and to always match the scale of the epidemic (Gilligan et al., 2007). We have shown how this improves management, echoing the results of other OCT and plant disease modelling studies (e.g. Forster and Gilligan, 2007; Ndeffo Mbah and Gilligan, 2011), as well as for other disease systems (Keeling and Shattock, 2012; Klepac et al., 2012), and control of invasive species (Carrasco et al., 2009).

Many different methods can be used to solve these optimal control problems, and we have tested both direct and indirect approaches. We tested the indirect forward-backward sweep method and the direct transcription method, chosen as they are both widely used to solve OCT problems. We did not test other approaches, such as multiple shooting methods, since the direct transcription method was sufficiently accurate, robust and reliable for the models considered here. Whilst the indirect methods do allow for more mathematical insight through analytic descriptions of the optimal control, applying the direct methods here allowed us to reliably and robustly solve larger problems where convergence of indirect methods would be limiting. The optimisation approaches we used allowed us to optimise deployment of many different control methods, with OCT identifying which of 9 different controls is optimal for control in a mixed stand.

Ultimately this led to optimising control of a fully spatial system around Redwood National Park, allowing the question of where control resources should be deployed to be answered with sufficiently resolved spatial strategies. Other studies have also optimised

similar spatial management strategies, for example work by Epanchin-Niell and Wilen (2012). The model used by Epanchin-Niell and Wilen (2012) though, required significant simplification to allow optimisation, with a discrete nearest-neighbour spread model in which cells are either invaded or not. Our work on the other hand, only reduced the resolution of the model to be optimised. Different approaches apply OCT to PDE models of disease spread, for example work by Neilan and Lenhart (2011), Christley et al. (2016), and Miyaoka et al. (2019). The reaction-diffusion type models used by these studies solve a similar problem to the problem of Chapter 7, and the work by Miyaoka et al. (2019) finds strategies that are significantly more spatially resolved than ours. These approaches could be used to extend our models, but importantly, none of these studies test results on more realistic models. Also, there is more scope with metapopulation type models to incorporate long distance dispersal, so that disease can still invade a fragmented landscape. Whether an ODE or PDE model is used for optimisation however, the simplifications made in these approximate models—to spatial structure for example—necessitate robustness testing of the optimal strategies.

This testing of OCT strategies can result in poor performance when there is uncertainty about parameters or the system is highly complex, as we showed in Chapters 3 and 4. Work by Forster and Gilligan (2007) showed that knowing the precise optimal time for a switch in an OCT strategy is important, and that inaccuracies can lead to management that is worse than using a much simpler strategy. Similarly, Carrasco et al. (2009) showed for control of invasive species that the optimal strategy is not always robust once parameter uncertainty is introduced. This and our comparisons of the open-loop and MPC strategies highlight the importance of simulation models for making accurate disease forecasts. For informing practical control, realistic complex models are needed to test and assess potential interventions. When data is limited though, these simulation models may not be as accurate as is required for policy. An advantage of a simulation model however, is that uncertainties can be included, for example through sampling from a parameter distribution (e.g. Cook et al., 2008; Parry et al., 2014; Cunniffe et al., 2015).

8.2.2 Optimising complex systems

Our approach to optimising complex simulation models was to start with a simpler approximate model for which the optimal control problem is more tractable: an approach commonly used in applications of OCT to engineering such as chemical plant control (Lee, 2011). The biggest factor affecting the size, and hence difficulty, of the optimal control problem is the spatial resolution of the model. To reduce the size of the state-space, we factored out space either completely in Chapter 6, into metapopulations in Chapters 3 and 4, or used reduced resolution models in Chapter 7. These approaches retain spatial detail when required for the problem at hand, allowing OCT to give the most spatially resolved

strategies possible. However, as with most OCT models, these approximations do lead to differences between the approximate model and the complex simulation model used for testing, and our solutions have limited spatial detail.

The MPC framework for disease control we have developed accounts for these inaccuracies in the approximate model. Other studies have considered MPC for disease control (Sélley et al., 2015), as well as for other applications such as chemical plants and the fields of manufacturing and aerospace (Qin and Badgwell, 2003). However, ours is the first study to apply the full system, with approximate and simulation models, to practical disease management questions. The feedback strategies allow the insights of OCT to be applied, whilst also adapting control to each epidemic realisation, and lead to improved control over an open-loop application of OCT results. We have demonstrated that MPC is a suitable framework for asking how to deploy control, showing in multiple scenarios that MPC provides the best strategies.

The main benefit of MPC is in the robustness of the framework to stochasticity, parameter uncertainty and imperfect surveillance. We have shown that adapting control through MPC handles these complexities, and minimises the risk of large-scale failures. We explained in Chapter 1 how poorly designed management strategies have led to the failure of control for dutch elm disease, citrus canker, and ash dieback. The strategies we have identified could limit failures such as these through adapting control. We have shown that MPC is of the most benefit during the most damaging epidemics. In Chapter 6 we saw that under parameter uncertainty MPC provides the greatest benefit in the worst-case scenarios. Similarly, the OCT strategies in Chapter 7 improve control in the worst-case epidemics. Finding robust control strategies is vital for effective decision making.

8.2.3 Objectives for disease management

Assessing strategies for potential policy decisions requires a clear quantification of the management goals; decision makers must ask why control is needed. For SOD, and arguably for management of plant diseases in the natural environment more generally, the consideration of solely large scale objectives has limited the utility of mathematical models. Here we have made progress by using management goals that focus on local disease control, prioritising protection of high value resources whether they be particular species or particular regions. These local goals have been shown to be achievable in the field (Hansen et al., 2019), and are still beneficial because of the high value of forest resources (UK government, 2018). Our frameworks apply the predictive power of simulation models and the insights of OCT to these management goals, and we have demonstrated the effectiveness of the resulting strategies.

The objective we used in Chapter 6 combined the goals of protecting tanoak and conserving biodiversity. This highlighted the importance of considering wider management goals when optimising disease control, since the choice of objective impacts on optimal strategies (Probert et al., 2016). However, our inclusion of biodiversity was still fairly simplistic, and other goals such as fire management and timber value were not accounted for. Nevertheless, we showed how goals such as these can alter the optimal control strategy, avoiding complete removal of bay and redwood. Similarly, in Chapter 7 control was optimised to protect the national park. But the objective did not account for the proximity of tribal lands, or commercially valuable stands. Whilst there is more scope in plant disease, as opposed to human or animal health, to protect one area or host at the detriment of others, wider implications must still be considered. This only highlights further the importance of the simulation model for assessing the resulting strategies. Any OCT informed strategies can be compared with other plausible interventions using the strategy testing approach. This allows the utility of the OCT informed strategies to be demonstrated as we did throughout this thesis, but also allows other consequences to be considered even when not formally included in the objective function. The simulation model can be used as a proxy for the real world, to illustrate the wider impacts of disease and control.

8.2.4 Practical management

In this thesis we have answered the questions of which control methods to use, and where, when, how and why to deploy these for disease management. Combining these solutions allows practical management advice. Using the frameworks we have developed for practical advice requires a fine balance between the realism of the simulation model, the level of approximation for optimising control, the choice of objective function and the constraints on the system. If all of these are aligned correctly then the resulting control strategies could provide practical and realistic advice.

The strategies we have found for local SOD management are arguably highly complex; the strategies show much time-dependence with many control switches and require a hierarchy of models. In their current form these results are unlikely to inform disease management in the real world because of the practical barriers to implementing such complex strategies. The results could still be of use though, through development of rules of thumb (e.g. Parnell et al., 2015; Hyatt-Twynam et al., 2017), or greater understanding of the drivers behind the optimal strategy. This could include optimal bay densities for thinning in mixed stands, or identification of areas that should be thinned and areas that should be rogued when protecting a valuable region. Information such as this could help inform practical management.

The variable level of surveillance in Chapter 6 allows a decision maker to vary the complexity of the MPC framework to suit the local disease risk. Regular updates are more important when there is a high risk of disease re-emergence, and when optimal control strategies are highly sensitive to changes in the system. We showed that in the mixed stand case the form of the optimal control was not sensitive to parameter changes, but because of the fine balance between tanoak retention and biodiversity conservation, small changes in the strategy did affect performance.

Alongside the complexity of the optimal strategies, there is the question of when the more complex MPC strategy is required. We showed the benefit of MPC when the approximate model degrades in accuracy, but also showed in Chapter 6 that when the approximate model closely matches the simulation throughout the epidemic (Section 6.3.3, p. 129), there is little benefit to MPC over open-loop. The benefit of MPC is to account for inaccuracies when the approximate model is unable to capture the full simulation dynamics, but how complex should the approximate model be? Here we made the approximate model as complex as possible whilst ensuring the optimisation was still tractable, but more systematic methods could be developed. The biggest problem with MPC more broadly, is the need of an underlying model (Camacho and Bordons, 1995). This model must be simple enough to optimise but complex enough to capture the main disease characteristics. If there is significant variation or stochasticity in the real world, then the surveillance and re-optimisation of MPC will be necessary regardless of the quality of the approximate model.

8.3 Scope for future work

There is scope in future work to extend the complexity of the underlying approximate models. In particular, our approximate models either factor out the heterogeneity of space entirely by assuming well-mixed dynamics (Chapters 5 and 6), or include only a restricted resolution via metapopulation approximations (Chapters 4 and 7). However, in all cases well-mixing is assumed at some level. There is a wide body of work showing how spatial heterogeneities can be handled empirically by using non-linear incidence functions to better approximate spatial dynamics (e.g. Liu et al., 1986; Clarke et al., 2013; Chowell et al., 2016a; Chowell et al., 2016b). However, since these forms are highly empirical, it is important to ensure they fit simulation dynamics well under different time-dependent control strategies. Doing so adaquately under the full range of possible controls with a single approximation might even prove impossible, as we found for the approximate model in Chapter 6 where control rates had to be scaled to match the control scheme. We also found that such complexity in the underpinning model can hinder optimisation convergence, particularly when derivatives are calculated through automatic differentiation, as was the case when the package BOCOP was used in Chapters 4 and 6.

Although here we used both network and structured metapopulation models for simulations, alternative approaches to modelling space could be used, both in the approximate model and the simulation model. For example, PDE type models as discussed in Section 8.2.1 (e.g. White and Gilligan, 2006; Neilan and Lenhart, 2011; Miyaoka et al., 2019) could be used as approximate models, potentially allowing for higher resolution spatial strategies because there would be no requirement to quantise space. Alternatively, the spatial detail of the simulation model could be improved by using individual-based models (e.g. Cook et al., 2008; Cunniffe et al., 2015), capturing the dynamics of each host individually, although this could be overly detailed. This would also allow the raster approximation we used in Chapter 7 to be tested, determining the appropriate resolution of the simulation model.

We have only tested a limited set of optimisation approaches in this thesis. In particular, we have used deterministic OCT to find robust control strategies for simulation models that in some cases were stochastic. It would be interesting to extend these analyses to use stochastic OCT (Bertsekas, 2001), and to compare the resulting optimal strategies with the deterministic analogues. Stochastic OCT can find more robust control strategies by incorporating noise into the differential equations, but this does significantly increase the complexity of the problem and only certain forms of noise can be included. In Chapter 6 we touched on optimising multiple objectives, there combining tanoak retention and biodiversity conservation. This could be extended to explicitly analyse multi-objective controls, including wider disease impacts such as fire risk and CO_2 capture in the OCT objective. However, combining multiple objectives meaningfully can be difficult (Probert et al., 2016).

It would also be interesting to investigate optimisation approaches other than OCT. Reinforcement learning (RL) is a field of machine learning concerned with computational agents learning actions to maximise rewards, and has many parallels with OCT (Recht, 2019). Where OCT determines the optimal strategy based on a model of the dynamics, RL approaches learn the underlying model whilst optimising the objective function. Recent advances have shown that RL can solve very large problems, most notably defeating a human professional player at the game of Go, for which exhaustive searching of solutions is infeasible (Silver et al., 2016). Using these approaches could find novel strategies in complex simulation models, solving problems with significantly more spatial detail for example. The approach has recently been shown to be effective in a proof of concept application to control of foot and mouth disease in animals (Probert et al., 2019). The disadvantage of this approach is a lack of understanding of how the decisions are made. This could hamper translation into policy.

We emphasised throughout this thesis the importance of robust control strategies, i.e. strategies that under stochasticity and uncertainty still ensure effective disease management. There is scope to extend this further in our work to more formally optimise robust strategies.

The strategies we have developed could be further tested under parameter uncertainties, potentially including estimation of parameters (Parry et al., 2014), as well as more precise estimates as the epidemic progresses (Thompson et al., 2018). Another potential avenue would be to use Robust MPC (Bemporad and Morari, 1999), an approach that explicitly accounts for uncertainty in the controller by testing worst-case scenarios. This could provide guarantees of management performance under defined levels of uncertainty, limiting the worst-case scenarios in a more formal framework than we have so far considered.

Finally, our work could be applied to other disease systems where control may be more effective than is now possible with SOD, for example management of the ongoing epidemic of olive quick decline syndrome in southern Europe caused by *Xylella fastidiosa* (Sicard et al., 2018), or to diseases with major food security implications such as cassava brown streak disease in East Africa (Legg et al., 2011) or maize lethal necrosis (Hilker et al., 2017). Additional real-world complexities could also be included into the models, for example the effects of human behaviour and non-compliance with disease management. Our spatial optimisations could be extended to a landscape of individual decision makers with contrasting management goals similar to the systems considered by Epanchin-Niell et al. (2010) and Milne et al. (2015) for example. Management could then be optimised to provide incentives and dis-incentives to align individual goals with the aim of wider disease management.

8.4 Concluding remarks

The setting of this thesis has been to combine methods from the fields of mathematics, biology and systems engineering. We have made progress in the difficult problem of optimising controls in complex epidemiological simulation models that capture enough realism to inform policy. The frameworks we have developed go some way to building robust decision making processes for plant disease control, but our work is only a start to applying approaches from systems engineering to epidemiology. There is still much scope to build more complex, more robust, and more rigorous approaches. With increasing disease threats, the effects of climate change, and a growing and increasingly global population, the significance and necessity of improving these methods is clear.

Bibliography

Adrakey, HK, G Streftaris, NJ Cunniffe, et al. (2017). Evidence-based controls for epidemics using spatio-temporal stochastic models in a Bayesian framework. *Journal of the Royal Society Interface* **14**, 20170386.

Al Basir, F, PK Roy, and S Ray (2017). Impact of roguing and insecticides spraying on mosaic disease in *Jatropha Curcas*. *Control and Cybernetics* **46**, pp. 325–344.

Aldrich, J (1997). R. A. Fisher and the making of maximum likelihood 1912–1922. *Statistical Science* **12**, pp. 162–176.

Alexander, J and CA Lee (2010). Lessons learned from a decade of sudden oak death in California: evaluating local management. *Environmental Management* **46**, pp. 315–328.

An, G, BG Fitzpatrick, S Christley, et al. (2017). Optimization and control of agent-based models in biology: a perspective. *Bulletin of Mathematical Biology* **79**, pp. 63–87.

Anderson, PK, AA Cunningham, NG Patel, et al. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, pp. 535–544.

Anderson, RM, GF Medley, RM May, and AM Johnson (1986). A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *Mathematical Medicine and Biology* **3**, pp. 229–263.

Baker, RHA, H Anderson, S Bishop, et al. (2014). The UK plant health risk register: a tool for prioritizing actions. *EPPO Bulletin* **44**, pp. 187–194.

Bansal, S, BT Grenfell, and LA Meyers (2007). When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface* **4**, pp. 879–891.

Basu, S and J Andrews (2013). Complexity in mathematical models of public health policies: a guide for consumers of models. *PLOS Medicine* **10**, e1001540.

Bateman, IJ, AR Harwood, GM Mace, et al. (2013). Bringing ecosystem services into economic decision-making: land use in the United Kingdom. *Science* **341**, pp. 45–50.

Behncke, H (2000). Optimal control of deterministic epidemics. *Optimal Control Applications and Methods* **21**, pp. 269–285.

Bellan, SE, JR Pulliam, CA Pearson, et al. (2015). Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis. *The Lancet Infectious Diseases* **15**, pp. 703–710.

Bellman, R (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bemporad, A and M Morari (1999). Robust model predictive control: a survey. In: *Robustness in Identification and Control*. London, UK: Springer, pp. 207–226.

Bertsekas, DP (2001). *Dynamic Programming and Optimal Control*. Vol. 2. Belmont, MA: Athena Scientific.

Betts, J (2010). *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Philidelphia, PA: SIAM.

Blayneh, K, Y Cao, and HD Kwon (2009). Optimal control of vector-borne diseases: treatment and prevention. *Discrete and Continuous Dynamical Systems Series B* **11**, pp. 587–611.

Bokil, VA, LJS Allen, MJ Jeger, and S Lenhart (2019). Optimal control of a vectored plant disease model for a crop with continuous replanting. *Journal of Biological Dynamics* **13**, pp. 325–353.

Bolker, BM (2008). *Ecological Models and Data in R*. Princeton, NJ: Princeton University Press.

Bowcutt, F (2013). Tanoak landscapes: tending a Native American nut tree. *Madroño* **60**, pp. 64–87.

Boyd, IL, PH Freer-Smith, CA Gilligan, and HCJ Godfray (2013). The consequence of tree pests and diseases for ecosystem services. *Science* **342**, 1235773.

Brasier, CM (2008). The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathology* **57**, pp. 792–808.

Brasier, C and J Webber (2010). Plant pathology: sudden larch death. *Nature* **466**, pp. 824–825.

Brown, L and B Allen-Diaz (2009). Forest stand dynamics and sudden oak death: mortality in mixed-evergreen forests dominated by coast live oak. *Forest Ecology and Management* **257**, pp. 1271–1280.

Brown, VL and KAJ White (2011). The role of optimal control in assessing the most cost-effective implementation of a vaccination programme: HPV as a case study. *Mathematical Biosciences.* **231**, pp. 126–134.

Bryson, AE (1996). Optimal control—1950 to 1985. *IEEE Control Systems Magazine* **16**, pp. 26–33.

Bussell, EH, CE Dangerfield, CA Gilligan, and NJ Cunniffe (2019). Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Philosophical Transactions of the Royal Society B* **374**, 20180284.

Byrd, RH, P Lu, J Nocedal, and C Zhu (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **16**, pp. 1190–1208.

Cadotte, MW, K Carscadden, and N Mirotchnick (2011). Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* **48**, pp. 1079–1087.

California Oak Mortality Task Force (2014). *Sudden Oak Death Guidelines for Arborists*. Available at: `http://www.suddenoakdeath.org/diagnosis-and-management/best-management-practices/` (visited on July 17, 2019).

Camacho, EF and C Bordons (1995). *Model Predictive Control in the Process Industry*. London, UK: Springer.

Carrasco, LR, R Baker, A MacLeod, JD Knight, and JD Mumford (2009). Optimal and robust control of invasive alien species spreading in homogeneous landscapes. *Journal of the Royal Society Interface* **7**, pp. 529–540.

Choi, YH, M Jit, N Gay, et al. (2010). Transmission dynamic modelling of the impact of human papillomavirus vaccination in the United Kingdom. *Vaccine* **28**, pp. 4091–4102.

Chowell, G, L Sattenspiel, S Bansal, and C Viboud (2016a). Mathematical models to characterize early epidemic growth: a review. *Physics of Life Reviews* **18**, pp. 66–97.

Chowell, G, D Hincapie-Palacio, J Ospina, et al. (2016b). Using phenomenological models to characterize transmissibility and forecast patterns and final burden of Zika epidemics. *PLOS Currents* **8**.

Christley, S, RM Neilan, M Oremland, R Salinas, and S Lenhart (2016). Optimal control of sugarscape agent-based model via a PDE approximation model. *Optimal Control Applications and Methods* **38**, pp. 473–497.

Clark, CW and GR Munro (1975). The economics of fishing and modern capital theory: a simplified approach. *Journal of Environmental Economics and Management* **2**, pp. 92–106.

Clarke, J, KAJ White, and K Turner (2013). Approximating optimal controls for networks when there are combinations of population-level and targeted measures available: chlamydia infection as a case-study. *Bulletin of Mathematical Biology* **75**, pp. 1747–1777.

Cobb, RC, JAN Filipe, RK Meentemeyer, CA Gilligan, and DM Rizzo (2012). Ecosystem transformation by emerging infectious disease: loss of large tanoak from California forests. *Journal of Ecology* **100**, pp. 712–722.

Cobb, RC, DM Rizzo, KJ Hayden, et al. (2013). Biodiversity conservation in the face of dramatic forest disease: an integrated conservation strategy for tanoak (*Notholithocarpus densiflorus*) threatened by sudden oak death. *Madroño* **60**, pp. 151–165.

Cobb, RC, P Hartsough, N Ross, et al. (2017). Resiliency or restoration: management of sudden oak death before and after outbreak. *Forest Phytophthoras* **7**, pp. 1–14.

Cobb, RC, N Ross, KJ Hayden, et al. (2019). Promise and pitfalls of endemic resistance for cultural resources threatened by *Phytophthora ramorum*. *Phytopathology* **109**, pp. 760–769.

Cook, AR, GJ Gibson, TR Gottwald, and CA Gilligan (2008). Constructing the effect of alternative intervention strategies on historic epidemics. *Journal of the Royal Society Interface* **5**, pp. 1203–1213.

Craig, AP, NJ Cunniffe, M Parry, FF Laranjeira, and CA Gilligan (2018). Grower and regulator conflict in management of the citrus disease Huanglongbing in Brazil: a modelling study. *Journal of Applied Ecology* **55**, pp. 1956–1965.

Cunniffe, NJ, FF Laranjeira, FM Neri, RE DeSimone, and CA Gilligan (2014). Cost-effective control of plant disease when epidemiological knowledge is incomplete: modelling Bahia bark scaling of citrus. *PLOS Computational Biology* **10**, e1003753.

Cunniffe, NJ, RO Stutt, RE DeSimone, TR Gottwald, and CA Gilligan (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLOS Computational Biology* **11**, e1004211.

Cunniffe, NJ, RC Cobb, RK Meentemeyer, DM Rizzo, and CA Gilligan (2016). Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. *Proceedings of the National Academy of Sciences* **113**, pp. 5640–5645.

David, J, H Tran, and HT Banks (2011). Receding horizon control of HIV. *Optimal Control Applications and Methods* **32**, pp. 681–699.

Davidson, JM, HA Patterson, and DM Rizzo (2008). Sources of inoculum for *Phytophthora ramorum* in a redwood forest. *Phytopathology* **98**, pp. 860–866.

Davidson, JM, AC Wickland, HA Patterson, KR Falk, and DM Rizzo (2005). Transmission of *Phytophthora ramorum* in mixed-evergreen forest in California. *Phytopathology* **95**, pp. 587–596.

Davis, FW, M Borchert, RK Meentemeyer, A Flint, and DM Rizzo (2010). Pre-impact forest composition and ongoing tree mortality associated with sudden oak death in the Big Sur region; California. *Forest Ecology and Management* **259**, pp. 2342–2354.

DEFRA (2014a). *Chalara Management Plan*. Available at: `https://www.gov.uk/government/publications/chalara-management-plan` (visited on Sept. 10, 2019).

DEFRA (2014b). *Tree Health Management Plan*. Available at: `https://www.gov.uk/government/publications/tree-health-management-plan` (visited on Aug. 29, 2019).

Epanchin-Niell, RS (2017). Economics of invasive species policy and management. *Biological Invasions* **19**, 3333—3354.

Epanchin-Niell, RS and A Hastings (2010). Controlling established invaders: integrating economics and spread dynamics to determine optimal management. *Ecology Letters* **13**, pp. 528–541.

Epanchin-Niell, RS and JE Wilen (2012). Optimal spatial control of biological invasions. *Journal of Environmental Economics and Management* **63**, pp. 260–270.

Epanchin-Niell, RS, MB Hufford, CE Aslan, et al. (2010). Controlling invasive species in complex social landscapes. *Frontiers in Ecology and the Environment* **8**, pp. 210–216.

Forrestel, AB, BS Ramage, T Moody, MA Moritz, and SL Stephens (2015). Disease, fuels and potential fire behavior: impacts of sudden oak death in two coastal California forest types. *Forest Ecology and Management* **348**, pp. 23–30.

Forster, G and CA Gilligan (2007). Optimizing the control of disease infestations at the landscape scale. *Proceedings of the National Academy of Sciences* **104**, pp. 4984–4989.

Fraser, C, S Riley, RM Anderson, and NM Ferguson (2004). Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* **101**, pp. 6146–6151.

Freer-Smith, PH and JF Webber (2017). Tree pests and diseases: the threat to biodiversity and the delivery of ecosystem services. *Biodiversity and Conservation* **26**, pp. 3167–3181.

Gamfeldt, L, T Snäll, R Bagchi, et al. (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature Communications* **4**, 1340.

Gao, F and L Han (2012). Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications* **51**, pp. 259–277.

Garbelotto, M and D Schmidt (2009). Phosphonate controls sudden oak death pathogen for up to 2 years. *California Agriculture* **63**, pp. 10–17.

Gillespie, DT (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**, pp. 2340–2361.

Gilligan, CA and F van den Bosch (2008). Epidemiological models for invasion and persistence of pathogens. *Annual Review of Phytopathology* **46**, pp. 385–418.

Gilligan, CA, JE Truscott, and AJ Stacey (2007). Impact of scale on the effectiveness of disease control strategies for epidemics with cryptic infection in a dynamical landscape: an example for a crop disease. *Journal of the Royal Society Interface* **4**, pp. 925–934.

Goheen, EM, A Kanaskie, S Navarro, and E Hansen (2017). Sudden oak death management in Oregon tanoak forests. *Forest Phytophthoras* **7**, pp. 45–53.

Goheen, EM, A Kanaskie, E Hansen, et al. (2010). Eradication effectiveness monitoring in Oregon tanoak forests. In: *Proceedings of the Sudden Oak Death Fourth Science Symposium*, pp. 233–235.

Gottwald, TR (2007). Citrus canker and citrus huanglongbing, two exotic bacterial diseases threatening the citrus industries of the Western Hemisphere. *Outlooks on Pest Management* **18**, pp. 274–279.

Gottwald, TR and M Irey (2007). Post-hurricane analysis of citrus canker II: predictive model estimation of disease spread and area potentially impacted by various eradication protocols following catastrophic weather events. *Plant Health Progress* **8**, 22.

Gottwald, TR, G Hughes, JH Graham, X Sun, and T Riley (2001). The citrus canker epidemic in Florida: the scientific basis of regulatory eradication policy for an invasive species. *Phytopathology* **91**, pp. 30–34.

Grünwald, NJ, M Garbelotto, EM Goss, K Heungens, and S Prospero (2012). Emergence of the sudden oak death pathogen phytophthora ramorum. *Trends in Microbiology* **20**, pp. 131–138.

Grünwald, NJ, JM LeBoldus, and RC Hamelin (2019). Ecology and evolution of the sudden oak death pathogen *Phytophthora ramorum*. *Annual Review of Phytopathology* **57**, pp. 301–321.

Hansen, EM, A Kanaskie, S Prospero, et al. (2008). Epidemiology of *Phytophthora ramorum* in Oregon tanoak forests. *Canadian Journal of Forest Research* **38**, pp. 1133–1143.

Hansen, E, P Reeser, W Sutton, et al. (2019). Efficacy of local eradication treatments against the sudden oak death epidemic in Oregon tanoak forests. *Forest Pathology*, e12530.

Hartl, RF, SP Sethi, and RG Vickson (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review* **37**, pp. 181–218.

Harwood, TD, X Xu, M Pautasso, MJ Jeger, and MW Shaw (2009). Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK. *Ecological Modelling* **220**, pp. 3353–3361.

Hilker, FM, LJ Allen, VA Bokil, et al. (2017). Modeling virus coinfection to inform management of maize lethal necrosis in Kenya. *Phytopathology* **107**, pp. 1095–1108.

Hill, L, G Jones, N Atkinson, et al. (2019). The £15 billion cost of ash dieback in Britain. *Current Biology* **29**, R315–R316.

Hocking, LM (1991). *Optimal control: an introduction to the theory with applications*. Oxford, UK: Clarendon Press.

Hovorka, R, V Canonico, LJ Chassin, et al. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement* **25**, pp. 905–920.

Hurst, JM, RB Allen, DA Coomes, and RP Duncan (2011). Size-specific tree mortality varies with neighbourhood crowding and disturbance in a montane *Nothofagus* forest. *PLOS One* **6**, e26670.

Hyatt-Twynam, SR, S Parnell, ROJH Stutt, et al. (2017). Risk-based management of invading plant disease. *New Phytologist* **214**, pp. 1317–1329.

Jeger, MJ, M Pautasso, O Holdenrieder, and MW Shaw (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist* **174**, pp. 279–297.

Jit, M, YH Choi, and WJ Edmunds (2008). Economic evaluation of human papillomavirus vaccination in the United Kingdom. *BMJ* **337**, a769.

Jones, E, T Oliphant, P Peterson, et al. (2001–). *SciPy: Open source scientific tools for Python*. Available at: `http://www.scipy.org/` (visited on July 29, 2019).

Kanaskie, A, E Hansen, W Sutton, P Reeser, and C Choquette (2011). Application of phosphonate to prevent sudden oak death in south-western Oregon tanoak (*Notholithocarpus densiflorus*) forests. *New Zealand Journal of Forestry Science* **41S**, S177–S187.

Keeling, MJ and CA Gilligan (2000). Metapopulation dynamics of bubonic plague. *Nature* **407**, pp. 903–906.

Keeling, MJ, MEJ Woolhouse, RM May, G Davies, and BT Grenfell (2003). Modelling vaccination strategies against foot-and-mouth disease. *Nature* **421**, pp. 136–42.

Keeling, MJ (2005). Models of foot-and-mouth disease. *Proceedings of the Royal Society B: Biological Sciences* **272**, pp. 1195–1202.

Keeling, MJ and P Rohani (2008). *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton University Press.

Keeling, MJ and A Shattock (2012). Optimal but unequitable prophylactic distribution of vaccine. *Epidemics* **4**, pp. 78–85.

Keeling, MJ, ME Woolhouse, DJ Shaw, et al. (2001). Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* **294**, pp. 813–817.

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

Keesing, F, LK Belden, P Daszak, et al. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**, 647.

Kelly, M, Q Guo, D Liu, and D Shaari (2007). Modeling the risk for a new invasive forest disease in the United States: an evaluation of five environmental niche models. *Computers, Environment and Urban Systems* **31**, pp. 689–710.

Kirschner, D, S Lenhart, and S Serbin (1997). Optimal control of the chemotherapy of HIV. *Journal of Mathematical Biology* **35**, pp. 775–792.

Kleczkowski, A and CA Gilligan (2007). Parameter estimation and prediction for the course of a single epidemic outbreak of a plant disease. *Journal of the Royal Society Interface* **4**, pp. 865–877.

Klepac, P, ON Bjørnstad, CJE Metcalf, and BT Grenfell (2012). Optimizing reactive responses to outbreaks of immunizing infections: balancing case management and vaccination. *PLOS One* **7**, e41428.

Kovacs, K, T Václavík, RG Haight, et al. (2011). Predicting the economic costs and property value losses attributed to sudden oak death damage in California (2010–2020). *Journal of Environmental Management* **92**, pp. 1292–1302.

Kowalski, T and O Holdenrieder (2009). Pathogenicity of *Chalara fraxinea*. *Forest Pathology* **39**, pp. 1–7.

Lee, C, Y Valachovic, and M Garbelotto (2010). *Protecting Trees from Sudden Oak Death before Infection*. Publication 8426. Richmond, CA: University of California Agriculture and Natural Resources.

Lee, JH (2011). Model predictive control: review of the three decades of development. *International Journal of Control, Automation and Systems* **9**, pp. 415–424.

Legg, JP, SC Jeremiah, HM Obiero, et al. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Research* **159**, pp. 161–170.

Lenhart, S and JT Workman (2007). *Optimal Control Applied to Biological Models*. New York: Chapman and Hall/CRC Press.

Liu, W, SA Levin, and Y Iwasa (1986). Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of Mathematical Biology* **23**, pp. 187–204.

Lofgren, ET, ME Halloran, CM Rivers, et al. (2014). Opinion: mathematical models: a key tool for outbreak response. *Proceedings of the National Academy of Sciences* **111**, pp. 18095–18096.

Long, J, A Gray, and F Lake (2018). Recent trends in large hardwoods in the Pacific Northwest, USA. *Forests* **9**, 651.

MacCleery, D (2015). Re-inventing the United States Forest Service: evolution from custodial management, to production forestry, to ecosystem management. In: *Re-inventing forestry agencies: Experiences of institutional restructuring in Asia and the Pacific*. Ed. by P Durst, C Brown, J Broadheada, et al. Bangkok: United Nations, FAO, Regional Office for Asia and the Pacific, pp. 45–77.

Maes, J, ML Paracchini, G Zulian, MB Dunbar, and R Alkemade (2012). Synergies and trade-offs between ecosystem service supply, biodiversity, and habitat conservation status in Europe. *Biological Conservation* **155**, pp. 1–12.

Magurran, AE (2004). *Measuring Biological Diversity*. Oxford, UK: Blackwell Science.

Maloney, PE, SC Lynch, S Kane, CE Jensen, and DM Rizzo (2005). Establishment of an emerging generalist pathogen in redwood forest communities. *Journal of Ecology* **93**, pp. 899–905.

Maloy, OC (2005). Plant disease management. *APSnet Education Center. The Plant Health Instructor* **10**.

Margosian, ML, KA Garrett, JS Hutchinson, and KA With (2009). Connectivity of the American agricultural landscape: assessing the national risk of crop pest and disease spread. *BioScience* **59**, pp. 141–151.

McDonald, PM and JC Tappeiner (2002). *California's hardwood resource: Seeds, seedlings, and sprouts of three important forest-zone species*. General Technical Report PSW-GTR-185. United States Department of Agriculture, Forest Service.

McPherson, BA, SR Mori, DL Wood, et al. (2010). Responses of oaks and tanoaks to the sudden oak death pathogen after 8 y of monitoring in two coastal California forests. *Forest Ecology and Management* **259**, pp. 2248–2255.

Meentemeyer, RK, NJ Cunniffe, AR Cook, et al. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in California (1990–2030). *Ecosphere* **2**, 17.

Meentemeyer, R, D Rizzo, W Mark, and E Lotz (2004). Mapping the risk of establishment and spread of sudden oak death in California. *Forest Ecology and Management* **200**, pp. 195–214.

Meentemeyer, RK, BL Anacker, W Mark, and DM Rizzo (2008). Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. *Ecological Applications* **18**, pp. 377–390.

Milgroom, MG and P Cortesi (2004). Biological control of chestnut blight with hypovirulence: a critical analysis. *Annual Review of Phytopathology* **42**, pp. 311–338.

Milne, AE, JR Bell, WD Hutchison, et al. (2015). The effect of farmers' decisions on pest control with Bt crops: a billion dollar game of strategy. *PLOS Computational Biology* **11**, e1004483.

Minor, ES and RH Gardner (2011). Landscape connectivity and seed dispersal characteristics inform the best management strategy for exotic plants. *Ecological Applications* **21**, pp. 739–749.

Miyaoka, TY, S Lenhart, and JFCA Meyer (2019). Optimal control of vaccination in a vector-borne reaction–diffusion model applied to Zika virus. *Journal of Mathematical Biology* **79**, pp. 1077–1104.

Morgan, O (2019). How decision makers can use quantitative approaches to guide outbreak responses. *Philosophical Transactions of the Royal Society B* **374**, 20180365.

National Park Service (2019). *RNSP Website*. Available at: `https://www.nps.gov/redw/index.htm` (visited on Aug. 21, 2019).

Ndeffo Mbah, ML and CA Gilligan (2011). Resource allocation for epidemic control in metapopulations. *PLOS ONE* **6**, e24577.

Ndeffo Mbah, ML and CA Gilligan (2010a). Balancing detection and eradication for control of epidemics: sudden oak death in mixed-species stands. *PLOS ONE* **5**, e12317.

Ndeffo Mbah, ML and CA Gilligan (2010b). Optimization of control strategies for epidemics in heterogeneous populations with symmetric and asymmetric transmission. *Journal of Theoretical Biology* **262**, pp. 757–763.

Ndeffo Mbah, ML and CA Gilligan (2014). Optimal control of disease infestations on a lattice. *Mathematical Medicine and Biology: A Journal of the IMA* **31**, pp. 87–97.

Neilan, RM and S Lenhart (2011). Optimal vaccine distribution in a spatiotemporal epidemic model with an application to rabies and raccoons. *Journal of Mathematical Analysis and Applications* **378**, pp. 603–619.

Neri, FM, AR Cook, GJ Gibson, TR Gottwald, and CA Gilligan (2014). Bayesian analysis for inference of an emerging epidemic: citrus canker in urban landscapes. *PLOS Computational Biology* **10**, e1003587.

Noss, RF (2000). *The Redwood Forest: History, Ecology, and Conservation of the Coast Redwoods*. Washington, D.C.: Island Press.

Oerke, EC (2006). Crop losses to pests. *The Journal of Agricultural Science* **144**, pp. 31–43.

Oregon Department of Forestry (2019). *Sudden Oak Death: Economic Impact Assessment*. Prepared by: Highland Economics | Mason, Bruce & Girard, Inc. Available at: `https://www.oregon.gov/ODF/Documents/ForestBenefits/sudden-oak-death-economic-impact-assessment.pdf` (visited on July 12, 2019).

Ostfeld, RS, GE Glass, and F Keesing (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* **20**, pp. 328–336.

Pandey, A, KE Atkins, J Medlock, et al. (2014). Strategies for containing Ebola in west Africa. *Science* **346**, pp. 991–995.

Panetta, JC and KR Fister (2003). Optimal control applied to competing chemotherapeutic cell-kill strategies. *SIAM Journal on Applied Mathematics* **63**, pp. 1954–1971.

Parnell, S, TR Gottwald, T Riley, and F Van Den Bosch (2014). A generic risk-based surveying method for invading plant pathogens. *Ecological Applications* **24**, pp. 779–790.

Parnell, S, TR Gottwald, NJ Cunniffe, V Alonso Chavez, and F van den Bosch (2015). Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20151478.

Parry, M, GJ Gibson, S Parnell, et al. (2014). Bayesian inference for an emerging arboreal epidemic in the presence of control. *Proceedings of the National Academy of Sciences* **111**, pp. 6258–6262.

Paul, C, N Hanley, ST Meyer, et al. (2020). On the functional relationship between biodiversity and economic value. *Science Advances* **6**, eaax7712.

Pautasso, M, T Harwood, M Shaw, X Xu, and M Jeger (2008). Epidemiological modeling of *Phytophthora ramorum*: network properties of susceptible plant genera movements in the nursery sector of England and Wales. In: *Proceedings of the Sudden Oak Death Third Science Symposium*. Vol. 214, pp. 257–264.

Perrings, C, C Castillo-Chavez, G Chowell, et al. (2014). Merging economics and epidemiology to improve the prediction and management of infectious disease. *EcoHealth* **11**, pp. 464–475.

Pimentel, D, R Zuniga, and D Morrison (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* **52**, pp. 273–288.

Plantegenest, M, C Le May, and F Fabre (2007). Landscape epidemiology of plant diseases. *Journal of the Royal Society Interface* **4**, pp. 963–972.

PlantMaps (2019). Available at: `https://www.plantmaps.com` (visited on June 4, 2019).

Pontryagin, LS, EF Mishchenko, VG Boltyanskii, and RV Gamkrelidze (1962). *The Mathematical Theory of Optimal Processes*. Translation by L.W. Neustadt. New York: Wiley.

Probert, WJM, K Shea, CJ Fonnesbeck, et al. (2016). Decision-making for foot-and-mouth disease control: objectives matter. *Epidemics* **15**, pp. 10–19.

Probert, WJM, S Lakkur, CJ Fonnesbeck, et al. (2019). Context matters: using reinforcement learning to develop human-readable, state-dependent outbreak response policies. *Philosophical Transactions of the Royal Society B* **374**, 20180277.

Qin, SJ and TA Badgwell (2003). A survey of industrial model predictive control technology. *Control Engineering Practice* **11**, pp. 733–764.

Ramage, BS and KL O'Hara (2010). Sudden oak death-induced tanoak mortality in coast redwood forests: current and predicted impacts to stand structure. *Forests* **1**, pp. 114–130.

Ramage, BS, KL O'Hara, and BT Caldwell (2010). The role of fire in the competitive dynamics of coast redwood forests. *Ecosphere* **1**, 20.

Ramage, BS, KL O'Hara, and AB Forrestel (2011). Forest transformation resulting from an exotic pathogen: regeneration and tanoak mortality in coast redwood stands affected by sudden oak death. *Canadian Journal of Forest Research* **41**, pp. 763–772.

Rao, AV (2009). A survey of numerical methods for optimal control. *Advances in the Astronautical Sciences* **135**, pp. 497–528.

Raue, A, C Kreutz, T Maiwald, et al. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, pp. 1923–1929.

Recht, B (2019). A tour of reinforcement learning: the view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems* **2**, pp. 253–279.

Rivers, C, JP Chretien, S Riley, et al. (2019). Using "outbreak science" to strengthen the use of models during epidemics. *Nature Communications* **10**, 3102.

Rizzo, DM and M Garbelotto (2003). Sudden oak death: endangering California and Oregon forest ecosystems. *Frontiers in Ecology and the Environment* **1**, pp. 197–204.

Rizzo, DM, M Garbelotto, and EM Hansen (2005). *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. *Annual Review of Phytopathology* **43**, pp. 309–335.

Robert, A, A Camacho, WJ Edmunds, et al. (2019). Control of Ebola virus disease outbreaks: comparison of health care worker-targeted and community vaccination strategies. *Epidemics* **27**, pp. 106–114.

Ross, N (2013). *SODDr*. `https://github.com/noamross/SODDr`. GitHub repository.

Rowthorn, RE, R Laxminarayan, and CA Gilligan (2009). Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface* **6**, pp. 1135–1144.

Roy, M and M Pascual (2006). On representing network heterogeneities in the incidence rate of simple epidemic models. *Ecological Complexity* **3**, pp. 80–90.

Sage, AP (1968). *Optimum Systems Control*. Englewood Cliffs, NJ: Prentice Hall.

Savary, S and L Willocquet (2014). Simulation modeling in botanical epidemiology and crop loss analysis. *APSnet Education Center. The Plant Health Instructor* **173**.

Savary, S, A Ficke, JN Aubertot, and C Hollier (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security* **4**, pp. 519–537.

Savary, S, L Willocquet, SJ Pethybridge, et al. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution* **3**, pp. 430–439.

Scheller, RM, JB Domingo, BR Sturtevant, et al. (2007). Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling* **201**, pp. 409–419.

Sélley, F, Á Besenyei, IZ Kiss, and PL Simon (2015). Dynamic control of modern, network-based epidemic models. *SIAM Journal on Applied Dynamical Systems* **14**, pp. 168–187.

Sethi, SP and PW Staats (1978). Optimal control of some simple deterministic epidemic models. *Journal of the Operational Research Society* **29**, pp. 129–136.

Sicard, A, AR Zeilinger, M Vanhove, et al. (2018). *Xylella fastidiosa*: insights into an emerging plant pathogen. *Annual Review of Phytopathology* **56**, pp. 181–202.

Silver, D, A Huang, CJ Maddison, et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484.

Simas, S de Ávila, MM Morato, DA Reynalte-Tataje, et al. (2019). Model-based predictive control for the regulation of the golden mussel *Limnoperna fortunei* (Dunker, 1857). *Ecological Modelling* **406**, pp. 84–97.

Skelsey, P, KA Garrett, et al. (2013). Pest and disease management: why we shouldn't go against the grain. *PLOS One* **8**, e75892.

Smith, DL, B Lucey, LA Waller, JE Childs, and LA Real (2002). Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences* **99**, pp. 3668–3672.

Stark, D, B Twieg, and Y Valachovic (2014). Sudden oak death in the Redwood Creek watershed and potential drought impacts. In: *Redwood Creek Symposium Proceedings*, p. 10.

Stenlid, J, J Oliva, JB Boberg, and AJ Hopkins (2011). Emerging diseases in European forest ecosystems and responses in society. *Forests* **2**, pp. 486–504.

Stephenson, B, C Lanzas, S Lenhart, and J Day (2017). Optimal control of vaccination rate in an epidemiological model of *Clostridium difficile* transmission. *Journal of Mathematical Biology* **75**, pp. 1693–1713.

Strange, RN and PR Scott (2005). Plant disease: a threat to global food security. *Annual Review of Phytopathology* **43**, pp. 83–116.

Stroud, PD, SJ Sydoriak, JM Riese, et al. (2006). Semi-empirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing. *Mathematical Biosciences* **203**, pp. 301–318.

Swiecki, TJ and EA Bernhardt (2013). *A reference manual for managing sudden oak death in California*. General Technical Report PSW-GTR-242. United States Department of Agriculture, Forest Service.

Tappeiner, JC, PM McDonald, and DF Roy (1990). *Lithocarpus densiflorus* (Hook. & Arn.) Rehd., tanoak. In: *Silvics of North America: Volume 2, Hardwoods*. Ed. by RM Burns and BH Honkala. Vol. 2. Washington, DC: United States Department of Agriculture, Forest Service, pp. 417–425.

Team Commands, Inria Saclay (2017). *BOCOP: an open source toolbox for optimal control*. Available at: `http://bocop.org` (visited on Aug. 30, 2017).

Thomas, PA (2016). Biological flora of the British isles: *Fraxinus excelsior*. *Journal of Ecology* **104**, pp. 1158–1209.

Thompson, ID, K Okabe, JM Tylianakis, et al. (2011). Forest biodiversity and the delivery of ecosystem goods and services: translating science into policy. *BioScience* **61**, pp. 972–981.

Thompson, RN and E Brooks-Pollock (2019). Preface to theme issue 'Modelling infectious disease outbreaks in humans, animals and plants: epidemic forecasting and control'. *Philosophical Transactions of the Royal Society B* **374**, 20190375.

Thompson, RN and WS Hart (2018). Effect of confusing symptoms and infectiousness on forecasting and control of Ebola outbreaks. *Clinical Infectious Diseases* **67**, pp. 1472–1474.

Thompson, RN, CA Gilligan, and NJ Cunniffe (2016). Detecting presymptomatic infection is necessary to forecast major epidemics in the earliest stages of infectious disease outbreaks. *PLOS Computational Biology* **12**, e1004836.

Thompson, RN, CA Gilligan, and NJ Cunniffe (2018). Control fast or control smart: when should invading pathogens be controlled? *PLOS Computational Biology* **14**, e1006014.

Tildesley, MJ, NJ Savill, DJ Shaw, et al. (2006). Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* **440**, pp. 83–86.

Tomlinson, I and C Potter (2010). 'Too little, too late'? Science, policy and Dutch elm disease in the UK. *Journal of Historical Geography* **36**, pp. 121–131.

Tonini, F, C Jones, BR Miranda, et al. (2018). Modeling epidemiological disturbances in LANDIS-II. *Ecography* **41**, pp. 2038–2044.

Twieg, B, Y Valachovic, R Cobb, and D Stark (2017). Reducing CO_2 emissions by managing for sudden oak death... is it possible? In: *Proceedings of the Sudden Oak Death Sixth Science Symposium*, pp. 19–22.

UK government (2018). *A Green Future: Our 25 Year Plan to Improve the Environment*. Annex 1: Supplementary Evidence Report. Available at: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/673492/25-year-environment-plan-annex1.pdf` (visited on Aug. 27, 2019).

Václavík, T, A Kanaskie, E Goheen, et al. (2010). Mapping the risk of sudden oak death in Oregon: prioritizing locations for early detection and eradication. In: *Proceedings of the Sudden Oak Death Fourth Science Symposium*. Vol. 229, pp. 126–132.

Valachovic, Y, L Quinn-Davidson, E Goldsworthy, and P Cannon (2013a). Novel approaches to SOD management in California wildlands: a case study of "eradication" and collaboration in Redwood Valley. In: *Proceedings of the Sudden Oak Death Fifth Science Symposium*, pp. 99–107.

Valachovic, Y (2006). *Humboldt County, 2006*. Available at: `http://www.suddenoakdeath.org/maps-media/photos/landscape-photos/` (visited on June 4, 2019).

Valachovic, Y, C Lee, J Marshall, and H Scanlon (2010). Forest treatment strategies for *Phytophthora ramorum*. In: *Proceedings of the Sudden Oak Death Fourth Science Symposium*, pp. 239–248.

Valachovic, Y, C Lee, B Twieg, et al. (2013b). Suppression of *Phytophthora ramorum* infestations through silvicultural treatment in California's north coast. In: *Proceedings of the Sudden Oak Death Fifth Science Symposium*, pp. 108–113.

Valachovic, Y, B Twieg, C Lee, RC Cobb, and D Stark (2017a). Forest stand conditions after *Phytophthora ramorum* management in northern California: post-treatment observations inform future responses. *Forest Phytophthoras* **7**, pp. 54–66.

Valachovic, Y, R Cobb, and B Twieg (2017b). How well has the spread of sudden oak death been predicted by the models in northern California? In: *Proceedings of the Sudden Oak Death Sixth Science Symposium*, pp. 11–12.

Von Stryk, O and R Bulirsch (1992). Direct and indirect methods for trajectory optimization. *Annals of Operations Research* **37**, pp. 357–373.

Wächter, A and LT Biegler (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* **106**, pp. 25–57.

Weber, TA and A Kryazhimskiy (2011). *Optimal Control Theory with Applications in Economics*. Cambridge, MA: MIT Press.

White, KAJ and CA Gilligan (2006). The role of initial inoculum on epidemic dynamics. *Journal of Theoretical Biology* **242**, pp. 670–682.

Whitehead, SJ and S Ali (2010). Health outcomes in economic evaluation: the QALY and utilities. *British Medical Bulletin* **96**, pp. 5–21.

Wikimedia Commons (2004). *Notholithocarpus densiflorus acorns*. Available at: `https://commons.wikimedia.org/wiki/File:Lithocarpus_densiflorus_acorns.jpg` (visited on June 4, 2019).

Yang, Y, SJ Titus, and S Huang (2003). Modeling individual tree mortality for white spruce in Alberta. *Ecological Modelling* **163**, pp. 209–222.

Zurakowski, R and AR Teel (2006). A model predictive control based scheduling method for HIV therapy. *Journal of Theoretical Biology* **238**, pp. 368–382.

Appendix to Chapter 4

A

A.1 Alternative fitting methods for network model

The best-fitting spatial deterministic model leads to disease progress curves in regions B and C that rise more quickly than the mean of the stochastic simulations, although the epidemic size within each region closely matches the simulations. This is because it is impossible for a deterministic model where the rates are fitted via maximum likelihood to adequately capture the dynamics of an epidemic in which the following three conditions are satisfied:

1. there is initially no infection;
2. there is a very small force of infection into the region (i.e. movement into the region is very restricted);
3. there are relatively fast within-region dynamics once disease has entered.

These conditions are true inside regions B and C and so we see reduced rate of spread in the stochastic case. This effect is due to stochastic fade outs after introduction events, as well as negative covariance between susceptible and infected hosts, leading to reduced infection rates in the simulations (Keeling and Rohani, 2008, pp. 227–229 and pp. 238–240). Figure A.1 demonstrates that this effect is seen in the simplest case of a metapopulation model with no risk structure. The rates are directly lifted from simulation to approximation to show the effect is purely due to the difference between deterministic and stochastic analogues. We also tested that our fitting procedures give these same values (data not shown).

The spatial approximate model could be extended and improved to account for these effects, for example by making use of moment closure techniques (Keeling, 1999) or nonlinear force of infection terms (Roy and Pascual, 2006) as used by Clarke et al. (2013) and Stroud et al. (2006), but we here focus on simplicity. Despite the limitations of the deterministic models, our proposed control frameworks allow the resulting controls to be used practically and successfully, particularly when approximating models are repeatedly reset. Since the benefits of the control frameworks should not depend on the exact fitting process used, we also fitted both approximate models by minimising the sum of squared

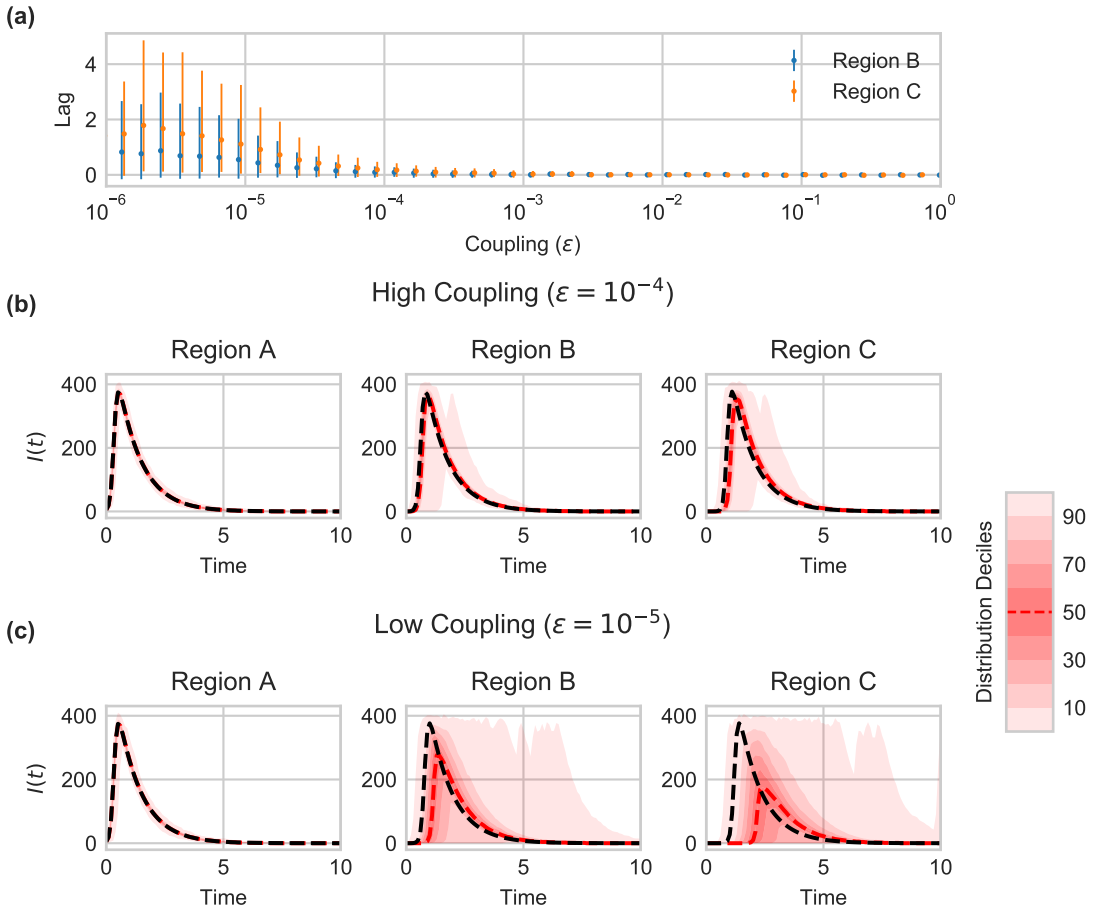

Figure A.1: Illustration of increased rate of spread in deterministic models when coupling is small. In this figure we use a 3 patch metapopulation model, with well-mixed dynamics within regions and coupling between regions A and B, and B and C. There is no risk structure. **(a)** shows the difference in peak infection time for deterministic and stochastic versions of the model with identical rates. Positive values indicate the deterministic model peaks earlier. Dots give the mean of 1000 stochastic simulations, with error bars showing the 95th percentile. **(b)** shows median stochastic dynamics and deciles in red, and deterministic disease progress curve in black, for a high coupling value (10^{-4}). The same is shown in **(c)** for a low coupling value (10^{-5})

errors from the simulation disease progress curves. For the risk-based model the following sum of squared errors was minimised:

$$\mathrm{SSE} = \sum_{i,j} \left(\Delta I_i^{\mathrm{H}}(t_j)\right)^2 + \left(\Delta I_i^{\mathrm{L}}(t_j)\right)^2 \tag{A.1}$$

where $\Delta I_i^{\mathrm{r}}(t_j)$ is the difference in the number of infected hosts in risk group r between the approximate model and simulation realisation i at time t_j. The sum is over all simulation realisations and across 51 times over the simulation time. The equivalent SSE function was used for the spatial approximate model, but summing differences for each region as well. Using this alternative fitting process did not change the ordering of control strategies as seen in Figure 4.10 in the main text (Figure A.2).

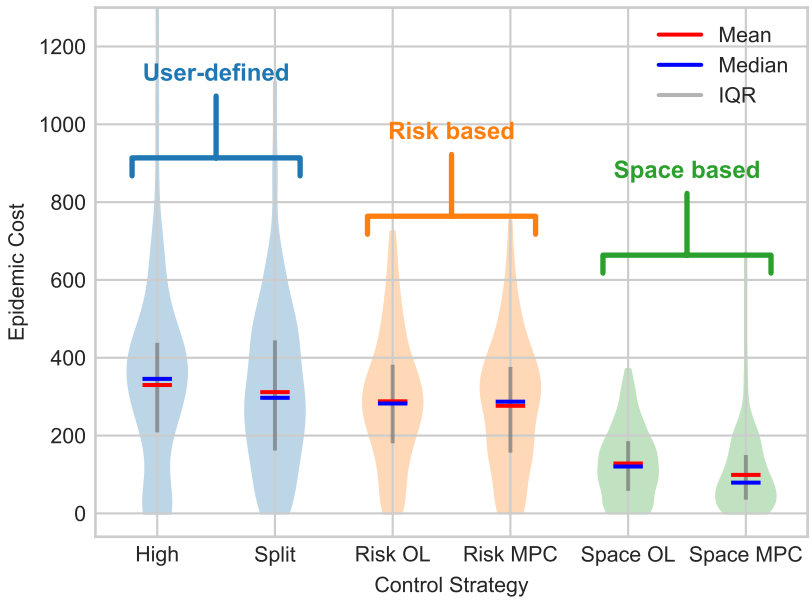

Figure A.2: Results of different control optimisation schemes on the illustrative simulation model using approximate models fitted by minimising SSE. Spatial MPC still performs best.

Appendix to Chapter 6

B

B.1 Parameter uncertainty scenarios

In Figure 6.12(c)–(d) in Chapter 6, reproduced for convenience in Figure B.1 below, 4 scenarios were highlighted to show the differences between open-loop and MPC under different parameter sets.

Scenario 1 Open-loop performs badly, MPC is significantly better.

Scenario 2 Average open-loop performance, moderate improvement by using MPC.

Scenario 3 Good performance using open-loop, marginal decrease in performance using MPC

Scenario 4 Good performance using open-loop, marginal increase in performance using MPC

Figures B.2–B.5 and the captions describe the situation in each scenario, explaining what drives the differences in performance.

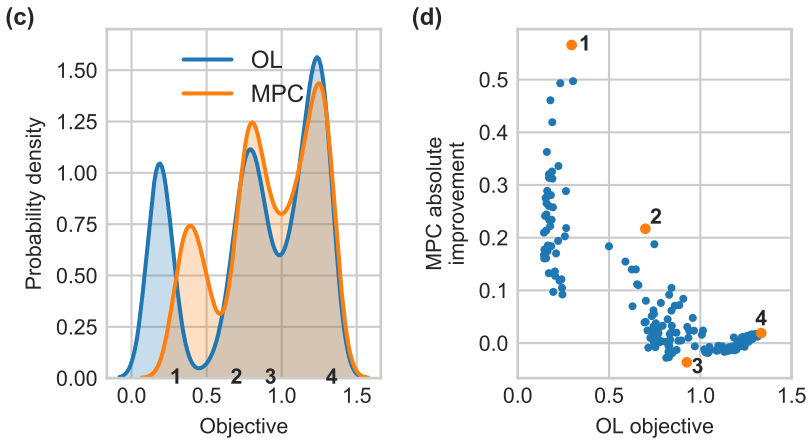

Figure B.1: Repeat of Figure 6.12(c)–(d) from Chapter 6. **(c)** shows the distribution of objective values using open-loop and MPC across 200 draws of simulation parameters. **(d)** shows the absolute improvement of the MPC strategy over open-loop, as a function of the open-loop objective.

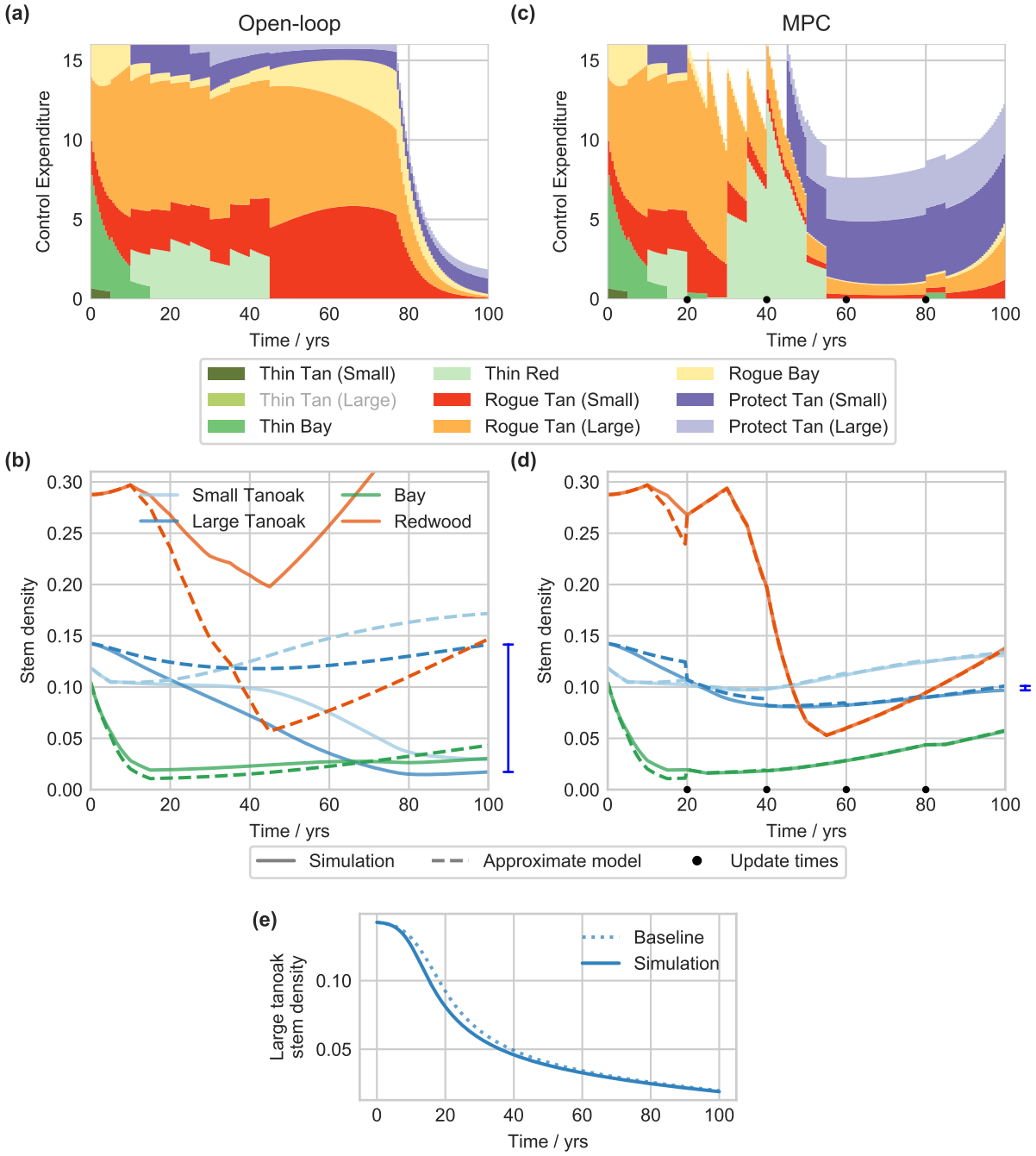

Figure B.2: Scenario 1. For each scenario, the open-loop control strategy and host dynamics are shown in **(a)** and **(b)**, and the same for MPC in **(c)** and **(d)**. The blue bars to the right of **(b)** and **(d)** highlight the difference between the simulation and approximate models in the number of large tanoak at the final time. **(e)** shows the large tanoak dynamics when there is no control compared with the baseline parameter case. Here, open-loop performs poorly because the disease spreads quickly, leading to significant tanoak decline in the first 20–40 years. **(e)** shows decline is faster than in the baseline case. In MPC, the framework can respond to this early decline and keep the disease under control.

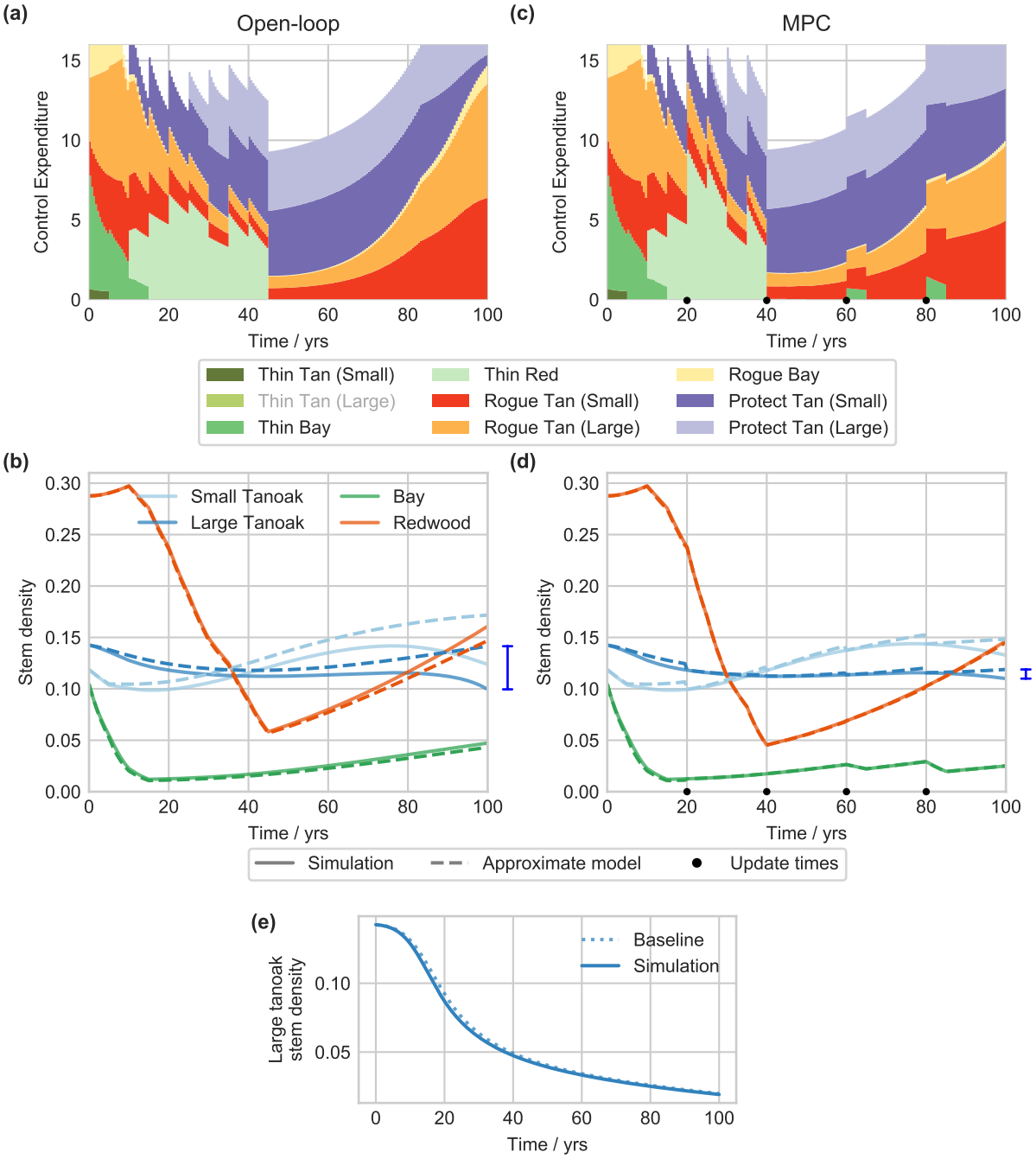

Figure B.3: Scenario 2. Here, under open-loop the approximate model slowly degrades and leads to differences between the simulation and approximate models. The control is relatively effective, but is not informed by the correct simulation state. Under MPC the approximate model is kept much closer to the simulation, leading to more informed control and better performance. **(e)** shows that tanoak decline under no control is similar to the baseline case.

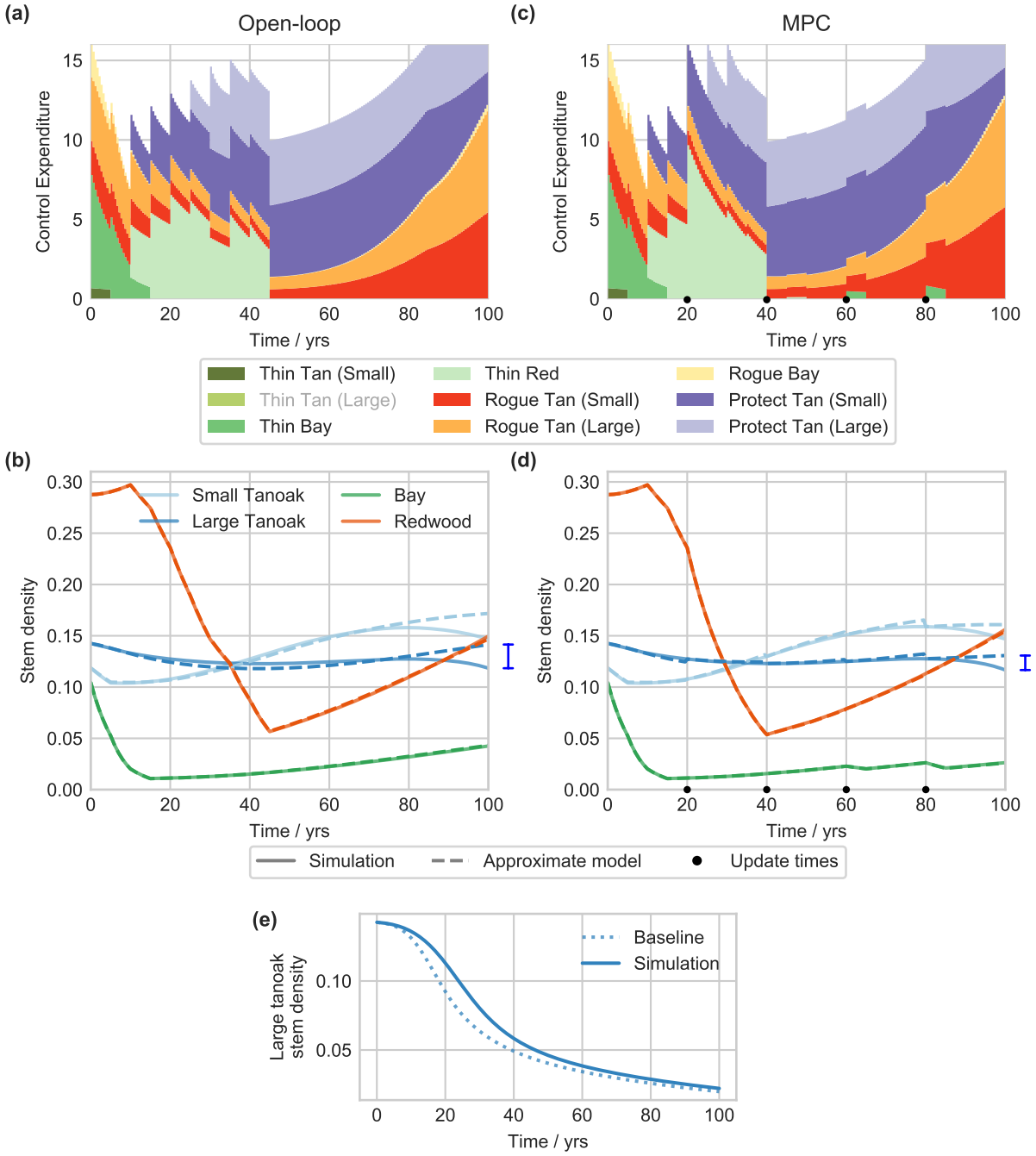

Figure B.4: Scenario 3. The disease is slow to spread, and therefore relatively easy to control. The approximate model stays close to the simulation under both open-loop and MPC as there are only small amounts of disease spread. The different thinning regime under MPC leads to slightly worse retention of tanoak than under open-loop, but the difference is very small. **(e)** shows that tanoak decline under no control is slower than in the baseline case.

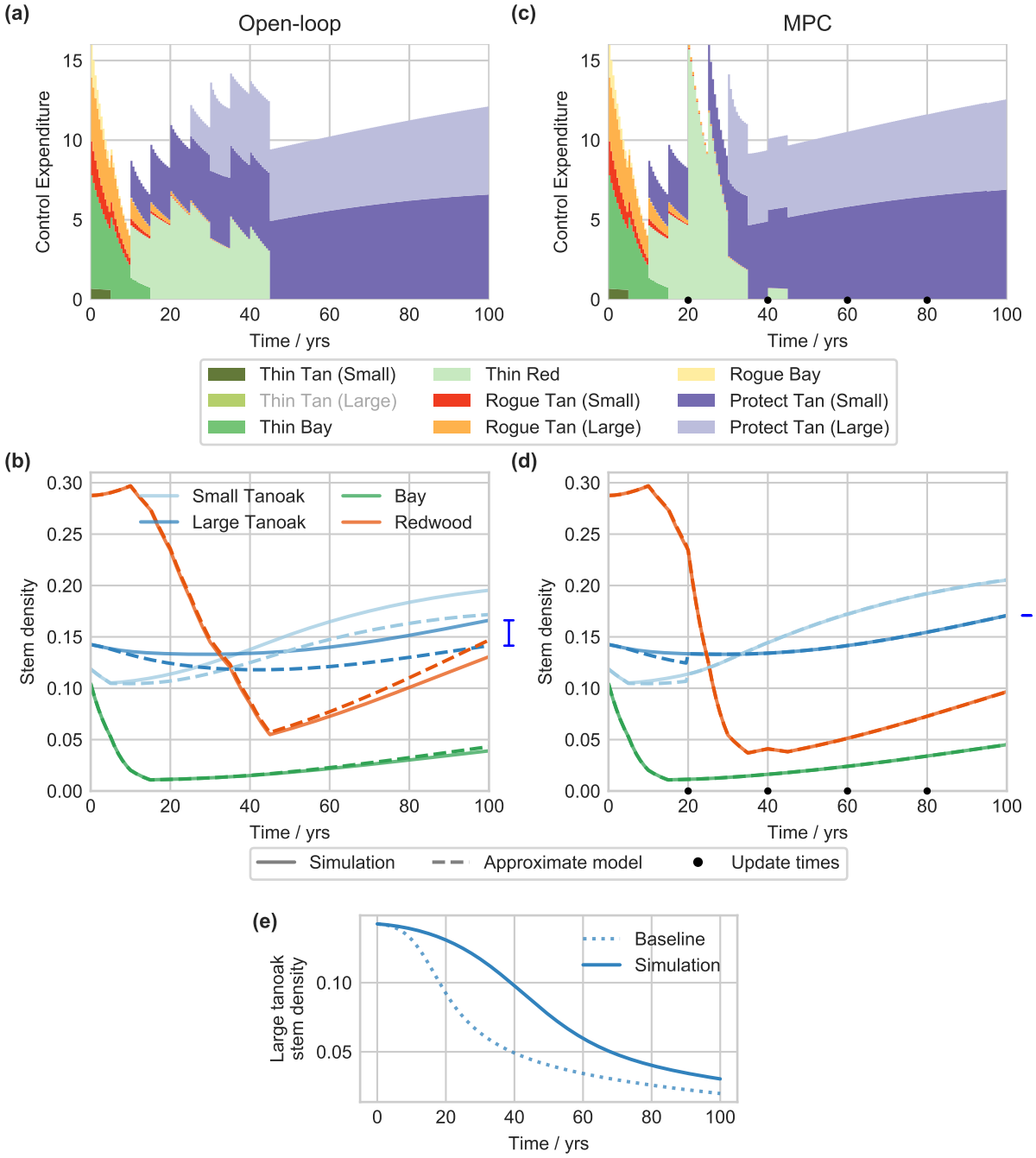

Figure B.5: Scenario 4. The disease is very easy to control, leading to minimal roguing under both frameworks. The thinning of redwood under MPC is better informed after the update time at 20 years, and so promotes additional recovery of tanoak. Here both frameworks increase the size of the tanoak population above the pre-disease introduction level. **(e)** shows that tanoak decline under no control is much slower than in the baseline case.

B.2 Efficacy of protectant methods

In Section 6.4.1 we stated that the protectant methods are unlikely to be applied in practice since they have a very small effect on the overall objective. We here verify this by running the MPC strategy with and without the protectant methods, as shown in Figure B.6. The protectant application marginally increases the objective function, but the effect is negligible.

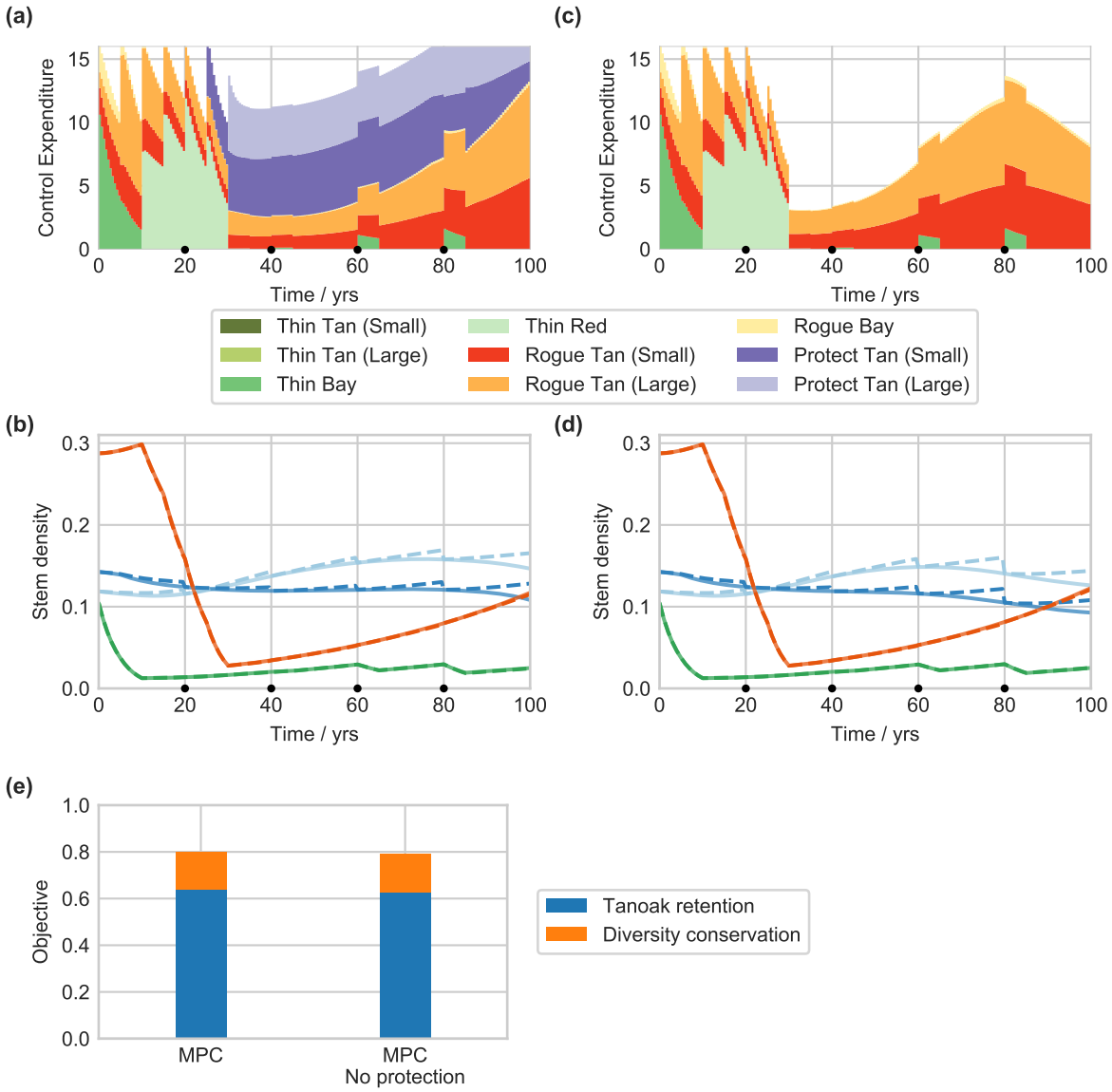

Figure B.6: Under the MPC framework, applying protectant methods does not have a significant effect on the control performance. **(a)** and **(b)** show the control and host dynamics under the standard MPC framework. **(c)** and **(d)** show the same but where the protectant method is not applied. Note that control expenditure, in particular for roguing, is not the same since the number of hosts has changed. **(e)** shows the objective function for MPC with and without protectant methods. The total objective values are 0.7996 (4 s.f.) and 0.7938 (4 s.f.) with and without protectant application respectively.

B.3 Time horizon robustness

The 100 year time horizon was chosen to be long enough to capture tanoak decline and show the differences between the open-loop and MPC frameworks. We here extend the time horizon by 5 MPC update periods to 200 years to verify that the difference between the frameworks is robust in the longer term. This longer time horizon is arbitrary but chosen as the disease has been eradicated by this time. We tested other end points to ensure the conclusions are consistent (data not shown). Figure B.7 shows the control strategies and host dynamics for both open-loop and MPC over this longer time horizon of 200 years. The MPC framework does show tanoak decline from late re-emergence, but this is delayed compared to open-loop. After 200 years, the disease has been eradicated and tanoak populations are recovering under both open-loop and MPC. However, the retained tanoak population is twice as large under the MPC framework.

We also test how robust the control strategies are to shorter time horizons. While a time horizon of 100 years is appropriate for considering tanoak restoration and biodiversity effects, decision makers may not optimise strategies over such a long time period. We optimised control over a time horizon of 20 years, as shown in Figure B.8. As with the longer time horizon, control still initially focusses on thinning of bay laurel. The control strategy looks very similar to the first 20 years of the optimal control using the 100 year time horizon, with a slight delay to thinning of redwood. This is because over the short term making space for tanoak regrowth is less important than managing infection levels.

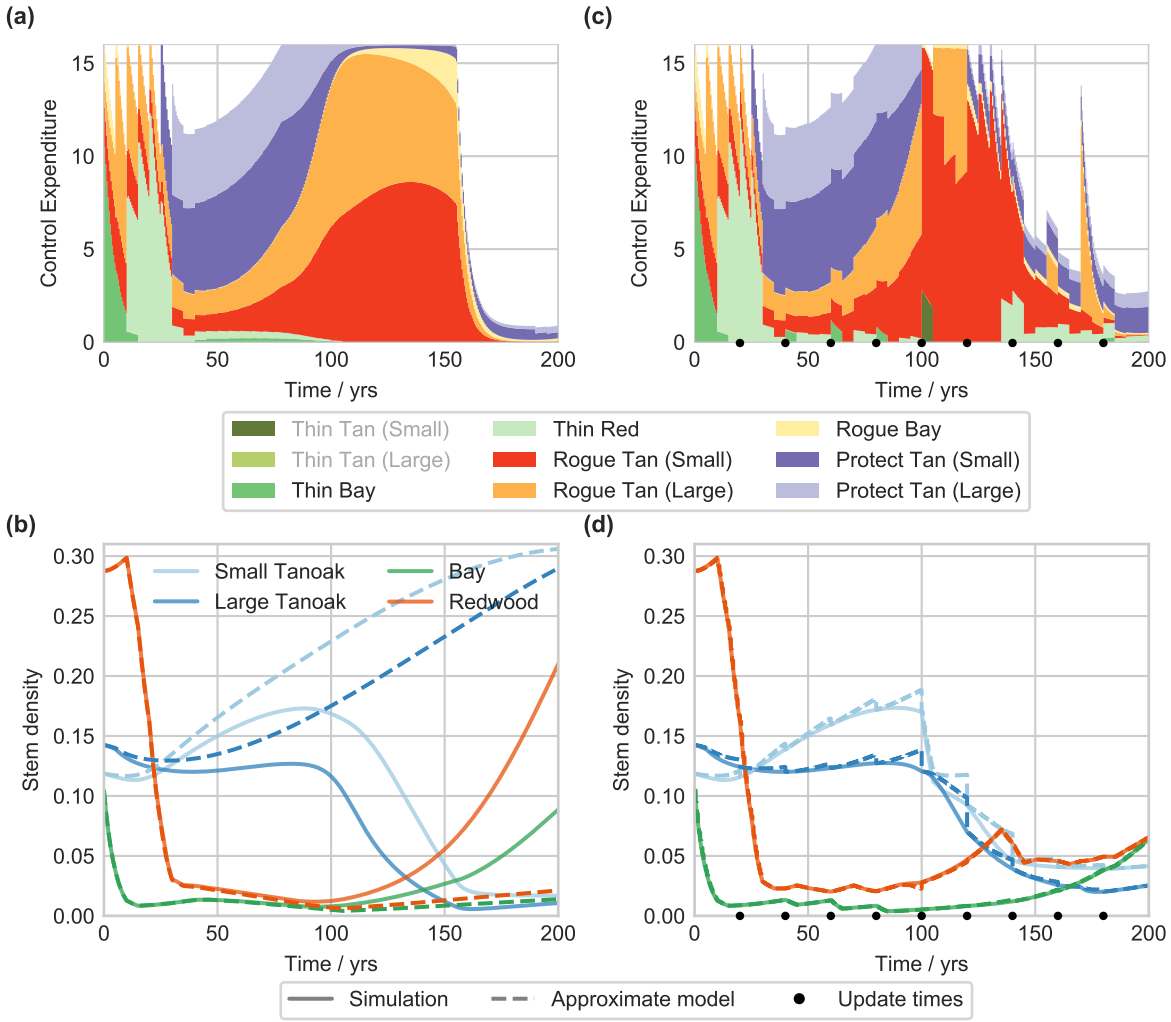

Figure B.7: Open-loop and MPC control strategies, and host dynamics, using a time horizon of 200 years. **(a)** and **(b)** are the control strategy and host dynamics for open-loop, **(c)** and **(d)** show the same for MPC. MPC slows down the tanoak decline seen using the open-loop framework and retains more tanoak.

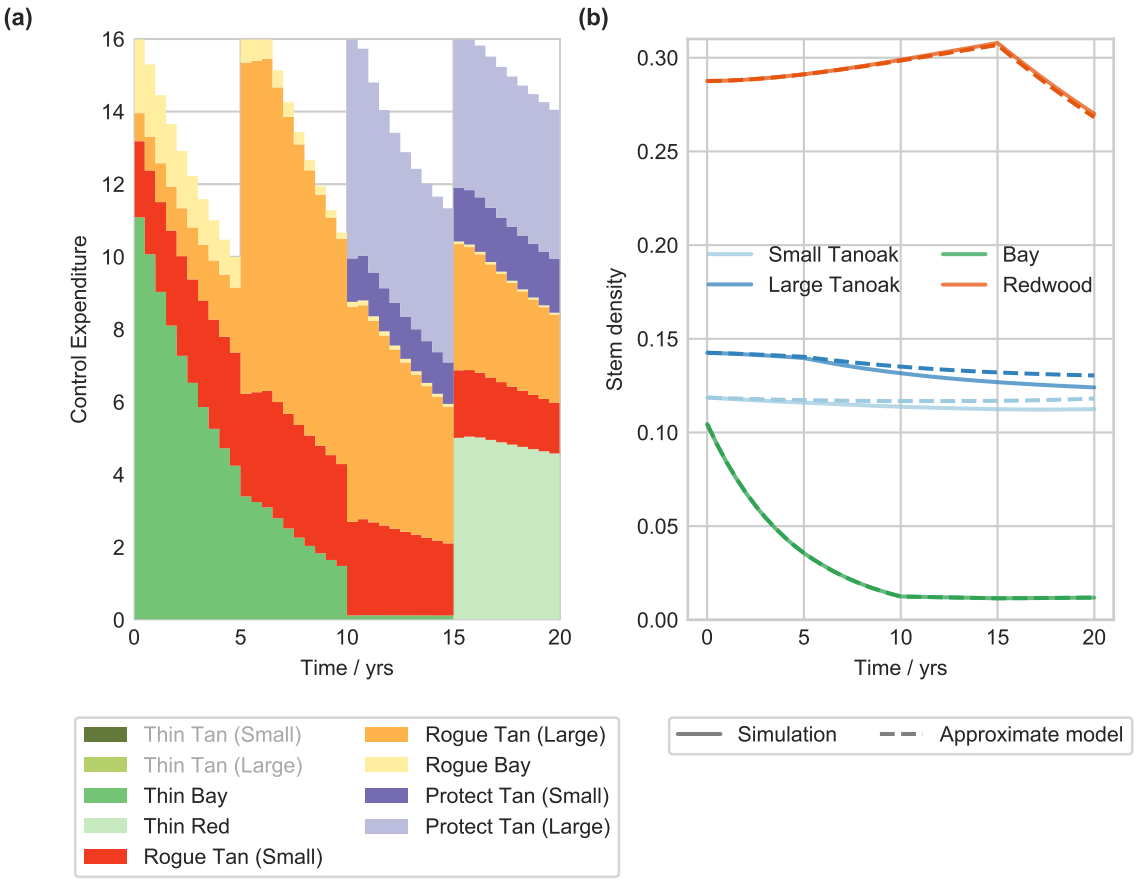

Figure B.8: Open-loop control strategy and host dynamics using a time horizon of 20 years. The optimal strategy over a short time horizon is similar to 100 year time horizon results with a delay to thinning of redwood.

Appendix to Chapter 7

C

C.1 Parameter and variable values

Table C.1: Table of parameter and variable meanings and default values.

Symbol	Meaning	Default value
β	Maximum spore production rate from each infected host unit	4.55 wk^{-1}
γ	Proportion of dispersal events distributed using short range kernel	0.9947
α_1	Scale parameter for short range kernel	20.57 m
α_2	Scale parameter for long range kernel	9.504 km
η	Maximum total control rate	50 km^2 $year^{-1}$
$\chi(f_i(t))$	Indicator for pathogen conduciveness in cell i of forest type f_i at time t	0 or 1
m_{it}	Pathogen conduciveness due to moisture for cell i at time t	$[0,1]$
c_{it}	Pathogen conduciveness due to temperature for cell i at time t	$[0,1]$
M_i	Averaged pathogen conduciveness due to weather and forest type mask for cell i	$[0,1]$
$N_{\max}$	Maximum number of host units in any single simulation cell	100
S_{it}	Number of susceptible host units in simulation cell i at time t	$[0,N_{\max}]$
I_{it}	Number of infected host units in simulation cell i at time t	$[0,N_{\max}]$
K_{ji}	Probability of spore dispersal from cell i to cell j	$[0,1]$
$\tilde{\beta}$	Fitted infection rate in approximate model	-
$\tilde{S}_i$	Number of susceptible hosts in aggregated cell i in approximate model	-
$\tilde{I}_i$	Number of infected hosts in aggregated cell i in approximate model	-
k_{ij}	Dispersal kernel between cells i and j in approximate model	$[0,1]$
$u_i(t)$	Thinning intensity at time t in cell i of the approximate model	$[0,1]$
$v_i(t)$	Roguing intensity at time t in cell i of the approximate model	$[0,1]$

C.2 Approximate model fitting results

Table C.2: Table of approximate model fitted parameters. All values are give to 3 significant figures.

	Exponential kernel		Cauchy kernel	
Resolution (m)	$\tilde{\beta}_1$ (host^{-1} t^{-1})	σ_1 (m)	$\tilde{\beta}_2$ (host^{-1} t^{-1})	σ_2 (m)
250	0.617	1090	0.0787	41.0
500	0.156	1100	0.022	91.6
1000	0.0392	1060	0.00626	189
1500	0.0164	1110	0.003	299
2000	0.00856	1220	0.00175	404
2500	0.00506	1310	0.00116	506
5000	0.000944	1830	0.000383	1420

C.3 Problem formulation

Here we show how the non-linear programming problem (NLP) was formulated for optimising spatial control to protect Redwood National Park in Chapter 7. The raster based disease system is described by 2 ODEs for each cell in the landscape:

$$\dot{S}_i = -\tilde{\beta} M_i S_i \sum_j (k_{ij} M_j I_j) - u_i(t) \eta S_i \tag{C.1a}$$

$$\dot{I}_i = \tilde{\beta} M_i S_i \sum_j (k_{ij} M_j I_j) - v_i(t) \eta I_i \tag{C.1b}$$

where the variables are as described in Chapter 7. As explained in Chapter 2, the direct method discretises the system of ODEs and optimises these variables subject to constraints that impose the correct dynamics. In our case, for a system of N cells discretised on a grid of $M+1$ time points, the vector of NLP variables to optimise is given by:

$$\begin{aligned} \mathbf{y}^{\mathrm{T}} = (&S_0^0, I_0^0, u_0^0, v_0^0, \dots S_N^0, I_N^0, u_N^0, v_N^0, \\ &S_0^1, I_0^1, u_0^1, v_0^1, \dots S_N^1, I_N^1, u_N^1, v_N^1, \\ &\dots S_N^M, I_N^M, u_N^M, v_N^M) \end{aligned} \tag{C.2}$$

where subscripts indicate the cell and superscripts indicate the time point.

Constraints must now be formulated to impose the dynamics of the ODE system on these NLP variables. Because of the size of the NLP problem, we use a simple midpoint

discretisation method. For an ODE system $\dot{x} = f(t, x(t))$, the implicit midpoint method is given by:

$$x_{n+1} = x_n + hf\left(t_n + \frac{h}{2}, \frac{1}{2}(x_n + x_{n+1})\right) \tag{C.3}$$

where h is the step size, such that $t_n = t_0 + nh$, and x_n is the approximation of $x(t_n)$ (Betts, 2010, p. 100).

For our system, applying the midpoint method results in the following state dynamics constraints:

$$\begin{aligned} S_i^{n+1} = S_i^n - h\Big[\tilde{\beta}\frac{1}{2}\left(S_i^n + S_i^{n+1}\right)\sum_j M_i M_j k_{ij}\frac{1}{2}\left(I_j^n + I_j^{n+1}\right) \\ + \eta\frac{1}{2}\left(u_i^n + u_i^{n+1}\right)\frac{1}{2}\left(S_i^n + S_i^{n+1}\right)\Big] \end{aligned} \tag{C.4a}$$

$$\begin{aligned} I_i^{n+1} = I_i^n + h\Big[\tilde{\beta}\frac{1}{2}\left(S_i^n + S_i^{n+1}\right)\sum_j M_i M_j k_{ij}\frac{1}{2}\left(I_j^n + I_j^{n+1}\right) \\ - \eta\frac{1}{2}\left(v_i^n + v_i^{n+1}\right)\frac{1}{2}\left(I_i^n + I_i^{n+1}\right)\Big]. \end{aligned} \tag{C.4b}$$

Alongside the initial condition constraints, and bounds on the control variables, these form the full constraint system for the NLP problem. These constraints are differentiated with respect to the NLP variables to give the Jacobian of the problem, and differentiated again to form the Hessian.

C.4 Problem optimisation

The NLP variables are optimised to maximise the number of healthy hosts at the final time, with a varying weight for each cell. In terms of the NLP variables, this objective function is given by:

$$J = \sum_i W_i S_i^M \tag{C.5}$$

where the weight for cell i is given by W_i, and time point M is at the final time.

The problem is optimised using Ipopt (Wächter and Biegler, 2006), an open source interior point optimiser designed for large-scale non-linear optimisation. The constraint system, objective function, and their derivatives are implemented in C++ as an interface to the Ipopt software.